Azure status history

This page contains all root cause analyses (RCAs) for incidents that occurred on November 20, 2019 or later. Each RCA will be retained on this page for 5 years. RCAs before November 20, 2019 aren't available.

Product: All | Region: All | Date: All

September 2020

9/3 RCA - Network Latency Issue – West Europe (Tracking ID 8KLC-1T8)

Summary of Impact: Between 09:21 and 17:32 UTC on 03 Sep 2020, a subset of customers may have experienced intermittent latency or issues connecting to resources hosted in West Europe. Retries may have worked during this timeframe.

Root Cause: Two separate events occurred in close succession prior to the start of impact from this incident:

- Approximately 4 hours before the impact start, some local activity (likely construction) in the vicinity of the data centre cause an increase in the number of packets corrupted during transmission over fiber optic cables between data centers in the West Europe region. These errored packets were detected and dropped, and our networking automation systems took the links out of service and opened tickets with the local site to have them repaired. This is a standard process, and our automated safety checks validated that there was no impact related to this.
- Separately, between 09:21 and 09:26 UTC a significant fiber-cut occurred approximately 5 kilometres from the data centre on one one of the other paths between the data centers. This cut impacted 50% of the capacity for that route, but again, this event on its own would have no impact on traffic overall in the West Europe region.

Each of the events in isolation would have had no perceptible impact on the networking operations for West Europe, but when combined, they resulted in 9 links between data centres receiving an unequal share of traffic, becoming congested, and dropping packets (the impact was to less than 2% of the total capacity on the impacted links). Connections that travelled over these congested links would have experienced increased packet loss and latency. As connections are spread over the available links, services that retried requests by opening new connections were likely to have been unaffected and successful.

The time to mitigate was extended by the need for on-call engineers to identify that there were multiple causes for down links and identify the best way to reduce congestion and rebalance traffic. During the initial response, the large number of concurrent alerts resulted in on-call engineers taking actions that moved the congestion from one link to another, but did not resolve it.

Mitigation: Mitigation was achieved by engineers manually determining which of the links that had experienced errors could be put back into service and rebalancing traffic across the links in service. Full mitigation was declared at 17:32 UTC, but most customers would have see improvement in advance of this time. Full restoration was achieved by September 4 02:00 UTC when the significant fiber cut was repaired.

Next Steps: We sincerely apologize for the impact to affected customers. We are continuously taking steps to improve the Microsoft Azure Platform and our processes to help ensure such incidents do not occur in the future. In this case, this includes (but is not limited to):

- Accelerate the readiness of additional fiber paths between these data centres to reduce the impact of future fiber cuts.
- Improve the tooling used by on-call engineers when responding to complex incidents with multiple causes of downed links, so that they can reduce congestion faster and achieve mitigation more quickly.

Provide Feedback: Please help us improve the Azure customer communications experience by taking our survey: https://aka.ms/AzurePIRSurvey

August 2020

8/21 Content Delivery Network (CDN) - Service Degradation - Mitigated (Tracking ID DLYY-ND8)

Summary of Impact: Between 18:05 and 19:55 UTC on 21 Aug 2020, a subset of customers using Azure CDN from Verizon may have experienced service degradation.

Preliminary Root Cause: We determined that a recent deployment task impacted connectivity to origins, causing dynamic or cache miss requests to fail.

Mitigation: The CDN provider rolled out an update that fixed the issue.

Next Steps: We will continue to investigate to establish the full root cause and prevent future occurrences. Stay informed about Azure service issues by creating custom service health alerts: https://aka.ms/ash-videos for video tutorials and https://aka.ms/ash-alerts for how-to documentation.

8/14 RCA - Degraded connectivity to Microsoft Services within the Southeast region of the United States (Tracking ID 9MDM-TT8)

Summary of Impact: Between approximately 02:20 UTC to 03:30 UTC and 04:07 UTC to 04:52 UTC on 14 Aug 2020, a subset of customers connecting through one of Microsoft's edge-nodes (https://aka.ms/MSGlobalNetwork) in the Southeast United States (US) may have experienced intermittent periods of degraded connectivity when attempting to connect to Azure, Microsoft 365, and Xbox resources.

Root Cause: Microsoft's Global Network consists of edge-nodes that connect to the Internet externally and two or more Backbone sites internally via diverse optical fiber paths for redundancy during failure scenarios.

On 14 Aug 2020 at 02:20 UTC, we experienced a dual fiber path failure isolating one of our edge-nodes in the Southeastern US. The initial fiber path incident occurred on 13 Aug 2020 at 18:34 UTC due to a fiber cut causing the path to be removed from Microsoft's Global Network. Traffic was then routed to our secondary fiber path per design. Meanwhile, our fiber provider had dispatched a technician to work on resolving the initial fiber incident. While working on that incident, the technician inadvertently disconnected our secondary fiber path at 02:20 UTC, which resulted in the secondary path to be removed from Microsoft's Global Network isolating this edge-node site.

