

GKE Autopilot による自動運用で スケーラビリティの向上と運用負荷の大幅な削減に成功



Beatrust株式会社

<http://corp.beatrust.com/>

〒101-0044

東京都千代田区鍛冶町 2-7-15 AD 神田駅東口ビル 3F

2020年3月設立。「誰もが最高の自分を実現できる世界をつくる」をビジョンとして掲げ、そのような出会いを創り出すためのプラットフォームとして、個人の経験やスキルを可視化して協業を促進するタレントコラボレーションプラットフォーム「Beatrust」を提供している。

インタビュー（写真左から）

・ Beatrust株式会社 SRE 中川 裕太 氏

・ Beatrust株式会社 プロダクトマネージャー 堀口 綾 氏

個人の経験やスキルを可視化して社員同士のつながりと協業を促進するタレントコラボレーションプラットフォーム「Beatrust」を提供する Beatrust株式会社（以下、Beatrust）。創業以来 Google Cloud を利用している同社では、ビジネスの拡大に合わせたシステム基盤のアップデートで GKE (Google Kubernetes Engine) Autopilot を採用し、大幅なスケーラビリティの向上と運用負荷の削減を実現しました。この移行プロジェクトを担当したお2人に話を伺いました。

利用しているサービス

[Google Kubernetes Engine Autopilot](#), [Cloud SQL](#), [Google Compute Engine](#), [BigQuery](#)

利用しているソリューション

[Application Modernization](#)

ビジネスの拡大に合わせて、GKE Autopilot の採用によるシームレスな自動スケールに挑戦

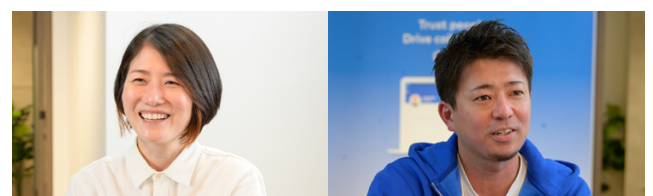
Beatrust による企業向けのコラボレーションプラットフォーム「Beatrust」では、「全ての出会いを最適化する」というミッションに基づいて、企業内における社員同士の自律的な共創をサポートするさまざまな機能を提供しています。「Beatrust」を利用することで、社員はチームの垣根を越えてより多くのメンバーと経験やスキルを共有したり、つながりを見つけて新しい価値をつくり出すことができます。同社の掲げるミッションについて、プロダクトマネージャーの堀口 綾氏は次のように説明します。「当社の掲げる『全ての出会いを最適化する』というミッションは、決してツールだけで実現できるものではないと考えています。組織文化の特性や、どういったプレイヤーが必要かなど、さまざまな要素を含めて検索する必要があり、『Beatrust』はそれを包括的にサポートする役割を担っています。」

同社では、この「Beatrust」のシステム基盤として2020年の創業当初より Google Cloud を利用してきました。サービス開始当初は開発速度を重視して Cloud Run を採用し、その後間もなくしてビジネスの要件に合わせて Google App Engine (GAE) Standard ベースのアーキテクチャに移行します。そして GAE への移行後に、Cloud Logging や Redash を活用したデータ分析基盤や、データドリブンな機能を提供するための機械学習基盤など、さまざまな機能拡張が行われました。

順調にビジネスを成長させてきた同社ですが、事業規模が拡大するにつれて、GAE Standard によって構築した既存システムでいくつかの課題が顕在化することになったと SRE の中川 裕太氏は説明します。

「主な課題は大きく分けて3つありました。1つ目は、ユーザーからのアクセスの増減

に対する自動スケーリングを柔軟にコントロールし、私たちの意図をより正確に反映させられるようにしたかったこと。2つ目は、Cloud SQL への接続に利用していた VPC アクセス コネクタがアクセスのピーク時に不安定になるという問題があったこと。そして3つ目が、データドリブンな機能のために使っている機械学習のワークロードが不安定で、レイテンシも悪化していたことです。これらの課題を解消するためには、場当たり的な対応では不十分でした。つまり、これまでは開発の容易性を重視して GAE Standard を採用していたのですが、今後のビジネスの拡大を見据えて、よりシンプルでコントロールしやすいアーキテクチャにする必要があったわけです。」同社がこれらの課題への対策として選択したのが GKE Autopilot でした。GKE Autopilot は、Kubernetes マニフェストに基づいてコンピューティングリソースを自動でプロビジョニングする GKE のモードの1つです。インフラの管理を Google Cloud に任せられることができるため、IT 担当者がアプリケーションの構築とデプロイに専念できるというメリットがあります。GKE Autopilot の採用理由に



