

Postmortem

Postmortem IR-27: Rubex not loading

Created by Rachel Coleman on Mar 21

FOR INTERNAL USE ONLY

INCIDENT PROPERTIES	
Status	resolved
Severity	SEV-1
Started	Mar 20, 2022 04:15 pm UTC
Incident Overview	IR-27

What Happened?

Impact on Customers

Customers can login to Rubex but the home page will not load and gives error messages. This lasted for 9 hours.

Why Did it Happen?

Root Cause

Database maintenance (vacuuming) is scheduled to be performed on a regular base based on a percentage change. We believe that the high number of tests to production (Which gets deleted every 48 hrs) likely caused the cleaning (vacuuming) to not be performed often enough which caused the locking we saw on the db.

Timeline

Mar 19 2022at 6:18 pm

Datadog started to alert that it could not perform its browser UI tests

Mar 19 2022at 6:53 pm

Rachel alerted Brian that Datadog was alerting and asked if he could check Rubex to see what was happening

Mar 19 2022at 7:40 pm

Brian logged on to investigate what was happening

Mar 19 2022at 8:00 pm

Brian restarted the function app that runs Project Hope migrations in hopes that it was what was causing the slow loads

Mar 19 2022at 8:13 pm

Brian reported no difference in load so he restarted the Batch Worker

Mar 19 2022at 8:23 pm

Restarting the Batch Worker did not resolve the issue so Brian started cycling the backend processes

Mar 19 2022at 8:33 pm

Brian stopped Indexer and OCR and reported that they were not the cause of the issue

Mar 19 2022at 8:39 pm

Brian started cycling the app servers 1 at a time to see if any one of them was the cause of the issue

Mar 19 2022at 8:57 pm

After cycling all of the app servers Brian started to see improvement

Mar 19 2022at 9:47 pm

Quinn reported that Rubex was slow for him

Mar 19 2022at 10:29 pm

Brian started cycling the app servers again

Mar 19 2022at 10:52 pm

Brian turned on verbose logs to IIS to see if we could get more details

Mar 20 2022at 8:46 am

Brian checked in in the morning to report the load is in queries against the read node

Mar 20 2022at 9:01 am

Brian took 3 servers off the ALB to see if it would make any difference

Mar 20 2022at 9:22 am

Brian made a config change in beanstalk to get all of the app servers to cycle IIS

Mar 20 2022at 10:11 am

Rachel updated the status page after seeing no resolution from the changes Brian had made

Mar 20 2022at 10:17 am

Brian restarted IIS again

Mar 20 2022at 10:17 am

Brian Rice has been **added as a Responder** by Rachel Coleman

Mar 20 2022at 10:23 am

Task created and assigned to Brian Rice by Rachel Coleman
Determine root cause

Mar 20 2022at 10:24 am

Task created and assigned to **Brian Rice** by Rachel Coleman
Determine fix

Mar 20 2022at 10:45 am

Emily called a war room

Mar 20 2022at 10:46 am

Video call started by Rachel Coleman
<https://meet.google.com/nxe-btrr-eix>

Mar 20 2022at 10:48 am z

Unknown User has been **added as a Responder** by Rachel Coleman
Unknown User has been **added as a Communications Lead**

Mar 20 2022at 11:00 am

Shon rebooted the reader node

Mar 20 2022at 11:08 am

Failover completed

Mar 20 2022at 11:09 am

Preview servers cycled

Mar 20 2022at 11:09 am

Shon rebooted B slot

Mar 20 2022at 11:11 am

Marc shut down the python reporting server

Mar 20 2022at 11:12 am

Python reporting server is shut down but reader sessions were **still** climbing

Mar 20 2022at 11:17 am

Shon initiated replacing servers on A slot

Mar 20 2022at 11:18 am

Beanstalk and beta are back to healthy on the app servers

Mar 20 2022at 11:20 am

Error message on a slot displayed: LWLock:MultiXactOffsetBuffer

Mar 20 2022at 11:26 am

locks appeared to go away

Mar 20 2022at 11:35 am

URL swap in route53 initiated by Shon

Mar 20 2022at 11:49 am

Shon attempted rebuild again of a slot following error

Mar 20 2022at 11:56 am

Curtis reported additional reports from customers via **livechat** that people can get in but they don't see any of their stuff

Mar 20 2022at 12:02 pm

Shon changing account.efilecabinet.net to go back to ALB 2

Mar 20 2022at 12:05 pm

LWLock:MultiXactOffsetBuffer error is back on app server

Mar 20 2022at 12:09 pm

Shon determined it seemed to be app servers causing the traffic to the DB

Mar 20 2022at 12:19 pm

Curtis reported customers were receiving 503 errors but Marc and Emily were unable to replicate

Mar 20 2022at 12:21 pm

Connections reduce for a time (with locks on DB) but are now increasing

Mar 20 2022at 12:29 pm

BoldBI turned off/shut down by Shon.