Our network is designed to withstand site isolation and all traffic should have rerouted to the next closest edge-node in the region. However, we identified a router in this edge-node site that continued to advertise a few local prefixes to the Internet, which resulted in the blackholing of all Internet traffic destined to those prefixes in the edge-node site. The route advertisement of the local prefixes should have been withdrawn by the router when the site was isolated from Microsoft Global Network during the secondary fiber path incident but that did not occur due to a missing configuration at this site to detect site isolation and resulted in an outage. In addition, customer notification of the event was delayed due to correlation of the event and the impact.

Mitigation: The outage was mitigated when the fiber provider technician completely restored the fiber connectivity at 04:52 UTC on 14 Aug 2020.

Next Steps: We sincerely apologize for the impact to affected customers. We are continuously taking steps to improve the Microsoft Azure platform and our processes to help ensure such incidents do not occur in the future. In this case, this includes (but is not limited to):

- Taking steps to prevent dual failures from occurring, reduce the degree of impact, and shorten time-to-mitigate by implementing improved failover operations to backup sites.
- Modifying our router configurations globally, to implement conditional prefix advertisement and withdrawal to ensure routers disconnect as expected during isolation events.
- Improving our alert correlation to notify fiber technicians in a timely manner, and to improve the overall notification experience.

Provide Feedback: Please help us improve the Azure customer communications experience by taking our survey: https://aka.ms/AzurePIRSurvey

July 2020

7/18 RCA - Azure DNS - Connectivity issues (Tracking ID TTPY-3P0)

Summary of Impact: Between 07:50 and 08:45 UTC (approx.) on 18 Jul 2020, Azure DNS experienced a transient resolution issue which in-turn impacted connectivity for some other Azure services. Authoritative and other DNS services were not impacted by this issue.

Root Cause: The decommissioning of a legacy (preview) DNS solution inadvertently caused some data streams for Azure DNS recursive resolver service to become out of sync with the resolver state. This was detected by a sync pipeline, which triggered a reset of the resolver instances to recover from the stale state. Unfortunately, this reset was not done in a staggered fashion and led to multiple resolver instances rebooting at the same time. This in turn led to degradation of the service and caused DNS resolution failures for the queries originating from virtual networks. Azure services dependent on the Azure DNS resolver service also saw degradation of service during this time.

The impact of the incident was observed across multiple Azure regions to varying degrees. While some instances of the service saw no impact, most impacted instances auto-recovered within 10 minutes, though some instances took up to 30 minutes to recover. The DNS resolution issues were fully auto-mitigated across all regions within 54 minutes. During this time, authoritative Azure DNS service was not impacted and DNS queries originating from the internet for zones hosted on Azure DNS were answered successfully.

Mitigation: The issue was self-healed as the restarts completed, and all services with dependencies on the recursive DNS service would have seen a restoration of functionality also. The DNS service was fully mitigated at 08:45 UTC, but some services with multiple dependencies may have taken longer for all customers to see full service restoration.

Next Steps: We sincerely apologize for the impact to affected customers. We are continuously taking steps to improve the Microsoft Azure platform and our processes to help ensure such incidents do not occur in the future. In this case, this includes (but is not limited to):

- Fixing the orchestration logic in the sync pipeline to help ensure that resolver instances are reset in a staggered, partitioned fashion
- Improving the resolver startup sequence to help ensure that a resolver instance can be up and running with 10 minutes after a reset

Provide Feedback: Please help us improve the Azure customer communications experience by taking our survey: https://aka.ms/AzurePIRSurvey

7/7 RCA - Virtual Machines - Virtual machine unexpected restarts (Tracking ID 8S8J-9T8)

Summary of Impact: Between 07:24 UTC on 07 Jul 2020 and 21:16 UTC on 17 Jul 2020, a subset of customers using Virtual Machines (VMs) may have experienced intermittent connection failures when trying to access some virtual machines. These virtual machines may have also restarted unexpectedly.

Root cause: We determined that an ongoing OS update deployment task inadvertently contained a code configuration error that resulted in a number of previously addressed bug fixes being reverted on a subset of clusters. This manifested as system deadlock on a subset of host nodes which were running VM workloads with heavy disk I/O. As a result, VMs on those nodes rebooted.

Mitigation: We stopped the ongoing deployment and subsequently developed and rolled out a new deployment task which contained a code fix to detect that a new patch needed to be applied. This fix was deployed to all impacted clusters, thereby mitigating the VM reboots and customer impact.

In parallel to deploying the permanent fix across all regions, we expedited mitigation for some customers by identifying affected nodes that were hosting the customers' VM workloads and reattaching patches to those nodes.

