CA Flowdock – Outage 22nd & 23rd April 2020

Root Cause Assessment

Document Title	CA Flowdock Outage RCA – 22/23 April 2020
Customer	Multiple Customers
Corporate Escalation	<Global Escalation Number>
Relevant Support Issue Numbers	<Support Ticket Numbers>
Last Saved Date	05/05/2020

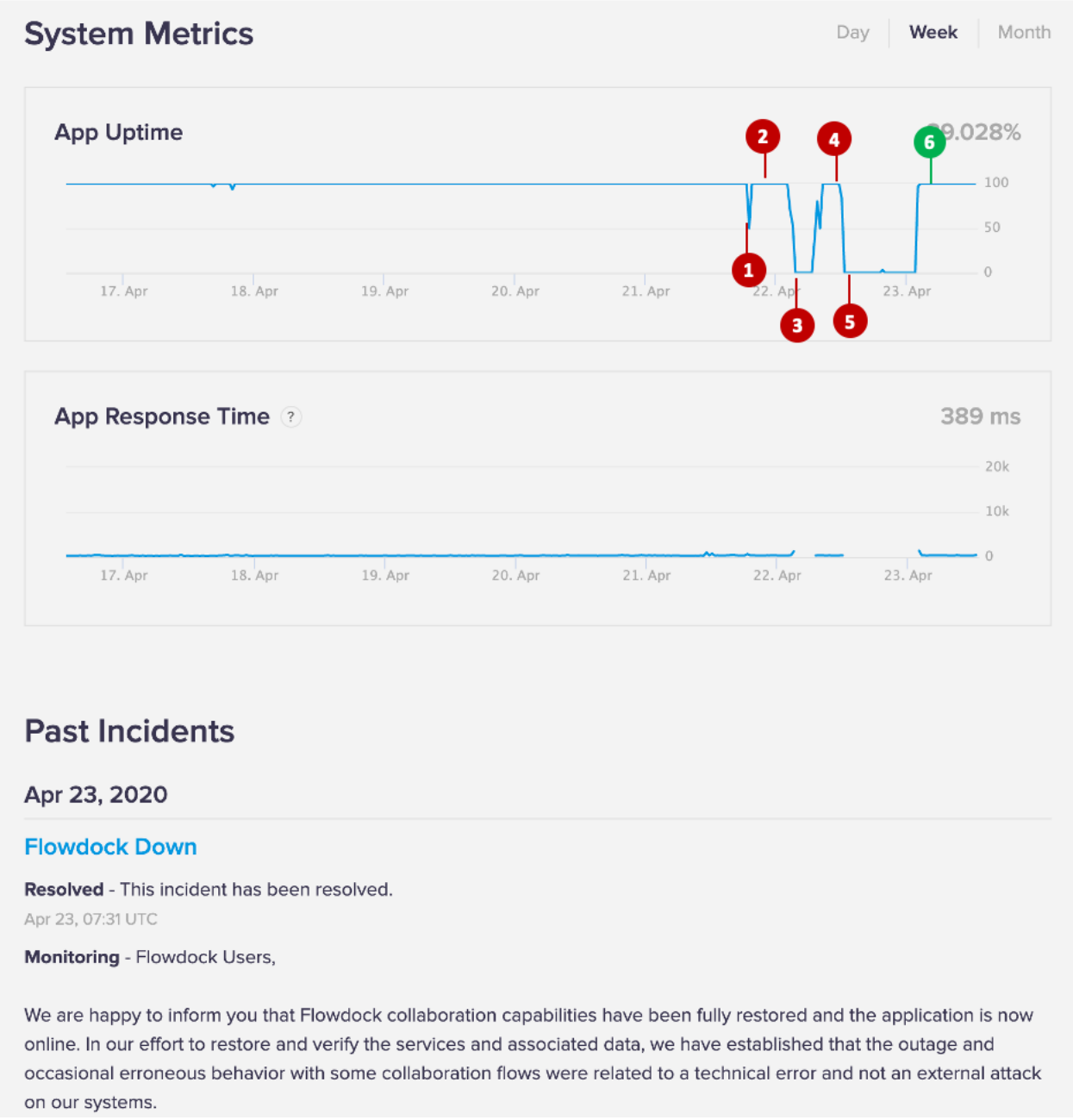
Overview

CA Flowdock is a team collaboration application offered as a SaaS (Software as a Service) capability to multiple CA Technologies, a Broadcom Company, customers. The application enables subscribed organizations to create collaboration flows for their teams to exchange messages and other artifacts, as well as for individuals within the same organization to have 1-1 conversations (private flows). The Product is not designed or meant for the specific purpose of managing, storing or transferring Personal Data. It will however process and store any content that the customer uploads to, or otherwise integrates with the Product

There was a major incident on Flowdock of about 32 hours (between April 21, 1:30 UTC to April 22, 21:30 UTC) leading to the unavailability of the application for about 22 hours (”Outage Window”), and in some cases, loss of data as well as third-party organizations’ individuals being able to potentially access other organizations’ collaboration flow during the “Incident Window” (between April 21, 13:30 UTC to 21:30 UTC, and April 22, 4:30 UTC to 6:30 UTC).

