Azure status history

Product: All | Region: All | Date: Most recent

July 2019

7/30 RCA - Issues logging in to the Azure Portal with a Microsoft Account

Summary of Impact: Between 21:00 UTC and 22:24 UTC on 30 Jul 2019, a subset of customers may have experienced intermittent error messages and failures when logging in to the Azure Portal with a Microsoft account. This issue affected a subset of MSA users who could not authenticate or manage their accounts. Retries may have been successful.

Root Cause: Engineers were performing standard maintenance on a standby module when the active module became unstable. It was determined that combination of factors including a device mismatch and code bug, resulted in the active device becoming unstable.

Mitigation: Engineers routed the user traffic to other redundant paths to recover the service, which restored Microsoft Account services for customers.

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):

- Updates to processes around manageability of all devices.
- Improve and increase the redundancy of the supporting infrastructure, including the manageability of failed devices.

Provide Feedback: Please help us improve the Azure customer communications experience by taking our survey https://aka.ms/HVZ7-JP8

7/8 RCA - Azure Active Directory - Password Changes

Summary of impact: Between 18:09 and 22:32 UTC on 08 Jul 2019, a subset of customers using Azure Active Directory may have experienced password change issues. For hybrid customers, passwords would have appeared to have changed successfully on-prem, but the sync with the backend AAD would have failed. For cloud only customers, password changes with AAD would have resulted in failures. In either case, customers may not have been able to log into AAD or the Azure portal. The following link was a workaround to reset passwords: https://passwordreset.microsoftonline.com

Root cause: Engineers determined that a recent deployment of a certificate caused backend instances responsible for password change functionality to reject the password change operation. The certificate deployment happened according to a schedule, however due to a bug, backend instances were not able to load the new certificate. In addition, the certificate rotation had an unrelated bug that increased the exposure of the issue to a larger number of customers than intended.

Mitigation: Engineers rolled back the recent deployment task and restarted associated frontend instances to mitigate the issue.

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):
- Improved automation and test coverage for rotation of the particular type of certificate to prevent re-occurrence of the same issue
- Verification of all certificate rotation procedures in AAD to eliminate exposure control issues

7/4 RCA - Connectivity Issues - UK South

Summary of impact: Between 06:18 UTC and 16:25 UTC on 04 July 2019, a subset of customers leveraging Storage in UK South may have experienced service availability issues. In addition, resources with dependencies on Storage may also have experienced downstream impact in the form of availability issues. These impacted services include App Services, Virtual Machines, Backup, KeyVault, App Insights, Logic Apps, and Azure Data Factory v2.

Root cause: During the last two weeks of June, the customer traffic on a single storage scale unit in the UK south region had increased more quickly than normal, and the automatic load balancing system had responded by beginning to shift load to other scale units. However, on 06:18 UTC on July 4th, 2019, the increased levels of resource utilization pushed the scale unit above its natural operating limits, which resulted in higher latency and failures. Services dependent on the storage scale unit experienced a high number of failures and latency which manifested in availability issues.

Mitigation: To recover the scale unit to healthy state, engineers applied following mitigation steps to reduce the resource utilization levels:

- Load balancing configuration changes
- Reducing internal background processes in the scale unit

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):

- Enhance existing load balancing and monitoring mechanism to maintain sustainable resource utilization at the storage scale unit level
- Improve service response to high resource utilization levels
- Improve automatic load balancing efficiency to respond more quickly to an unusual increase in traffic

Provide feedback: Please help us improve the Azure customer communications experience by taking our survey https://aka.ms/HMD0-LP0

7/3 Azure Monitoring (Diagnostic logs, Autoscale, Classic Alerts (v2))

Summary of impact: Between 02:00 and 08:00 UTC on 03 Jul 2019, a subset of Azure monitor customers may have received failure errors while performing service management operations - such as create, update, delete - for Autoscale settings, Classic alerts and Diagnostics settings. Existing autoscale, classic alerts & diagnostics were not impacted.

Root cause: As part of service engineering improvements, a configuration change was incorrectly applied to the underlying KeyVault resources which are used by Azure Monitor's management services (Autoscale, Classic Alerts, Diagnostic Settings) during start up. Subsequently when the management services recycled, the services were not able to correctly access the KeyVault resources due to the misconfiguration and hence could not startup correctly.

Mitigation: Engineers performed a roll back of misconfiguration to mitigate the issue.

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:

- Review and improve configuration change, testing & rollout process to prevent a reoccurrence.

7/2 Azure Services - Intermittent Service Availability Issues

Summary of Impact: Between 19:20 and 22:20 UTC on 02 Jul 2019, a subset of customers using Microsoft Azure Services may have intermittently experienced degraded performance, latency, network drops or time outs when accessing Azure resources due to a network event. This impact would have potentially spanned multiple Azure services. During the impact window, traffic peering through San Jose route would have been impacted.

Root Cause: One of the network devices in San Jose had a hardware grey failure at 19:20 UTC, causing traffic going to one of Microsoft peer networks to experience intermittent failures. This partial failure caused intermittent connectivity issues to services peered through San Jose for a subset of customers connecting through this faulty peer.

Mitigation: Engineering localized and isolated the faulty router and took the device out of service at 20:15 UTC. The traffic engineering systems immediately redirected the traffic and services started to recover. The additional load caused by the redirection out of the device was spread across multiple devices to ensure optimal latency and throughput to customers after the failure. At 22:20 UTC all traffic optimizations were completed by the automation systems, mitigating the impact.