Next Steps: We understand that the time to mitigate for this incident was longer than desired, and we sincerely apologize for the impact to affected customers. We are continuously taking steps to improve the Microsoft Azure Platform and our processes to help reduce the duration of such incidents. This includes (but is not limited to):

- Incorporating the missed combination of software versions and system configurations in our validation matrix before deploying similar updates.
- Improving rollout monitoring and correlation capabilities to detect such faults and halt the rollout sooner.

Provide Feedback: Please help us improve the Azure customer communications experience by taking our survey: https://aka.ms/AzurePIRSurvey

7/4 Azure DevOps - Service Outage - Mitigated (Tracking ID SNDL-NS8)

Summary of Impact: Between 02:26 am and 03:40 am UTC on 04 Jul 2020, customers using Azure DevOps in multiple regions may have observed connectivity errors to DevOps services.

Preliminary Root Cause: We identified an inadvertent error with a configuration change in the back-end service which caused the outage.

Mitigation: We applied a configuration update which has fully mitigated the issue.

Next Steps: We will continue investigations to establish the full root cause and prevent future occurrences. Stay informed about Azure service issues by creating custom service health alerts: https://aka.ms/ash-videos for video tutorials and https://aka.ms/ash-alerts for how-to documentation.

7/1 RCA - Azure SQL Database - Japan East (Tracking ID CLCK-LD0)

Summary of Impact: Between 09:24 and 11:15 UTC on 01 Jul 2020, a subset of customers using Azure SQL Database, Azure SQL Data Warehouse/Synapse Analytics, Azure Database for MySQL, Azure Database for PostgreSQL, and Azure Database for MariaDB in Japan East may have experienced service connection failures or possible timeouts. Services utilizing SQL Databases may have also been impacted.

Root Cause: Connections to Azure SQL Database and related data services go through a load balanced set of front-end nodes (Gateways) that provide directory lookup services and reroute the incoming connections to the intended backend nodes hosting the database. For scalability and zone redundancy purposes, there are multiple active SQL Gateway clusters in a region. During this incident, one of the SQL Gateway clusters became unhealthy, having an intermittent impact on login availability. A specific network traffic pattern combined with a networking stack configuration on the SQL Gateway instances triggered an imbalance on the CPU processing of new connection requests. The persistence of such CPU imbalance over a long period of time-caused high response latency and increased timeouts on connection requests. The error condition propagated across multiple instances of the SQL Gateway cluster in this region, sometimes causing a service restart.

Mitigation: Multiple SQL Gateway instances became healthy upon the triggered service restart. On further investigation, we were able to isolate the specific network pattern and the configuration setting that caused this incident and were able to reconfigure the traffic to prevent a recurrence.

Next Steps: We apologize for the impact to affected customers. We are continuously taking steps to improve the Microsoft Azure platform and our processes to help ensure such incidents do not occur in the future. In this case, this includes (but is not limited to):

- Fix the underlying issue that causes service restart when such a condition occurs.
- Improve the alerting logic and add identified telemetry to diagnose this kind of issues faster.
- Activate a newer SQL Gateway cluster in this region with a more efficient networking stack configuration that reduces the chances of hitting a processing imbalance.

Provide Feedback: Please help us improve the Azure customer communications experience by taking our survey: https://aka.ms/AzurePIRSurvey

June 2020

6/14 RCA - Azure Active Directory - Authentication Errors (Tracking ID PMHH-NS0)

Summary of Impact: Between 23:00 UTC on 14 Jun 2020 and 01:40 UTC on 15 Jun 2020, a subset of customers using Azure Active Directory may have experienced authentication issues when accessing resources. Customers may have received the following error message "AADSTS90033: A Transient error has occurred. Please try again."

Root Cause: An unexpected increase in traffic volume and resource utilization of infrastructure in the region responsible for acquiring authentication tokens resulted in regional contention which exceeded operational thresholds; resulting in authentication issues for a subset of customers.

Mitigation: The backend infrastructure was scaled out to increase resources and traffic was redistributed.

Next Steps: We are continuously taking steps to improve the Microsoft Azure Platform and our processes to help ensure such incidents do not occur in the future. This includes, and is not limited to:

- Improvements to our systems that dynamically scale resources in this scenario accordingly with corresponding monitoring
- Improvements to monitoring to better detect increases in traffic for similar scenarios
- Review and enhance procedures to improve customer communication experience

Provide Feedback: Please help us improve the Azure customer communications experience by taking our survey: https://aka.ms/AzurePIRSurvey

6/11 RCA - Storage - East US (Tracking ID 9VHK-J80)

Summary of Impact: Between 11:57 and 14:20 UTC on 11 Jun 2020, a subset of Storage customers in East US may have experienced connection failures when trying to access some of their resources hosted in this region. Services with dependencies on the impacted storage resources, such as Virtual Machines, may also have experienced downstream impact during this time.

