※ 実際に弊社でインシデントが発生した際に作成した PoMo です。
[ ] の箇所はマスクになるので、全体像だけご参考にしてください

[TicketNo] Trade-in service Outage

Date: 2024-12-19

Authors: ke-no@

Reviewer: CTO, COO, PjM, PdM

Status: Close

Summary / 概要:

- 2024-12-19 08:00 (JST) 頃から 2024-12-19 13:34 (JST) 頃まで買取サービス内の認証エラーが急増し、サービス全体が利用できない状態になっていた
    - [Other Incident] の対応により、リトライ処理を実施していたが、リトライでは解決できない頻度で発生するようになった

Impact / 影響:

- 買取サービスを利用する全ユーザーに影響
    - [社内サービス A]からのアクセスや、[社内サービス B]からのステータス更新なども対象
- 主なエラー件数（2024-12-19 08:00-13:34 JST）
  ※内訳とその他は「Supporting information > エラー件数詳細」に記載
    - TOP ページ　254 件（定期監視 132 件）
    - [Feature A]　148 件
    - [Feature B]　271 件
    - [Feature C]　33 件
    - [Feature D]　14 件（[Followup] 1 件）
        - 当機能のエラーにより[ステータス A]への更新遅延が発生
    - [Feature E]　94 件
    - [Feature F]　131 件
    - [Feature G]　637 件（[Followup] 123 件）
        - 当機能のエラーにより[ステータス B/ステータス C/ステータス D]への更新遅延が発生

Root Causes / 原因:

- 買取サービスバックエンド間の通信時に認証エラーが発生していたため

    - API リクエストを集約し共通処理を行う gateway サービスから内部の実処理を行う gRPC Server を呼び出す際に認証エラー（401 Unauthorized）が発生していた
    - [Other Incident Link] の際は再実行で解決していたが、レート上昇により再実行ではカバーしきれなくなった
        - Google Cloud 側に [Case Ticket] にて確認
            - Google Cloud 側の認証チェック時の挙動が変わり、買取サービスで送っているトークン形式では認証が行えなくなっていた
            - 技術詳細については「Supporting information > 技術詳細」に記載

Trigger / 引き金:

- 2024-12-19 08:00（JST）頃に発生した認証エラーの急増により発生

Resolution / 解決策:

- 一次対応 (2024-12-19)
    - 買取サービスバックエンド間の認証を Google Cloud 側でなく Belong 側で行う方式に切り替えた
        - 詳細
            - gateway からトークン付きでリクエストを受ける [Service A] と [Service B] の各サービスで IDToken の検証を実施するように変更
            - 一時的な対応であるため上記検証に加え、IDToken 発行元の ServiceAccount が想定されたものであるかの確認も実装
            - パッチ反映前に Dev/Stg 環境にて IDToken 有無/正誤での検証を実施してから変更をリリースした
    - [社内サービス] からの通知で失敗していたものは翌日 2024/12/20 に再送処理を実施
- 恒久対応 (2025-01-16)
    - 「Supporting information > 技術詳細」に記載の通り、gateway サービスにて Authorization ヘッダーが重複しない形でバックエンドへのリクエストが行えるように修正

Detection / 発見:

- 買取サービス TOP ページの自動監視によるアラートにより通知が発生
- L1 サポートから Eng チームへのエスカレーションがあったため

Action Items / 対策:

- パートナーが増え対応の重要度が増してきた買取サービスのサポート体制の再考
- エラーにならない形で認証トークンを渡せるように買取サービスを修正し、元の Google 認証を利用できるように設定を戻す

    - [Ticket A]
    - [Ticket B]

Lessons Learned / 振り返り:

What went well / 上手く行ったこと:

- サービスが全てダウンする規模の大きいインシデントではあったが、対応開始から 2 時間半以内にパッチ実装〜検証を行い、3 時間半でサービスの復旧までを行えた
    - サービスの内部理解だけでなく、Google Cloud や Go 言語といった使用している技術についての知識も活用し、密度の高いインシデント対応が行えていた
- システム的に [Other system] からのイベント再通知も問題なく受け取れるようになっていたため、データメンテナンスなど個別に対応をする必要がない構成であった

What went wrong / 上手く行かなかったこと:

- 解決策の検証において、Dev->Stg->Prod というステップを取れず、環境間に差分が生まれ、Prod への適応時に検証時と異なる状態になってしまった

Where we got lucky / 幸運:

- インシデントが発生した 12/19 は CTO, 買取サービス開発チーム, PjM/PdM が全員出社しており、スムーズにコミュニケーションを取りながら解決まで事を進めることができた
- 休日や年末年始など対応の難しい日ではなく平日に発生し、かつ年末年始の長期休暇前に認証エラーの懸念を払拭できたこと

Timeline / 時系列:

2024-12-19（以下は全て JST）

08:00 買取サービスバックエンドで設定していたリトライ上限を突破するリクエストが発生
以降、同日 09:08 頃まで少しずつエラー発生頻度が増加

09:09 全てのリクエストがリトライ上限を突破、買取サービス全体がダウン

09:39 インシデント宣言、対応を開始

11:58 [Service A/Service B] の修正を含む PullRequest を作成

12:07 上記修正を Dev 環境にデプロイし、動作を確認

12:24 Dev 環境での動作確認を完了、Stg 環境へのデプロイ開始

12:35 Stg 環境へのデプロイ完了、動作を確認

12:47 Stg 環境での動作確認を完了、本番リリース作業を開始

13:02 Prod 環境にデプロイ完了、動作を確認中 API 呼び出しで 403 エラーになることを発見

13:26 Prod 環境にて買取サービスの [Account A] に対して [Role A] 権限の付与
（Prod 操作時に意図しない権限削除が発生してしまったため再付与を実施）

13:34 買取サービス全体が復旧

15:01 全環境でインシデント対応時に消えてしまった権限を再付与

15:32 同構成の [Service A] に対しても同様の対応を行い、エラーを解消

2024-12-20

10:30 [Other System A] からの通知エラーA 1 件を再送完了

11:29 [Other System A] からの通知エラーB 13 件を再送完了

15:30 [Other System B] からの通知エラー 123 件を再送完了

2025-01-16

15:42 本インシデントの根本修正を含む [Service A], [Service B]をリリース

Supporting information / 関連情報:

エラー件数詳細

- TOP ページ
    - 254 件（定期監視 132 件）
- [Page A]
    - [Page A-1] 129 件
    - [Page A-2] 19 件
- [Page B]
    - 271 件
        - [Partner A] 145 件
        - [Partner B] 32 件
        - [Partner C] 29 件
        - [Partner D] 65 件
- [Page C]
    - 33 件

        - [Partner A] 27 件
        - [Partner B] 6 件
- [Page D]
    - 2 件
        - [Partner A] 1 件
        - [Partner B] 1 件
- [Page E]
    - 33 件 ※[Page E] 遷移前のデータ取得が行えていないため、発生数が少ない
        - [Partner A] 28 件
        - [Partner B] 1 件
        - [Partner C] 1 件
        - [Partner D] 3 件
- [Page F]
    - 1 件
- [Page G]
    - 14 件（要フォローアップ 1 件）
- [Page H]
    - 94 件
        - [Partner A] 83 件
        - [Partner B] 2 件
        - [Partner C] 3 件
        - [Partner D] 6 件
- [Page I]
    - 131 件
        - [Partner A] 100 件
        - [Partner B]2 件
        - [Partner C]16 件
        - [Partner D] 13 件
- [Page J]
    - 637 件（要フォローアップ 123 件）

技術詳細

1. grpc-gateway を立ち上げているサービスに認証がかかっている場合、runtime/context.go の処理にあるように、gateway サービスに向けた Authorization header をそのままバックエンドサービスにもパススルーする挙動になっている

2. #1 での挙動に加え、grpc.PerRPCCredentials でバックエンドサービスへのトークンを Authorization header に設定するため、2 つのトークンが Authorization header として登録され、送信される
3. Google Cloud サーバー側は複数の Authorization header がある場合、それらを「,」区切りで結合する
    a. これにより、バックエンドサービスで受け取る Authorization header は「{bearer バックエンド向けトークン},{bearer 使用された gateway 向けトークン}」の形式となる
4. Cloud Run の認証チェックでは Authorization header がカンマ区切りの場合は最初の値を使用していたが、エラー該当期間内でアップデートが発生
   カンマ区切りなどは見ず、Authorization header 全体を 1 トークンとして扱うようになったため、#3-a の形式が不正な形式となり、401 エラーが発生していた
    a. 上記変更があった旨は Google Cloud [Case Ticket] でも確認済み
    b. 段階的にエラーが増えていったのも、Google Cloud 内部の更新を徐々に解放していったものと思われる