Mar 20 2022at 12:30 pm

Connections still increasing

Mar 20 2022at 1:00 pm

Deletion of pangea user session in user sessions and usersessionstobdroles table

Mar 20 2022at 1:10 pm

customer for pangea migration IP blocked

Mar 20 2022at 1:10 pm

shon sent imports.efilecabinet.net to rubex.null.efilecabinet.net

Mar 20 2022at 1:16 pm

shon directed imports, import, and beta to rubex.null.efilecabinet.net

Mar 20 2022at 1:24 pm

Marc added a reader node to see if it helps with performance

Mar 20 2022at 1:26 pm

Marc rebooted the reader node to try to remove locked tables

Mar 20 2022at 1:34 pm

restarting Rubex-Prod-3-ALB

Mar 20 2022at 1:42 pm

Restarting Rubex-Prod-2-ALB - B environment

Mar 20 2022at 1:49 pm

Shon restarting previewers due to locks

Mar 20 2022at 1:54 pm

shon restarted previewer 2

Mar 20 2022at 1:55 pm

shon putting preview 2 back onto LB

Mar 20 2022at 1:56 pm

previewer 2 is back in service, previewer 3 is still off LB

Mar 20 2022at 2:00 pm

B slot is prod 3

Mar 20 2022at 2:00 pm

Shon initiated swap of b and a slot and is rebuilding B slot now

Mar 20 2022at 2:07 pm

attempting to rebuild previewer

Mar 20 2022at 2:11 pm

applying the hotfix that was applied to the UK on Thursday to US so that US and UK match. (edited)

Mar 20 2022at 2:21 pm

A slot is updated with the latest code

Mar 20 2022at 2:24 pm

The locks continue to show up in the DB.

Mar 20 2022at 2:28 pm

shon initiating failover to node 2 (on database)

Mar 20 2022at 2:33 pm

sessions have normalized, shon is initiating reboot readers one at a time (reader 1 reboot en route)

Mar 20 2022at 2:38 pm

locks have returned

Mar 20 2022at 2:39 pm

failover en route

Mar 20 2022at 2:41 pm

sessions have normalized

Mar 20 2022at 2:45 pm

locks are no longer present

Mar 20 2022at 3:05 pm

Previewer has been fixed (though still slow)

Mar 20 2022at 3:07 pm

got under max cpus

Mar 20 2022at 3:51 pm

The app is still slow to respond.

Mar 20 2022at 3:53 pm

Brian is swapping A and B to get the older version on A slot
we are rolling back to the code released on the 8th to rule out code changes

Mar 20 2022at 3:59 pm

Brian is rebooting the servers receiving traffic to old code

Mar 20 2022at 4:06 pm

Brian is now failing over DB

Mar 20 2022at 4:08 pm

Brian is rebooting node 1

Mar 20 2022at 4:08 pm

brian is cycling app servers

Mar 20 2022at 4:18 pm

A and B slot swapping back to bring A slot back to the current version to match UK

Mar 20 2022at 4:44 pm

Brian on phone with AWS

Mar 20 2022at 5:15 pm

brian going to change maintenance mem to 2 gb from "4179MB"

Mar 20 2022at 5:17 pm

Brian undid that since our previous value was already higher

Mar 20 2022at 5:17 pm

vacuum freeze verbose "DbNodeToFileInfoes";

Mar 20 2022at 5:20 pm

first vacumm finished, doing dbnodes now

Mar 20 2022at 5:36 pm

it's still vacuuming dbnodes, after that we'll vacuum dbfileinfoes

Mar 20 2022at 5:44 pm

it's still running the cleanup of nodes

Mar 20 2022at 5:57 pm

dbnodes finished, starting vacuum freeze verbose "DbFileInfoes"; (dbnodes took 35 minutes to run)

Mar 20 2022at 6:05 pm

dbfileinfoes has completed

Mar 20 2022at 6:07 pm

Recommended vaccuming is complete

Mar 20 2022at 6:08 pm

shon started batch worker

Mar 20 2022at 6:09 pm

indexer restarted by shon

Mar 20 2022at 6:12 pm

OCR Components have been restarted, all Rubex components back online

Mar 20 2022at 6:15 pm

Attribute updated by enash@efilecabinet.com

Root Cause: Auto vacuuming may not have been running at the rate of frequency needed for the amount of test Project Hope migrations that were taking place and being purged.

Mar 20 2022at 6:15 pm

Task completed by Rachel Coleman

Determine fix

Task completed Determine root cause

Mar 20 2022at 6:16 pm

Unknown User has been added as a Responder by tchadwick@efilecabinet.com

Mar 20 2022at 6:22 pm

Attribute updated by Rachel Coleman

Detection Method: Monitor

Mar 20 2022at 6:26 pm

Incident set to stable by Rachel Coleman

Mar 20 2022at 6:36 pm

Route53 stuff has been updated (beta, imports, and import back to normal routing) and the migration customer IP address has been updated.

Mar 20 2022at 6:50 pm

Incident set to resolved by Rachel Coleman

Mar 20 2022at 6:59 pm

Root Cause: Database maintenance (vacuuming) is scheduled to be performed on a regular basis based on a percentage change. We believe that the high number of tests to production (Which gets deleted every 48 hrs) likely caused the cleaning (vacuuming) to not be performed often enough which caused the locking we saw on the DB.

Mar 21 2022at 12:48 pm

Summary: Customers saw degraded performance and the database had multiple table locks. At one point the table locks caused the application to be unreachable.

How do we prevent it in the future?

Action Items

- Determine root cause

Created by: Rachel Coleman

- Determine fix

Created by: Rachel Coleman