堀口氏

中川氏

ついて、中川氏は次のように語ります。

「ビジネスの拡大に備えてスケーラビリティを向上させたいというニーズを考慮して、Cloud Run に戻るか、新たに Kubernetes に挑戦するかという 2 つの選択肢があり、将来的なことも踏まえて新しいことにチャレンジしたいと考えて GKE を採用しました。Autopilot の利用については、開発メンバーが限られているため、運用負荷を下

げたいという思いで検討したのです。当社の要件で自動スケールが適切に動作するかや、ワークロードが安定するかなどを検証した結果、当時の本番環境のワークロードに耐えられることを確認したので採用を決定しました。その後構築を進めるなかで、パラメータ チューニング 等によって最終的に当時の約 10 倍のワークロードにも耐えられることが明らかになり、将来への期待も高まりました。」

GKE Autopilot に移行したことで、スケーラビリティの向上と信頼性の高さを実感

GKE Autopilot への移行におけるポイントとして中川氏は、Balloon Pod を利用することや、HPA (Horizontal Pod Autoscaler) の指標を低めに設定することなどを挙げています。GKE Autopilot では、クラスタに余剰のキャパシティがあれば新しい Pod を高速に割り当てることが可能になるという性質があります。この性質を有効活用するために、あらかじめ優先度の低い Pod を作って余剰キャパシティを確保しておくのが Balloon Pod と呼ばれるテクニックです。

「事前検証の段階で、余剰キャパシティがない状態だと Pod の割当に最大で数分程度の時間がかかることが判明しており、Balloon Pod を使うことでこの問題を解消してスケーリングの性能を確保しました。HPA を低く設定するというのも、あえてスケールしやすくすることで、高トラフィック時にレイテンシが高くなるのを未然に防止する意味があります。その他にも、Pod の Disruption Budget を指定して最低限必要な Pod 数を確保しておくことや、Pod Topology Spread Constraints を利用して Pod が複数のノードに分散配置されるようにしておくことなど、安定運用に向けて GKE Autopilot ならではのベスト プラクティスを探っていきました。」(中川氏)

