the Availability Digest

www.availabilitydigest.com
@availabilitydig

British Airways Downed by Fat Finger and Testing Shortsightedness

June 2017

Say it ain't so, British Airways. On Saturday, May 27, 2017, a systems crash in a British Airways (BA) data center caused a massive disruption of air travel at Heathrow and Gatwick Airports and at airports in 70 countries.

The outage was caused by human error. We often refer to such an event as a "fat finger" blunder. A technician mistakenly turned off the UPS system in BA's active data center and shut down all of BA's servers.

Unfortunately, British Airways suffered from a syndrome that exists in many organizations. BA never had fully tested their failover capabilities to their backup data center. When their production data center was brought down by the errant technician, their only option was to do whatever was necessary to bring the production center back online. Unfortunately, this took the better part of the weekend.

British Airways

British Airways is the flagship airline of the U.K. It is the largest U.K. airline based on fleet size and the second largest, behind easyJet, based on passengers carried. BA transports about 120,000 passengers each day.

BA has two data centers located near Heathrow airport. The BA servers contain everything from customer and crew information to operational details and flight plans.

The Fat Finger Error

On that Saturday morning in May, a maintenance contractor mistakenly shut down the Uninterruptible Power Supply (UPS) feeding BA's active data center. The purpose of the UPS is to guarantee power to the data center in the event of a main power source failure. The UPS contains blocks of batteries that can provide sufficient power to the data center until the backup diesel generators can be started and phased in to the data center power feeds.

Turning off the UPS resulted in the immediate loss of power to the data center. This action bypassed the batteries and the backup diesel generator, thereby preventing a smooth transition from the main power source to the backup power source.

To further aggravate the catastrophe, the power was turned back on in an unplanned and uncontrolled sequence. The resulting power surge caused physical damage to many of BA's servers, storage devices, and network systems.

The power surge had a devastating effect on the airline's communication systems. It affected messaging across all platforms, including operations systems, baggage systems, and passenger processing.

The engineer who turned off the UPS worked for CBRE Global Workplace Solutions. This is the firm that is now assisting BA with its investigation into the outage.

Fortunately, there appears to have been no data corruption or loss due to the incident. However, this is not to say that corrupted data will not surface later.

The Consequences

The outage led to chaos for 75,000 passengers in 170 airports across 70 countries. With the loss of communication, all 200 systems across the airline's network were affected. They included baggage systems, bookings, online check-in machines, call centers, and mobile apps.

At least 1,000 flights were cancelled at Heathrow and Gatwick over the weekend. All flights out of these airports were cancelled on Saturday. All flights out of Gatwick operated on Sunday, but more than a third of the flights out of Heathrow were cancelled. Many travelers spent Saturday night sleeping on yoga mats provided by the airline.

BA spent the weekend rebuilding servers and recovering databases. The only way the airline could communicate with its customers was via Twitter. Operations returned to normal on Monday. However, many who departed discovered belatedly that their luggage had not accompanied them.

The airline is expected to face a bill of more than £100 million for transport, accommodation, and meals provided to stranded passengers.

Why No Failover?

BA has two data centers about a kilometer apart. Why did their applications not fail over to the backup data center? The airline could provide no explanation as to why the backup system did not kick in.

From what I have observed in the industry, the reason for the failover fault was probably because BA had never tested their failover procedures thoroughly. Testing failover of a system in production is a complex and costly endeavor. Since the system is providing services to users on a 24x7 global basis, there is no maintenance window in which the production system can be taken down to test failover.

In order to determine that the backup system will take over processing seamlessly, the production system must be shut down. If the backup system fails to come up, users may be without services until the production system can be brought up again. This could take anywhere from minutes to hours. Therefore, thorough failover testing often is replaced with what I call "faith and hope."

It is common practice for management to be consulted when a failover must be executed. Is it better to try to bring up the backup system, or is it better to try to restart the production system that has just failed? More often than not, the decision is to take the time to restore the production system to operation. At least it is known that it is up-to-date and is capable of providing proper services. Configuration drift, in which updates to the production system are not made to the backup system, is a common reason for a backup system not to behave properly if it is brought into service.

Summary

British Airways chief Alex Cruz is leading an inquiry into the outage. He is being assisted by several power supply specialists. Certainly, part of the inquiry will focus on failover and why it did not protect BA's services in this outage.

As is true with most airlines, BA is no stranger to IT outages. In 2016, IT failures with check-in systems caused widespread passenger delays in June and again in September.

Certainly, one of the issues that BA will have to address is that of failover testing so that it knows it has a viable backup data center in place. One of the best failover testing plans that I have seen is performed by one company that undertakes a full failover every three months. It then runs on that system for three months and repeats the failover to the original system. Since each system is running production for three months at a time, there is a high degree of confidence that each system is operating properly and that failovers will succeed.

Acknowledgements

Information for this article was taken from the following sources:

What Went Wrong in British Airways Datacenter in May 2017?, *UP2V*; May 29, 2017.
BA won't discuss IT crash which grounded thousands of flights, *Datacenter Dynamics*; May 29, 2017.
British Airways IT Systems Glitch Causes Bank Holiday Flight Chaos, *Silicon*; May 30, 2017.
Inquiry to determine whether BA outage was human error, *Datacenter Dynamics*; May 31, 2017.
British Airways IT Outage Could Be Down To Human Error, *Silicon*; June 2, 2017.
British Airways says IT chaos was caused by human error, *BBC*; June 5, 2017.
BA's boss admits human error caused outage, *Datacenter Dynamics*; June 6, 2017.