This document provides a complete description of the incident, details of the root cause analysis, and the approaches put in place to prevent such occurrences in the future.

Timeline of the incident



1. Flowdock went down and was restarted **(April 21, ~13:30 UTC)**
    - The application database was hung due to high CPU usage caused by a sudden surge in user activities from multiple organizations potentially due to an increased

higher usage of Flowdock due a multitude of users working remotely, despite having already increased resources available to the systems
- The application was restarted including the database services.

2. Some customers started reporting issues citing strange behaviors in their collaboration flows **(April 21, ~13:30 -21:30 UTC)**
    - Customers reported issues such as users missing from conversation flows, some users not being able to log in, some users found users from another organization in their flows, etc. Users with active sessions prior to the crash were not facing this problem.

3. Flowdock application was taken down manually to investigate the problem **(April 21, ~21:30 UTC)**
    - Flowdock services were brought down and investigated for cross-organizational access issues.
    - It was found that a database table was corrupted in step #1. This table maintains user ID sequences and was reset to a number that is almost 80 days old
    - When the database was restarted in step #1, due to this corrupted table, the application assigned some already generated user ID's which were created recently, to the users of other organizations during the log-in process.
    - To remediate the problem, a stable snapshot of the database (pre-crash) was restored, which also repaired the corrupted table, and the application services were restarted.
    - Flowdock was brought back online for users.

4. A few users and internal teams still reported issues **(April 22, ~4:30 -6:30 UTC)**
    - A few internal users and customers still seemed to experience views of cross-organizational flows.

5. Flowdock application was taken down manually for a full verification **(April 22, ~6:30 UTC)**
    - It was found that cross-organizational access for some users still occurred as the user IDs generated in #2 were still in the cache.
    - A stable snapshot of the primary database (pre-crash) was restored again to remove erroneous transactions from the Incident Window. Due to the data restoration, user activities during the Incident Window could not be preserved. These activities included:
        - New flows added during the Incident Window
        - New users created or updated under an organization during the Incident Window
        - New users added to an existing collaboration flow, including 1-1 conversations, during the Incident Window

        - The integration calls from external applications e.g. Rally to Flowdock that could not be completed during the Incident Window and resulted in the removal of the integration from their respective flows.
    - All active user sessions were invalidated.
    - A full verification was performed. It was established through the end-to-end tests that there is no cross-organizational access of flows for any user.
6. Flowdock was brought online **(April 22, ~21.30 UTC)**
    - Flowdock services were made live. Customers were notified. Customers were also notified about potential user activity loss during the IncidentWindow. Users would need to re-create those activities.
    - Failed Rally integration calls that resulted in the removal of the integration from their respective flows would need to be re-added by the flow admins.
    - No other issues were reported so far and the application has been stable since.

Impact Analysis

- The application was unavailable for its users for ~22 hrs
- The fact that only recently defined user ID's were accidentally affected, suggests a very limited exposure of 1-1 conversations
- The restoration of database resulted in the loss of user activities performed during the Incident Window
    - New flows added during the Incident Window
    - New users created or updated under an organization during the Incident Window
    - New users added to an existing collaboration flow, including 1-1 conversations, during the Incident Window
    - The integration calls from external applications e.g. Rally to Flowdock that could not be completed during the Incident Window and resulted in the removal of the integration from their respective flows.
- Further analysis of data and log diagnostics revealed that no .doc, .xls, .pdf or .csv files, which could have been part of the flows, were downloaded.

Retrospection and Next Steps

In our effort to ensure that we prevent such an occurrence from happening in the future, we are introducing the following additional measures in our application hosting and monitoring procedures that are already defined in accordance with security and data integrity best practices.

- Increased h/w resources for database server to handle higher loads in addition to already increased resources
- Improved monitoring of database server response time at more fine-grained time intervals.
- Enhancing the current failover mechanism to switch the application to a secondary database in the event of an unrecoverable failure on the primary database server.
- Inclusion of a Database Administrator in the emergency change approval committee to evaluate database failures before applications are restarted.
- Process and technology improvements to invalidate logged-in user sessions and cached data, by providing them a maintenance notification in advance, when the application is restored or restarted.