CU*ANSWERS HIGH AVAILABILITY PROGRAM REVIEW

EVENT DATE: 5/17/2014 – 5/27/2014

SUMMARY

As part of an ongoing business continuity program, CU*Answers actively maintains a high-availability (HA) core-processing environment with real-time CU*BASE/GOLD data replication between identical hosts located at two geographically dispersed, state-of-the-art datacenters. A minimum of twice each year, HA rollover events are scheduled to redirect core-processing and operations to the secondary datacenter (located in Muskegon, MI) for a minimum period of three business days. At the completion of each event, core-processing is redirected back and operations resumed at the primary datacenter (located in Kentwood, MI). These HA rollover events are invaluable in our effort to validate our procedures and ensure our ability of recovering CU*BASE/GOLD core processing in an effective and timely manner.

Case in point: An unplanned emergency HA "failover" was conducted on the morning of Sunday, May 18, 2014 to bring CU*BASE/GOLD core-processing online at the secondary datacenter. This response became necessary when the primary production host became unresponsive without warning (hence the term "failover"). A rollover event to bring CU*BASE/GOLD core-processing back to the primary production datacenter was scheduled and performed on the evening of Tuesday, May 27, returning to normal daily operations.

Notable characteristics of this event include:

- Represented the second "unplanned" rollover event performed at CU*Answers in the last six months.
- Provided a test of our ability to quickly and effectively failover CU*BASE/GOLD production to the secondary datacenter without notice and without a 'live' production host.
- Required the coordination of daily (24x7) operations from systems at the secondary datacenter throughout the extended Memorial Day weekend.
- Required last minute adaptation and planning of a scheduled disaster recovery test to be performed during the same period of this HA rollover exercise.
- Five credit union library files were scheduled to be upgraded to the FEP beta software version on the same day the failover event occurred. The decision was made to proceed and upgrade performed successfully within three hours after the host at the secondary datacenter was brought online.

This unplanned event demonstrates the value of the investment in high-availability technology resources and rigorous hours of planning and testing to best prepare us for recovering CU*BASE/GOLD core-processing in a timely and effective manner. This report identifies the scenario and circumstances leading up to the decision to perform the emergency rollover as well as any challenges observed, lessons learned, and recommendations for consideration related to this event.

EVENT DETAILS

At approximately 11:05 PM ET on Saturday, May 17, while preparing for end-of-day (EOD) processing, Operations personnel observed that the production CU*BASE/GOLD host suddenly became unresponsive. An immediate call went out to the iSeries Administration Team to alert them of the incident. Upon physical inspection, it was noted that the system did have power but that disk activity was abnormal and attempts to obtain a system console session were unsuccessful.

IBM support was promptly engaged to troubleshoot the incident. Recovery teams were contacted and put on alert status to be ready in case further action was required. During this assessment period, it was critical to determine the status of data replication before any changes were made. By 1:00 AM, the integrity of the data had been confirmed and the decision was made to promote the HA system at the secondary datacenter as the production host.

The process for bringing the HA host online as the primary production system was delayed due to the fact that the replication service still considered the unresponsive host as active. With the assistance of developers at iTERA, this was resolved and recovery teams were given the green light to continue with emergency failover procedures. On the morning of Sunday, May 18, at approximately 3:20 AM, all core-processing applications were back online at the secondary datacenter.

At that time, attention was again directed to recovering the unresponsive host. Later that morning, IBM support arrived on-site and began to physically inspect the host and extract available system log files. All system components were fully tested and certified (no components were replaced). It was concluded that the controllers for the disk array had initiated a form of hibernation mode that resulted in the unresponsive host. The manufacturer stated this condition has been extremely rare on this make and model (cited approximately 12 known cases globally). Based on the test results and vendor recommendations, the decision was made to reintroduce the host to the replication pool at 3:00 PM. By that time, more than 400 million journal entries were generated on the production host that needed to be synchronized with the HA host. By 5:00 PM, full data replication was restored.

Due to the holiday weekend, the roll-back event (usually scheduled for a Sunday evening) was performed on Tuesday, May 27 at 10:00 PM ET with no problems reported.

CHALLENGES

This incident represents the second unplanned, emergency rollover performed at CU*Answers within the last six months. The previous incident (November, 2013) involved a production host that was experiencing performance degradation due to a failed hardware component. This particular incident (May 2014) involved a production host that was completely unresponsive. Both incidents have many similar circumstances, but a few, very important differences. These differences resulted in some unique challenges faced for this recovery and, moving forward, require a fresh perspective and an enhanced, proactive approach to the way high-availability is viewed.

