POL00163019

**From:** Chris Broe GRO
**Sent:** Tue 10/05/2016 4:09:44 PM (UTC)
**To:** Alisdair Cameron GRO; Paula Vennells GRO; Kevin Gilliland GRO; Mark R Davies GRO
**Cc:** Rob Houghton GRO; Sharon Gilkes GRO; Andy Garner GRO
**Subject:** Horizon Incident 9th May 2016 - Diagnosis, Lessons Learned & Actions

Dear All,
Please find below first-cut Lessons Learned and Actions from yesterday's incident, which is complementary to Mark's Business Comms note issued earlier.
Best
Chris

---

**From:** Chris Broe
**Sent:** 10 May 2016 14:52
**To:** Sharon Gilkes; Rob Houghton; Andy Garner
**Subject:** RE: Horizon Incident 9th May 2016 - Diagnosis & Lessons Learned - V2

Overview
At approx. 9:15 am on 9th May Horizon Users experienced "transaction timeouts" consistent with volume bottle-necks in the Primary Branch Database. This was taken out of service, a configuration parameter was changed, re-launched at 10:10 and normal service was progressively resumed by 10.30. See the timeline in Fig 1.

- A change had been approved during weekend maintenance to transfer back to the Primary database following a week of successfully running on the Secondary to test resilience-related improvements after the incident in February (the abortive fail-over from Primary to Secondary)
- The Primary has been operational, trouble free, for at least 12 months prior and it is not yet understood why the configuration-change should have been necessary. FJ and Oracle are in urgent technical consultation to establish the Root-Cause.
- In the meantime the system will be kept under close monitoring to re-assure regarding operations and has remained stable.
- A Change Freeze has been applied until the RCA is confirmed.

Lessons Learned
This Major Incident has highlighted urgent need for improvement in several aspects of IT and Business Response:

1) IT Incident Alerting
    a. The IT Helpdesk quickly became overloaded by the volume of calls. This didn't impact resolution but gave a poor Customer experience. The Help Line will be improved with a recorded message to reassure callers that we are already aware of a Major Incident and dealing with it. Expansion of the usage of Grapevine (providing a 'text blast' communication service to > 5,000 registered branches) will also be considered as interim communications whenever Horizon is out of service.
    b. Unrelated changes to the Incident Alert system (text/email) mis-directed and omitted key individuals initially. This didn't impact resolution either in this instance, but circulation lists will be rectified immediately.
    c. Social Media overtook Internal Communications causing understandable frustration and confusion. Service Delivery Teams (starting with Horizon and PostOffice.co.uk) will add Social Media Monitoring Alerts to their procedures with immediate effect.
    d. FJ became aware of the issue from their own operational monitoring, and were already acting on resolution, but were slow to alert POL. Service Delivery Teams will add FJ Branch Activity and Transaction Monitoring to their procedures with immediate effect.

POL00163019

2) Incident Handling
    a. Technologically the recovery was handled effectively. Technical & Service teams were promptly stood-up and conference calls coordinated by Atos, the problem was diagnosed before 9.45, the fix applied by 10.00 and service back on stream by 10:30.
    b. Lessons from the February incident regarding integration of IT recovery activities and POL Comms had been implemented and insofar as the initial recovery was concerned worked well. POL Comms reps were represented on all conference calls and took ownership of the internal messaging. However, as the media and other enquiries came in post-incident (which were individually handled well as evidenced by the consistency of content reported eg on the Web), POL GE was not kept informed of the precise nature and content of these enquiries or their aggregate impact and we're left exposed.

3) Future Avoidance
    a. Both this incident and the previous one have reinforced the need for a "load test" capability in the Live Production Environment. The planning and preparation for the resilience-related improvements to the Secondary Database took the necessary lessons from the previous incident and were successfully executed.

    b. It has yet to be explained why the Primary Database , which had been working for so long, and previously subject to several successful fail-overs and fail-backs, became over-loaded when it was brought back online on this particular occasion. However Horizon is an extremely large and complex system and the best way to increase levels of reassurance surrounding impact of change for changes of this nature will be to replicate BAU levels of activity within the LIVE environment, before the service is handed back to users.

    c. This is not a trivial undertaking, likely to involve simulating c50-100k transactions (representing c5-10 minutes of peak activity) either of a read-only nature, auto-reversed, or via dummy accounts/products/services. Design of such a capability was commenced after the last incident, has not yet been completed (early proposals rejected on grounds of prohibitive cost/complexity) but must be implemented as a matter of priority. We will also explore automated, specialised testing software for this purpose.

4) Business Communications
    a. See separate note from Mark Davies.

Fig 1. Timeline (draft tbc)
Timeline of incident alert:
09:10 – the Service Desk reported to the Atos Service Delivery Director a surge of calls, with 21 calls waiting
09:12 – PV receives email and tweets alerting Horizon crashed nationally, and that help desk constantly engaged.
09:15 - Fujitsu stood up first Technical Bridge to commence diagnosis
09:15 – Business alerts to twitter feed & NFSB conference
09:30 – POL Comms informed by IT of Issue
09:48 – Fujitsu second Technical Bridge held, decision to implement mem-lock configuration on primary server agreed.
09:57 – P1 text update sent to distribution list
10:00 – FJ fix implementation commenced
10:11 – P1 email update sent to distribution list
10:15 – Branches start to come back online
10:30 – Normal Service Restored
10:37 – Business Protection Team process invoked via email (delayed due to the telephony outage in Atos)