1h Dec 20, 9:10 am – Dec 20, 10:10 am

Open Notebook in Datadog

Postmortem

Postmortem IR-1: Rubex prod is down and showing 500 errors

Created by **Rachel Coleman** on Sep 27, 2021

INCIDENT PROPERTIES	
Status	resolved
Severity	SEV-1
Started	Sep 27, 2021 06:14 pm UTC
Commander	Rachel Coleman
Incident Overview	IR-1

You can generate a postmortem from any resolved incident with these fields pre-filled, along with incident metadata and timeline.

What Happened?

Impact on Customers

Customers were unable to log into the US Prod instance of Rubex as a result of a Database issue on the core infrastructure stack.

This lasted for 43 minutes, beginning at 18:00 UTC -6 (12:00 PM MDT)

Why Did it Happen?

Root Cause

Long Running and performance intensive queries from migration jobs to Rubex from On-Prem caused severe degradation of the RDS Database similar to past events we had seen with the DB performance issues. Similar to previous queries that had caused failover events. We manually failed over prior to an automated failover event today. This caused a temporary lack of access as well as required services to be restarted to restore functionality.

Timeline

Sep 27, 2021 at 11:34 am

Kevin Searcy said he was getting internal server errors in a migration going to imports.efilecabinet.net, then said he was unable to log into Rubex Prod.

Sep 27, 2021 at 11:39 am

Brian Rice was able to log into US Prod and imports.efilecabinet.net endpoint.

Sep 27, 2021 at 11:42 am

Brian Rice was seeing multiple App servers in Beanstalk's health check were returning 5% or greater of responses as 5xx errors. I had Rubex-Prod-4 environment (currently in the A slot) make a minor config change in 2 batches to cause IIS to cycle. A minor change like this normally only takes a minute or 2 to apply. By 12:01 it had registered that the change had been applied to batch 1 but failed in the instance health check and wasn't proceeding to batch 2.

Sep 27, 2021 at 11:46 am

Monitors in Datadog started showing the errors as well. All API tests and the HTTP test to accounts.efilecabinet.net alerted.

Sep 27, 2021 at 11:47 am

Brian Rice made a config change to Rubex-Prod-3 to remove the EnhancedLogging flag to prepare this environment to be failed over to as its status in Beanstalk was still healthy (it's in the B slot which would have only been getting requests from beta.efilecabinet.net, imports.efilecabinet.net, etc).

Sep 27, 2021 at 11:54

Reports of Rubex Down

Sep 27, 2021 at 11:55 am

War Room Meeting started by Emily Nash
https://meet.google.com/nxe-btrr-eix

Sep 27, 2021 at 11:55 am

Testing connection of application servers to database

Sep 27, 2021 at 12:11 pm

Investigation began to confirm why app servers could not connect to the DB.

Sep 27, 2021 at 12:14 pm

Customer impact updated by Rachel Coleman

Scope: Customers cannot login to Rubex
Started at: 09/27/2021 12:00 pm

Sep 27, 2021 at 12:13 pm

Brian Rice performed a DB failover between rbx-prod-pgdb-aug30-node01 in us-east-1c to rbx-prod-pgdb-sept16-node02 in us-east-1d in the prod-db-rubex cluster. This was marked as complete by 12:14

Sep 27, 2021 at 12:18 pm

Field updated by Rachel Coleman

Detection Method: Monitor

Sep 27, 2021 at 12:20 pm

Failover to writer Aurora db complete. Connection test successful. Restarting application servers on B environment to test connection B test successful

Sep 27, 2021 at 12:21 pm

Brian Rice did a swap between A/B Beanstalk environments and confirmed that at least in Incognito that he was able to authenticate to the live app servers.

Sep 27, 2021 at 12:37 PM

Ticket opened with AWS Support with Production System Down Status (8956439801) AWS Advised that:

- 'rbx-prod-pgdb-aug30-node01':
  - underlying hardware is healthy and active, so no relation to the on-going EBS/EC2 event
  - had restarted on 2021-09-27 18:13:55 UTC
- rbx-prod-pgdb-sept16-node02
  - underlying hardware is healthy and active
  - had restarted on 2021-09-27 18:13:32 UTC

DBLoadCPU had spiked (the spike is due to non-CPU activity like IO, locks, etc.) and FreeLocalStorage had dipped (meaning temporary storage was heavily used).