Root Cause: Engineers determined that an incident during a planned power maintenance activity at the datacenter caused an impact to a single storage scale unit, which then became unhealthy. The incident caused power to be lost to to a subset of racks comprising 60% of this single storage scale unit.

The maintenance activity itself did not impact the storage scale unit, but it caused the scale unit to have reduced redundant power options at the time of the incident. All racks and network devices have two sources of power for redundancy, but it is standard procedure in some types of maintenance to isolate some resources to a single source of power for a short period. After the isolation had been completed on this scale unit, but before maintenance could begin, a distribution breaker in the redundant power source tripped open unexpectedly and the power was lost to the subset of racks.

Mitigation: The site engineers paused all maintenance work and inspected the electrical distribution system to ensure there were no apparent equipment fault issues. They found the tripped breaker and determined it had failed. Power was restored by closing the other breaker that had previously been opened to commence the isolation for the scale unit, and this restored a single power source to the impacted racks. A new breaker was located and fully-tested before installation. The bad breaker on the redundant power supply was replaced with the new breaker and redundant power was then also restored to the affected racks.

Once power was restored to the impacted storage racks, the automated restart process for storage resources began, and restored the scale unit to full operation. The restart process for storage clusters follows a series of structured steps to ensure full integrity of customers' data is preserved, and access to storage resources on this scale unit would have become available over a short period of time. Final mitigation was declared at 14:20, but most customers would have seen recovery prior to this time.

Subsequent testing showed that the breaker had an internal failure on one phase and it has been sent to the manufacturer for full forensic analysis.

Next Steps: We sincerely apologize for the impact to affected customers. We are continuously taking steps to improve the Microsoft Azure Platform and our processes to help ensure such incidents do not occur in the future. In this case, this includes (but is not limited to):

- Forensic investigation into the breaker that failed.
- Review of breaker maintenance and testing requirements
- Improving monitoring and alerting visibility when in maintenance modes.
- Ensuring that electrical maintenance activities associated power switching are monitored closely with automated processes to manage unintended impact.

Provide Feedback: Please help us improve the Azure customer communications experience by taking our survey: https://aka.ms/AzurePIRSurvey

6/4 RCA - Azure Resource Manager - Failures creating or deleting resources (Tracking ID DLZG-7C0)

Summary of Impact: Between 07:45 and 16:57 UTC on 04 Jun 2020, a subset of customers across all Public Azure regions may have experienced deployment failures when attempting to create or delete certain service based resources via Azure Resource Manager (ARM) deployment and management service due to an underlying networking issue. While the related networking resources for the impacted services were actually being created or deleted during this time, ARM was not notified of the deployment status and hence was failing the service creation or deletion. This issue may have impacted some GET or READ action on the resources. Less than .01% of users would have experienced this issue.

This issue was initially detected an hour after the impact start time and was identified and escalated by an underlying service experiencing end user impact. Once detected, multiple engineering teams were engaged to investigate the cause of the issue to understand what needed to be fixed. By 11:00 UTC, the appropriate networking team was engaged and began investigating. The underlying cause was identified by 13:00 UTC. We identified the appropriate fix and rolled it out to a single region to validate success. We confirmed success of the roll out and began deploying to other regions in 3 batches. At the end of each batch we validated the success of the fix. By 16:57 UTC, the fix was rolled out to all regions and mitigation was confirmed.

Root Cause: A recent ARM deployment contained a configuration file that stores the URL endpoint that ARM connects to for operation status query calls. The configuration file had an incorrect endpoint for networking resources. Due to this wrong setting, the ARM status query for networking service management operations failed, which customers saw as failures when attempting to create or delete networking resources. The faulty configuration file was not caught prior to production because the staged test cause that the network resource failures was applied after testing was performed on a then healthy configuration file. When picking up the latest configuration file for deployment, the faulty file was assessed for production and not testing. The faulty configuration file was then manually rolled out without testing being performed with the newest configuration, breaking change.

Mitigation: We corrected the incorrect URL endpoint within the configuration file and safely re-deployed to mitigate the issue.

Next Steps: We apologize for the impact to affected customers. We are continuously taking steps to improve the Microsoft Azure platform and our processes to help ensure such incidents do not occur in the future. In this case, this includes (but is not limited to):

- Networking service will onboard to SDP (Safe Deployment Practice) endpoint configuration rollout process immediately, to ensure enough testing is done and enough time occurs between deployment batches to catch any misconfigurations or changes prior to deployment.
- Networking service will immediately plug-in testing and monitoring holes to make sure we immediately identify an issue like this on the networking end as failures were only seen on the ARM end.
- Networking service will work with ARM team to streamline configuration rollout process, to guard against errors that may occur with the current manual deployment process.

Provide Feedback: Please help us improve the Azure customer communications experience by taking our survey: https://aka.ms/AzurePIRSurvey