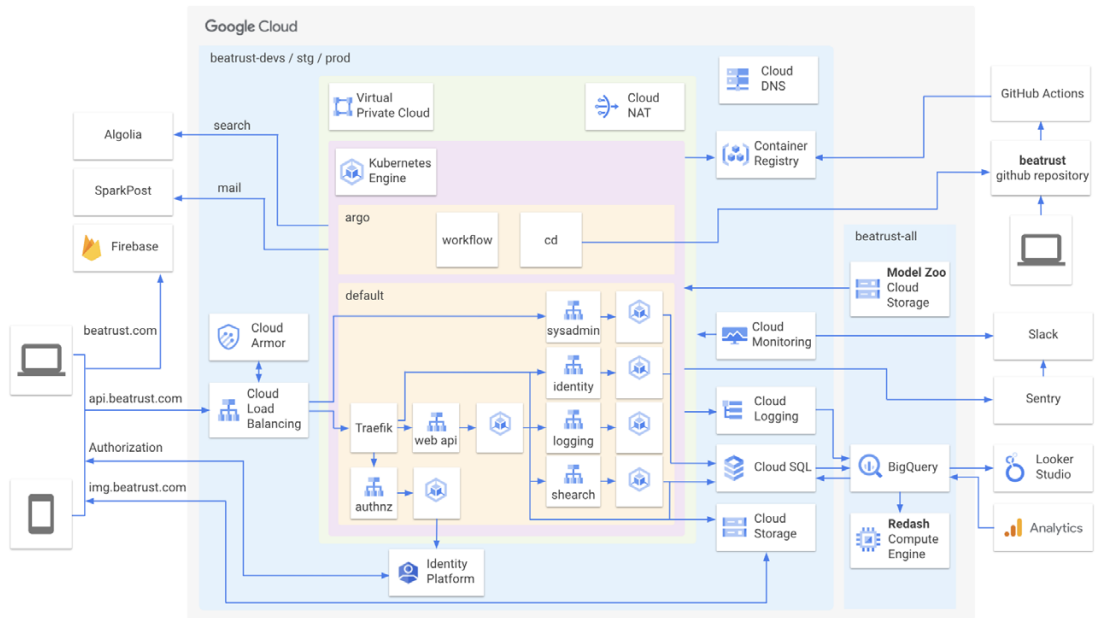
GKE Autopilot への移行プロセスは、既存のシステム構成を維持しつつ、GAE および Cloud Run で動作していたコア機能を優先的に GKE に載せ換えるという方針で実施しました。その背景には、移行トラブルのリスクを最小限に抑えること、スモールスタートで運用ノウハウを蓄積するという 2 つの目的がありました。また、GKE Autopilot によるエコシステムへの対応が進むなかで、利用可能になったツールを順次取り込んでいきたいという狙いもあったと中川氏は説明します。

「当時、GKE Autopilot は正式リリースから間もないこともあり、

当社で使用している周辺ツールでもサポートされていないものがありました。それらについては、将来どのような形にしたいかという設計だけは作っておいて、あとはパーツごとに正式にサポートされ次第、個別に対応できる構成にしておくという方針で進めました。結果的にコア機能の移行を短期間で行うことができ、その後エコシステムのサポートも充実してきたので、GKE Autopilot の管理対象は徐々に広がっています。」

中川氏は、GKE Autopilot への移行による効果として、スケーラビリティが向上したことに加えて、運用負荷が大幅に削減されたことを挙げています。

「日々の運用に力を割かなくてよいということが想像以上に大きな効果でした。日常的なリソース管理などはすべて Autopilot に任せておけるので、我々は他の作業や新機能の開発に時間を割くことができるようになりました。移行してから 1 年以上にわたって特に大きな問題は発生しておらず、Autopilot の信頼性の高さを実感しています。いつ障害が起きるかわからない状態だと他の作業にも集中できなくなってしまうのですが、そういった心配はまったくなくなりました。」



GKE Autopilot のスケーラビリティを海外を見据えたビジネス展開に生かしていく

Beatrust では今後、海外市場に向けてビジネスを拡大させることを計画しています。また、現在は企業内でのコラボレーション プラットフォームという位置づけで展開していますが、将来的には企業同士や個人同士、あるいは企業と顧客など、さまざまな形のコラボレーションの可能性を模索していると堀口氏は語ります。

「ビジネスのバリエーションが広がることで、システム面でも多様なアップデートが必要になってくるはず。国が変われば法律や商習慣も変わってくるので、データの持ち方やプライバシーの問題など、考慮しなければならない課題が増えてきます。当然、ユーザー数の大幅な増加も考えなければなりません。そういった先の展開を見据えたときに、GKE Autopilot のスケーラビリティは大きな強みになると思っています。」

Beatrust では、Google Cloud の活用にあたってスタートアップ企業向けのサポートプログラムである Google for Startups を利用しています。今回の移行プロジェクトにおける Google Cloud チームのサポートについて、中川氏は次のように話します。

「Google Cloud チームからは、構成のレビューやベスト プラクティスに関する情報提供などについて、プロダクトのロードマップなども踏まえたアドバイスをもらうことができ非常に助かりました。質問に対するレスポンスも早く、悩んだときの意思決定が格段に楽になりました。当初は我々のようなスタートアップが Kubernetes を使いこなすことができるのかという不安もありましたが、結果的に小さな組織でも十分に Kubernetes とそのエコシステムの恩恵を受けられるということが確認できました。」

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。

