

Google Cloud の活用による徹底した負荷対策によって世界的スポーツ イベントにおいて安定した視聴体験を提供



株式会社AbemaTV

<https://abematv.co.jp/>

〒150-0042

東京都渋谷区宇田川町 40-1 Abema Towers

2015年4月、株式会社サイバーエージェントと株式会社テレビ朝日によって共同で設立。テレビのイノベーションを目指し「新しい未来のテレビ」として「ABEMA」を運営。登録は不要で、国内唯一の24時間編成のニュース専門チャンネルをはじめ、オリジナルのドラマや恋愛番組、アニメ、スポーツなど、多彩なジャンルの約25チャンネルを24時間365日放送している。

インタビュー（写真左から）

株式会社AbemaTV

- ・ 開発本部 SRE 海野 嘉和 氏
- ・ 開発本部 プロダクト開発テックリード 辻 純平 氏

株式会社AbemaTV（以下、AbemaTV）が運営する「新しい未来のテレビ」 「ABEMA」では、サービス開始当初よりバックエンド システムの構築に Google Cloud を採用しています。過去最大の WAU（週間アクティブ ユーザー数）を記録した世界的スポーツ イベントの全試合無料生中継で安定した配信を実現した背景について、負荷対策を担当したお二人にお話を伺いました。

利用しているサービス

Anthos Service Mesh, Google Kubernetes Engine, Cloud Bigtable, Cloud Run, Pub/Sub, Cloud Trace

高トラフィックに対応するためにシステム構成の見直しを実施、Google Cloud の性能をフル活用へ

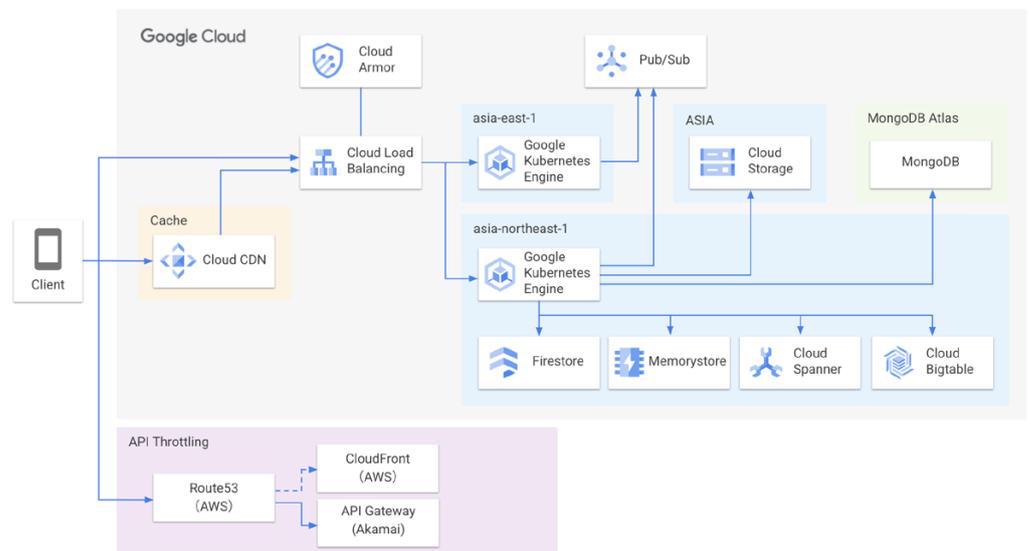
AbemaTV では、2015年のサービス スタート当初より、API やデータ分析、レコメンド システム、広告配信といった周辺サービスの基盤として Google Cloud を利用しています。ABEMA のサービスが大規模になることを想定して、そのトラフィック