Looking at Performance Insights, the following queries were ran between 16:30 - 17:00 UTC, causing the spike in DBLoadCPU:

```
AWS ID: 77460C71972C832AF22D90961ED9AC3E8E1D64A4
AWS ID: 17A15EDB600E83D6083ADCD5019500FCFD7B304B
```

```
SELECT "Extent1"."Id", "Extent1"."AccountID", "Extent1"."CreatedByUserID",
"Extent1"."FileExtension", "Extent1"."FileIdentifier", "Extent1"."SizeInBytes",
"Extent1"."UploadedBytes", "Extent1"."CreatedOn", "Extent1"."UploadSuccessful",
"Extent1"."Deleted", "Extent1"."FilePasswordHash",
"Extent1"."FilePasswordEncryptionType", "Extent1"."TotalPagesPreviewer",
"Extent1"."NodeCommentID", "Extent1"."GeneratedFromFileInfoID",
"Extent1"."FormFillDefinitionID", "Extent1"."EncryptionVersion" FROM "public"
```

and

```
SELECT "Project2"."C1", "Project2"."ParentID", "Project2"."C2", "Project2"."Id",
"Project2"."AccountID", "Project2"."CreatedByUserID", "Project2"."ParentID1",
"Project2"."Name", "Project2"."SystemType", "Project2"."FileInfoID",
"Project2"."ProfileID", "Project2"."CreatedOn", "Project2"."ModifiedOn",
"Project2"."DeletedOn", "Project2"."PurgedOn" FROM (SELECT "Alias1"."ParentID", 1 AS
"C1", "Join2"."Id", "Join2"."AccountID", "Join2"."CreatedByUserID",
"Join2"."ParentID_Alias2" AS "ParentID1", "Join2"."Name", "Join2"."SystemType",
"Join2"."FileInfoID", "Join2"."ProfileID", "Join2"."CreatedOn", "Join2"."ModifiedOn",
"Join2"."DeletedOn", "Join2"."PurgedOn",  CASE  WHEN ("Join2"."AccountID_Alias3" IS
NULL) THEN (CAST (NULL AS int4)) ELSE (1) END  AS "C2" FROM (SELECT DISTINCT
"Extent2"."ParentID" FROM "public"."DbNodeClosures" AS "Extent1" INNER JOIN
"public"."DbNodes" AS "Extent2" ON "Extent1"."AccountID" = "Extent2"."AccountID" AND
"Extent1"."ChildID" = "Extent2"."Id" WHERE "Extent1"."ParentID" = $1 AND
"Extent1"."ChildID" != $2 AND "Extent2"."DeletedOn" IS NULL) AS "Alias1" LEFT OUTER
JOIN (SELECT "Extent3"."ParentID", "Extent3"."ChildID", "Extent4"."ParentID" AS
"ParentID_Alias2", "Extent4"."Id", "Extent4"."AccountID", "Extent4"."CreatedByUserID",
"Extent4"."Name", "Extent4"."SystemType", "Extent4"."FileInfoID",
"Extent4"."ProfileID", "Extent4"."CreatedOn", "Extent4"."ModifiedOn",
"Extent4"."DeletedOn", "Extent4"."PurgedOn", "Extent3"."AccountID" AS
"AccountID_Alias3" FROM "public"."DbNodeClosures" AS "Extent3" INNER JOIN
"public"."DbNodes" AS "Extent4" ON "Extent3"."AccountID" = "Extent4"."AccountID" AND
"Extent3"."ChildID" = "Extent4"."Id" WHERE "Extent4"."DeletedOn" IS NULL) AS "Join2"
ON "Join2"."ParentID" = $1 AND "Join2"."ChildID" != $2 AND ("Alias1"."ParentID" =
"Join2"."ParentID_Alias2" OR "Alias1"."ParentID" IS NULL AND "Join2"."ParentID_Alias2"
IS NULL)) AS "Project2" ORDER BY "Project2"."ParentID" ASC ,"Project2"."C2" ASC
```

Sep 27, 2021 12:39 PM

Brian Rice launches a 3rd Node to the US Production Cluster to increase resiliency this instance is named

```
rbx-prod-pgdb-sept27-node-03
```

The instance was online by 12:43 local time

Sep 27, 2021 at 12:34 pm

Incident set to stable by Rachel Coleman

Sep 27, 2021 at 12:43 pm

Incident set to resolved by Rachel Coleman
Customer impact updated
Ended at: 09/27/2021 12:43 pm

How do we prevent it in the future?

Action Items

Work with AWS to see how we can prevent this in the future