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:

- Replaced the faulty route. [COMPLETED]
- Enhanced monitoring and alerting to detect hardware grey failure. [COMPLETED]
- Improve auto-mitigation logic to quickly redirect traffic across multiple devices to avoid future instances of issue. [COMPLETED]
- Continuing the investigation into the network router with the vendor.

June 2019

6/14 RCA - Virtual Machines and Virtual Networks (Management) - North Europe

Summary of impact: Between 15:45 and 20:58 UTC on 14 Jun 2019, a subset of customers using Virtual Machines, Virtual Machine Scale Sets, and/or Virtual Networks in the North Europe region may have experienced failures when attempting service management operations including create, update, delete, and scale. Azure Cloud Shell customers may have seen the error message "Failed to provision a Cloud Shell."

Root cause: The Fabric Controllers are the services that control the allocation of resources in an Azure scale unit. The Fabric Controllers are reached via a Virtual IP Address (VIP) implemented by a Software Loading Balancing (SLB) system. In order to function properly, the Software Load Balancing servers must be configured with the identities of the network switches to which they are attached so that they can create connections to the network switches and send the switches information on how to reach the VIPs.

During a configuration change to decommission old equipment in one Availability Zone, the switches attached to the active SLB servers were incorrectly marked as no longer part of their active scale unit, and the SLBs ceased to connect with the network switches as designed. This resulted in loss of connectivity to the Fabric Controllers in that Availability Zone, which blocked the ability to allocate, deallocate, or change resources in that Availability Zone.

There was no impact to VIPs or Load Balancing services for Azure tenants during this incident, as the SLB system that provides connectivity to the Fabric Controller VIPs is separate from the one that provides VIPs and Load Balancing to Azure tenants.

The rollout of the bad configuration information to the SLB servers was not automatically stopped and rolled back by the Azure Deployment system as the SLB monitoring and alerting components did not determine this was an incorrect and undesirable configuration.

Mitigation: Azure Engineers identified and rolled back the incorrect configuration.

It should be noted that resilience options existed that would have made this issue non-impactful for customers. Specifically, this issue affected one Availability Zone in the region. Services that leverage multiple Availability Zones as part of their service availability design would have been unaffected during this incident. Customers interested in learning more regarding resilience options should visit:

https://azure.microsoft.com/en-us/features/resiliency/

Next steps: We apologize for the impact this incident may have caused you and your services. Among the repair actions being taken to prevent recurrence are the following:

- Verify that no other changes have been made to network metadata that incorrectly removed a switch from an active scale unit.
- Improve the resilience of the VIP in front of the Fabric Controllers by distributing responsibility for it across multiple SLB instances.
- Improvement to the SLB system monitoring components so that any configuration that causes SLB components to no longer connect to network switches is considered an error and rolled back.
- Implement validations on the network-metadata so that switches cannot be removed from membership in an active scale unit without additional review and approval steps.

Please help us improve the Azure customer communications experience by taking our survey https://aka.ms/F5_P-TVG

May 2019

5/22 RCA - Service Management Operations - West Europe

Summary of Impact: Between 15:10 and 21:00 UTC on 22 May 2019, a subset of customers in West Europe experienced intermittent service management delays or failures for resources hosted in this region. Impacted services included Azure Databricks, Azure Backup, Cloud Shell, HDInsight, and Virtual Machines.

Between 23:20 and 23:50 UTC on 22 May 2019, during the deployment of the permanent fix, a small subset of customers using Virtual Machines and Azure Databricks experienced increased latency or timeout failures when attempting service management operations in West Europe.

Root Cause: The issue was attributed to performance degradation in the Regional Network Manager (RNM) component of Azure software stack. The RNM component, called the partition manager, is a stateful service and has multiple replicas. This component saw an increase in latency due to a build up of replicas being created for the service. Prolonged operational delays triggered a RNM bug which caused the primary replica to re-build on two occasions. This caused service management operations to fail while one of the other replicas was taking on the primary role.

Mitigation: Engineers identified the RNM bug and applied a hotfix to the region which helped resolve network operation failures. During the outage, an security validation process on RNM nodes was performing scans which slowed the replica buildout for impacted nodes. Engineers terminated the scans to improve performance. The network operation job queues began to drain and latency returned to normal.

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):

- Develop and roll out dedicated partition for large customers (in progress)
- Implement automatic throttling to balance load (in progress)
- Root cause and fix the cause for high commit latency (in progress)

Provide feedback: Please help us improve the Azure customer communications experience by taking our survey https://aka.ms/XKB-5FZ

5/13 RCA - Network Connectivity - Increased Latency

Summary of impact: Between 09:05 and 15:33 UTC on 13 May 2019, a subset of customers in North America and Europe may have experienced intermittent connectivity issues when accessing some Azure services.

Root cause and mitigation: The impact was the result of inconsistent data replication in a networking infrastructure service. This resulted in unexpected throttling of network traffic to our name resolution servers. Once the issue was detected, engineers mitigated it by updating the configuration of the affected network infrastructure service to override the effect of this data inconsistency. Simultaneously, engineers performed operations to repair the data inconsistency.

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):

- Improve our monitoring to detect data inconsistencies similar to the one that caused this issue.
- Improvements in the system to help ensure such inconsistencies do not occur in the future.

Provide feedback: Please help us improve the Azure customer communications experience by taking our survey: https://aka.ms/7CC-HXG