に対応できる基盤が必要であったことや、大規模でも高速な開発サイクルを実現できることなどから、当時の大手クラウド サービスでは唯一 Kubernetes のマネージド サービスを提供していた GKE (Google Kubernetes Engine) を採用したと開発本部 SRE 海野 嘉和氏は語ります。「当時はまだコンテナ オーケストレーションの導入事例そのものがあまり多くは無かったのですが、通常の VM ベースの開発に比べると、コンテナを利用した開発は圧倒的に開発速度が速かったという実感があります。実際にサービスが拡大した現在でも、マイクロサービスがどんどん増えて管理が複雑になってくると、当時 GKE を選択したのは間違いではなかったと実感しています。マイクロサービスでよく課題になるネットワークの複雑性について

も、Anthos Service Mesh を導入したことである程度は解消できているので、大きな問題にはなっていません。」

今回のスポーツ イベントの中継でも基本的なアーキテクチャは既存のものが利用

システム構成図（概要）



できるものの、急激なトラフィックの変化に備えてこれまで積み重なった技術的負債を解消することも必要だったと、開発本部 プロダクト開発テックリード 辻 純平氏は説明します。

「今回は過去に無い規模のトラフィックが予想されたので、従来から抱えていた技術的な課題を解消して、Google Cloud の性能をフル活用できる形に改修しながら負荷対策を進めていく必要がありました。例えば Bigtable に関しては、データ整合性を重視してシングルゾーンで運用していたものを、アプリ プロファイルを活

用して整合性を維持しつつマルチゾーン構成に変更しました。インメモリ データストアに関しては、もともとは Redis Cluster をセルフ ホスティングで運用していたのですが、高負荷でのレイテンシー対策としてマイクロサービスごとに分割するように構成し直しています。MongoDB も同様にマイクロサービス単位での分割を行いました。また、それまで台湾リージョンで運用していたのですが、レイテンシーなども考慮して今回の機会に東京リージョンへの移設も行っています。」

当日に起こりうる問題を洗い出すために、PSO チームと連携して徹底した事前対策を実施

実際の試合当日におけるトラフィックは ABEMA にとって過去最大のものとなりましたが、結果的に全試合大きなトラブル無く安定した配信を行うことができました。その成功の背景には、前述のシステム改修と並んで、Google Cloud の Professional Services Organization (PSO) チームと連携して行った徹底した事前対策がありました。

「当日に向けた負荷対策として最初に行ったのが、トラフィックの見積もりと、我々がクリティカル ユーザー ジャーニーと呼んでいた「絶対に停止させてはいけない部分」の特定です。トラフィックについては、まず直前に開催された大規模な格闘技イベントの配信からアクセスログを分析し、それを参考にしてリクエスト数の見積もりを行いました。この見積もりはあくまでもテストシナリオ作成の参考にするための大ざっぱなもので、予測したトラフィックを基にして十分に大きな負荷で試験を行い、クリティカル ユーザー ジャーニーが落ちないように対策していく、というのが全体の大きな流れです。」(海野氏)

テストシナリオの作成にあたっては、試合中に起こり得る出来事とそれに対するユーザーの行動のモデル化を行いました。これは、リアルタイムに変動するアクセスに対してシステム側にどのような現象が起こるのかをシミュレートし、問題を洗い出すためのものです。

「実際の中継では、試合の展開に合わせて刻一刻とトラフィックが変動します。例えば試合開始直後やハーフタイム、ゴールが決まったときなど、さまざまな状況でのユーザーの行動を想定してシナリオを作り、それに基づいてクリティカル ユーザー ジャーニーの負荷対策や、トラフィックのスパイクへの対策などを考えました。」(海野氏)

PSO チームからは、キャパシティ プランニングにあたってのアドバイスや、負荷試験のレビュー、障害試験の提案などといった支援を受けることで、負荷対策の確実性が大幅に向上したと海野氏は語ります。

「社内のチームだけではどうしてもさまざまなバイアスがかかって工程の取りこぼしや優先順位付けの誤りなどが起こりがちになるため、PSO チームから率直なアドバイスをもらったことは非常に助かりました。我々の負荷対策のアプローチに誤りがないかといったことをはじめとして、Google Cloud の内部的なコンポーネントの特性を考慮した問題点の洗い出しなど、プロダクトの中身を知る専門家ならではの視点で確認してもらいました。」(海野氏)

