the Availability Digest

Windows Azure Cloud Succumbs to Leap Year
March 2012

Shades of Y2K! Microsoft's Windows Azure Cloud went down for over a day on Wednesday, February 29, 2012. Starting around midnight as the clock ticked to Leap Day, various subsystems of the Azure Cloud started to fail one-by-one. Soon, applications for many customers became unresponsive. By 8 AM Thursday morning, thirty-two hours later, Microsoft reported that recovery efforts were complete but that "a small number of customers … may face long delays during service management operations."

Smells like a leap-year bug. Was it?

Windows Azure

Windows Azure is Microsoft's Platform as a Service cloud-computing platform. Azure became commercially available in February, 2010, and is currently supported by six Microsoft service centers worldwide. Azure Regions include:

- North America
    - North-Central US (Chicago, IL)
    - South-Central US (San Antonio, TX)
- Europe
    - West Europe (Amsterdam, Netherlands)
    - North Europe (Dublin, Ireland)
- Asia
    - East Asia (Hong Kong, China)
    - Southeast Asia (Singapore)

On the Azure web site (www.windowsazure.com), Microsoft describes Windows Azure as follows:

"Windows Azure delivers a 99.95% monthly SLA and enables you to build and run highly available applications without focusing on the infrastructure. It provides automatic OS and service patching, built in network load balancing and resiliency to hardware failure. It supports a deployment model that enables you to upgrade your application without downtime."

"Windows Azure enables you to easily scale your applications to any size. It is a fully automated self-service platform that allows you to provision resources within minutes. Elastically grow or shrink your resource usage based on your needs. You only pay for the resources your application uses. Windows Azure is available in multiple datacenters around the world, enabling you to deploy your applications close to your customers."

An availability SLA of 99.95%? That's four hours per year of downtime. Microsoft has just used up eight years. Fortunately for Microsoft, it has violated its SLA for only one month since that is its SLA period.

The Failure Timeline

The following is the failure timeline as reported by ZDNet,[1] complied in part from the Azure dashboard (when it was accessible) and partly from Microsoft reports.

1:45 AM GMT Wednesday, February 29, 2012

According to a report posted at 1:45 AM GMT on Wednesday, February 29, by ZDNet, the Azure Service Management subsystem was down in the U.S. and Northern Europe regions. Windows Azure Service Management lets customers manage their application deployments, hosted services, and storage accounts in the Azure cloud.

10:30 AM GMT

Microsoft reported that "As we proceed through the rollout [of a hotfix patch], we will progressively enable service management back for customers."

10:55 AM GMT

Microsoft began reporting problems on its Service Status dashboard with Windows Azure Compute across North America. "Incoming traffic may not go through for a subset of hosted services in this sub-region." It assured customers that applications would continue to run and that there would be no impact on storage.

12:36 PM GMT

The Service Status dashboard became intermittently unavailable. Customers were in the dark as to what was happening or as to any projections of return to service.

2:20 PM GMT

Microsoft claimed that the Service Management issue had been "mitigated" and that service had been restored to the majority of its customers. However, "We still need to work through some issues before we can completely restore service management services." Microsoft said that it is in the process of "verifying the most probable cause" of the outage.

3:30 PM GMT

Microsoft was trying to restore Azure Compute service in the North Europe and U.S. regions. It reported that Service Management was still unavailable for some customers in these regions.

5:00 PM GMT

Customers reported that they could see that their applications were running in the cloud, but they were completely inaccessible. They would not even respond to a ping.

7:50 PM GMT

Microsoft acknowledged that Azure failures were spreading:

---

[1] Microsoft grapples with Windows Azure outage, *ZDNet*; February 29, 2012.

"Due to attempting to restore Compute functionality in the North Central US, South Central US and North Europe regions, functionality has been downgraded or squashed on a range of Azure services, including the Windows Azure Marketplace Datamarket in South Central US; SQL Reporting in North Europe; SQL Azure Data Sync across the East Asia, North Central US, North Europe, South Central US, Southeast Asia and West Europe regions; the Service Bus, Access Control & Caching Portal worldwide; the Service Bus in South Central US, and there are continuing problems in Access Control 2.0 across the North Europe and South Central US and Access Control in the South Central US region."