The following challenges were observed during this incident:

1. After the initial failover, we were not able to reestablish data communications with FIS/Certegy (third party EFT vendor).

a. By coincidence, FIS had scheduled a maintenance window of 11 PM – 6 AM (during which the rollover was performed). The remedy involved support personnel from both parties to troubleshoot services and restore communications by 9:00 AM.

2. After the initial failover, we were not able to reestablish data communications with CUSC (third party shared branching vendor).

    a. An initial configuration error message received when bringing the subsystem online was first thought to be informational only and not responded to properly. While reviewing the log files during troubleshooting, this error message was located and the proper response applied.

    b. Procedures for handling subsystem error messages are being reviewed and revised.

3. On the morning following the initial failover, we received early reports that GOLD sessions were not available. It was discovered that the service that validated GOLD licensing was not active. Starting this services resolved the problem.

    a. One of the post-roll procedures for confirming GOLD access was not completed accurately. The recovery team member was able to authenticate but stopped short of accessing menus that require GOLD licensing.

    b. Procedures for confirming GOLD functionality have been revised (with additional detail) and reviewed by recovery team members for future rollover events.

4. On the morning following the initial failover, we received reports from several credit union of a backlog of batch jobs causing delays once submitted (report requests, teller balancing, payroll posting, etc.).

    a. This ended up being an incorrect configuration setting for all active JOB queues. The default setting of "1" only allowed one active job to process at a time. The configuration setting was adjusted and service restarted to resolve the reported problem.

    b. A procedure to confirm the QBATCH configuration setting following all rollover events has been added to the appropriate documentation.

5. On the morning following the initial failover, two credit unions reported that they were not able to communicate with the secondary datacenter. Neither of these two credit unions have networks managed by CU*Answers Network Services.

    a. As a result of the previous unplanned rollover (November 2013), Email announcements are issued every 30 days encouraging all credit unions to regularly test connectivity to the secondary datacenter in preparation for the next planned/unplanned rollover. This announcement includes links to documentation that detail the network routing requirements necessary for connectivity. We will continue to issue these types of announcements to increase awareness.

6. On the morning following the rollover, we received reports of old documents and reports in the print OUTQs.

    a. During regular planned rollovers, a procedure is performed that flushes print OUTQs on the destination host before syncing up current print OUTQs. Since the circumstances of this event included an unresponsive production host, this procedure was not completed (not able to sync

print OUTQs until the host was recovered and replication resumed). This resulted in old documents and reports left in the print OUTQs from the previous rollover (November, 2013).

b. Procedures for handling print OUTQs during failover events are being reviewed and revised.

7. On June 2, it was reported that as many as (24) EFT transactions (at or around the time of the production host becoming unresponsive) were not properly posted.

    a. During regular planned rollover events, EFT subsystems are taken offline (vendor stand-in mode) so that no new transactions are submitted allowing existing transactions in the JOBQ to complete processing. Once the JOBQ has cleared, the rollover process is initiated.

    b. However, in this instance, the production host was unresponsive, preventing these transactions from being processed completely. Typically, these types of transactions are not replicated between hosts until posting has completed. The EFT Programming Team was engaged to identify and trace the transactions, then process each one manually.

    c. For future unplanned rollovers, a more proactive measure will be taken that engages the EFT Programming Team early in the process to confirm all EFT transactions.

CONTINUING EFFORTS AND RECOMMENDATIONS

Whether planned or unexpected, each recovery test and high availability rollover exercise provides us the opportunity to continually seek to improve the process and adjust our procedures. The best way to accomplish this is to “Practice. Learn. Repeat”. The following is a list of action items and projects that we are pursuing to get us closer to that goal:

1. For future incidents, recovery teams will be expanded to include members from all departments (business units) that provide critical business functions. Each department will develop a comprehensive list of post-roll procedures they will follow during rollover events (whether planned or unplanned).
2. System and programming changes happen daily. It is important that recovery teams understand what data is replicated, what data is not replicated, and the implications this can have during planned and unplanned rollover events. Meetings are being scheduled to identify and communicate this information and to review and refine response procedures accordingly.
3. The previous two unplanned incidents resulting in rollover events occurred during non-peak business periods. Recovery teams and department leaders will review what was learned to understand the implications should a future unplanned rollover event occur during peak business hours. The findings will be presented to the Executive Council and any procedural changes merged into existing documentation.
4. Continue the communications campaign to educate clients on the requirements and importance of complying with standard network configuration settings for connecting to CU*BASE/GOLD and testing connectivity to the secondary datacenter.
5. Develop a regular review and audit process to ensure all documented recovery (HA/DR) procedures are current and accurate.

Report submitted by: Jim Lawrence, CBCP | CU*Answers | Manager of Disaster Recovery and Business Resumption Services