負荷試験には Kubernetes で動作する負荷テストツールの k6 を使用し、クリティカル ユーザー ジャーニーに至るさまざまな API のシーケンスコールをテストシナリオに

基づいて再現。段階的にクライアント数を増やしながら問題が発生したら対策を施すというサイクルを繰り返して、最終的にビジネス目標とされていた同時視聴者数を達成しました。具体的には、データベースのスケールアウトやサーキット ブレーカーの追加、Kubernetes ノードの増強、必要なリソースの追加などの対策を行ったとのこと。障害試験では、特にリスクが高いコンポーネントが停止したり、外部サービスとの連携が遮断された場合などに、その影響範囲やシステム全体の挙動がどのように変わるかなどを確認しました。障害試験の組み立て方や、意図的に障害の状態を作り出す方法などについては、PSO チームのアドバイスが非常に役に立ったと辻氏は語ります。

「クラウドを意図的に壊すことは原理的にできないので、どのようにして障害の状況をシミュレートするかという問題がありました。その点について、PSO チームからは各サービスの特性を考慮した上で、仮想的に障害の挙動を発生させる方法などをアドバイスしてもらいました。クラウド内部の特性を含めた知見を共有してもらえたことには大きな価値がありました。」(辻氏)

今回の負荷対策に向けた改修項目のひとつに、分散トレース システムである Cloud Trace の全面的な導入がありました。この Cloud Trace が、負荷試験や障害試験において発生した問題の分析に非常に役立ったと辻氏は語ります。

「マイクロサービスではどうしてもどこで問題が発生しているのかを特定しにくいという特性があるのですが、Cloud Trace を導入したことで細部のテレメトリーまで正確に把握できるようになったので、問題の特定にかかる時間を大幅に削減できました。試験中に明らかになった問題の中には、Cloud Trace が無ければ発見できなかったような性質のものもあります。今回の負荷試験は複合的な機能に着眼したもので、障害も連鎖的に発生するため非常に分析が難しいという特徴があり、分散トレースの仕組みが本当に役に立ちました。」(辻氏)



海野氏

辻氏

負荷試験・障害試験を最小限の手間で実施できるようにプラットフォームの改善を進めていく

これらの事前対策の成果により、大会中の配信では大きなトラブルは発生せず、始終安定して中継を届けることができました。実際には、想定していた最大同時接続数を上回るリクエストが発生した時間帯もあったものの、結果的に全試合に対して完全な機能を提供することができたと海野氏は語ります。

「同時接続数に対しては、目標最大数の半分までのリクエストに対しては確実に全機能を提供し、それを超えた場合でも最大数まではクリティカルな機能を維持して配信を継続するという要件を設定していました。結果的に、ピーク時には最大数に近い同時接続数を記録したのですが、システムは想定以上に安定しており、全機能の提供を継

続することができました。これは、PSO チームの協力の下で事前に各種ボトルネックを解消できた成果だと考えています。」(海野氏)

この成功を踏まえて、同様の水準での負荷試験や障害試験を継続的に行えるようにプラットフォームの改善を進めていきたいと、辻氏は次のように語ります。

「今回の経験で、疎結合なマイクロサービスの複合的な試験は非常に手間がかかり、認知負荷も高いことを実感しました。この負担を取り除いて、負荷試験や障害試験を最小限の手間で継続的に実施できるプラットフォームを構築していく必要があると考えています。」

Google Cloud を活用することで、ビジネスの将来に注力できるようになります。インフラストラクチャの管理やサーバーのプロビジョニング、ネットワークの構成などに起因する負担を軽減することができます。つまり、インベーターもプログラマーも、自分の本来の仕事に集中することができます。

お問い合わせはこちら
<https://goo.gl/CCZL78>



Google Cloud の詳細については、右記 URL もしくは QR コードからアクセスしていただくか、同ページ「お問い合わせ」よりお問い合わせください。
© Copyright 2023 Google
Google は、Google LLC の商標です。その他すべての社名および製品名は、それぞれ該当する企業の商標である可能性があります。