11:35 PM GMT

Azure customers continued to be impacted by the outage. Microsoft wrote,

"Recovery efforts are still underway. Further updates will be published to keep you apprised of the situation. We apologize for any inconvenience this causes our customers."

11:50 PM GMT

Microsoft claimed that it was 70% complete with the recovery operations in the affected regions.

1:25 AM GMT Thursday, March 1, 2012

Microsoft reported that it had restored all Service Management functionality for customers in the North Europe region.

4:30 AM GMT

85% of the U.S. regions had been returned to service, according to Microsoft.

7:30 AM GMT

Recovery efforts had been completed in both U.S. regions. However, "a small number of customers [in both regions] may face long delays during service management operations."

The Leap Year Fiasco

Microsoft has not been very forthcoming as to what caused this massive outage. However, clues abound; and the general consensus is that the outage was due to one of the most embarrassing bugs a programmer can make: Azure did not handle the February 29th Leap Day correctly.

Clue #1: The outage occurred just after the start of Leap Day, GMT time.

Clue #2: Though customers noted that their Azure applications continued to run, they could not be reached. They would not even respond to pings.

Clue #3: Microsoft said the Azure service management problems were caused by "a cert issue triggered on 2/29/2012."

Clue #4: Microsoft's lead Azure engineer, Bill Laing, Corporate Vice President, Server and Cloud Division, noted in his blog post that "While final root cause analysis is in progress, this issue appears to be due to a time calculation that was incorrect for the leap year."

These clues lead to the conjecture that the outages resulted from a date-related glitch in SSL (Secure Sockets Layer) certificates. These certificates are used by systems to authenticate each other. Without valid certificates, the systems are unable to communicate.

All SSL certificates include the date the credential was issued and the date it expires. Before an application accepts a certificate as valid, it computes the current time to make sure it falls inside that range. Many administrators prefer that their certificates remain valid for relatively short periods of time, sometimes for only one or two years.

If the Azure certificates relied on years consisting of only 365 days rather than the 366 days required every four years for leap year, the cloud platform may have shut down as systems were unable to confirm that they were connected to other trusted nodes. This would explain why applications that were, in fact, running in the Azure cloud could not be reached, even with a simple ping.

Summary

It is troubling that after the Y2K hysteria, we should be experiencing once again a calendar-related failure. A raft of date-simulation products were developed back then to allow systems to simulate dates without changing the system clock, thereby allowing the Y2K transition to be tested while the system remained in production. Many of these products are still around today. If the Azure cloud had been tested for the leap-year problem to the extent that most systems were checked for the Y2K problem, Microsoft may have avoided this disaster.

On the other hand, perhaps the problem was preordained. On the Azure web site quoted earlier, Microsoft stated that Azure provides "resiliency to hardware failure." There is no mention of software failure.

Acknowledgements

We thank our subscriber, Jack Di Giacomo, President of TANDsoft Inc., for bringing this article to our attention. Jack is particularly interested in this incident as his company provides its product Opta2000 (http://www.tandsoft.com/files/products/OPTA2000.html) for clock simulation and time-zone simulation for HP NonStop servers.

In addition to references made above, material for this article was drawn from the following sources:

Microsoft's Azure Cloud Suffers Serious Outage, *IDG News*; February 29, 2012.
Microsoft's Windows Azure has a meltdown, *ZDNet*; February 29, 2012.
Windows Azure Cloud Hit By Downtime, *Data Center Knowledge*; February 29, 2012.
Yes, Microsoft's Azure Was Downed By Leap-Year Bug; *Wired Enterprise*; March 1, 2012.
Leap Day brings down Microsoft's Azure cloud service, *Digital Trends*; March 2, 2012.
Why would "leap day" take down Windows Azure?, *IT World*; March 7, 2012.