Partial Draft Root Cause Analysis
Microsoft Azure Storage Service Major Interruption
2014-Nov-19, 00:51 to 11:45 AM (UTC)

Information retrieved from following URL 2014-Nov-24
http://azure.microsoft.com/blog/2014/11/19/update-on-azure-storage-service-interruption/

LEGEND
Fact, or question unlikely to lead to important cause(s)
Further analysis could lead to contributing cause(s)
Further analysis could lead to root cause(s)

11 hour interruption to Azure cloud storage service
Cold shutdown/restart takes considerable time?
No way to warm boot / restart affected processes?
Patch reversion does not restore BFEs immediately
Cold shutdown/restart of BFEs required

Patch installation to BFEs reverted
Standard post-patch-problem practice?
Service disruption detected
Traffic requests going unserved
Traffic requests coming in at normal rate
BFEs not taking traffic

Infinite loop prevents BFE from handling traffic requests?
Loop limits missing or misconfigured or overwritten or...
Latent, unknown, pre-existing bug

Patch not discovered prior to code release to production
New patch triggers BFE bug
New patch ran fine on production Table Front Ends

Decision to install patch on BFEs
Incremental batch rollout process not used

Traffic requests going unserved
All BFEs continuously executing infinite loop

No other means to serve traffic requests. BFEs required

“Human Error”
No. That’s a useless conclusion. People do what they did for reasons that made sense to them at the time. If somebody did something that, in hindsight, looks like an error, you need to figure out why they thought it was okay. What is it about your organization/system that made the “error” seem like the correct action? See the following webpages for a better way to think about this.

http://www.bill-wilson.net/b64 http://www.bill-wilson.net/b87

Performed by Bill Wilson
Http://www.bill-wilson.net

More info about this analysis available at
Http://www.bill-wilson.net/azure-root-cause-analysis