Google Cloud

# 次世代セキュリティ DWH を BigQuery、Cloud Data Fusion で構築し、ログ分析の民主化を実現


17,000 名を超える従業員を擁し、プラットフォーマーとしても幅広くサービスを展開する株式会社リクルート（以下、リクルート）にとって、セキュリティの確保は最重要課題のひとつです。同社 SOC（Security Operation Center）は、そうした中で外部からの攻撃や内部不正などのリスクに対する仕組み作りや、ツールの開発などを行う部門。その最新の取り組みである「次世代セキュリティ DWH」について、開発を主導した日比野 恒氏に話を伺いました。

**利用しているサービス**

BigQuery、Cloud Data Fusion、Cloud Storage、Dataproc、VPC Service Controls、Looker など

RECRUIT

## 株式会社リクルート

https://www.recruit.co.jp

〒100-6640
東京都千代田区丸の内 1-9-2
グラントウキョウサウスタワー

「Follow Your Heart」のビジョンのもと、人材領域と販促領域からなるマッチング＆ソリューション事業を主軸に幅広いサービスを展開。新規事業も積極的に開拓し、常に新しい領域に挑戦し続けている。現在の従業員数は 17,327 名（2022 年 4 月 1 日現在 / アルバイト・パート含む）。

**インタビュイー**

セキュリティ統括部 セキュリティ戦略グループ
日比野 恒 氏

## 専用のソリューションではなく、BigQuery を中心にログ分析基盤を実現

企業内でセキュリティを管轄する部署のメンバーが正しい施策を行うためには、今何が起きているのかを、ログなどのファクトによって正確に把握する必要があります。約 3 年間、SOC でそのための仕組み作りに携わってきた日比野氏（現在はより高次なセキュリティ戦略を立案するセキュリティ戦略グループに在籍）は、同社の新たな武器となる次世代セキュリティ DWH のコンセプトについて次のように説明します。

「ログを使ったセキュリティ監視ツールというと、多くの人が SIEM（Security Information and Event Management）と呼ばれる、リアルタイムにインシデントを検知するツールをイメージすると思うのですが、今回、われわれが開発したのは、ログをもとに攻撃の兆候を見つけだしたり、起きてしまったインシデントを後追いでフォレンジックしたり、レポーティングしたりといった、監視以外の目的でログを有効活用するための仕組みです。特にこだわったのが、専門のエンジニアではない人でもログを使った分析業務を行えるようにすることでした。」

従来、そうした分析は担当者がログ分析のスキルを持つ SOC のエンジニアにオーダーして、出力してもらったレポートをもとに行われていました。日比野氏は、そのやり方では技術力の高いエンジニアに負荷が集中したり、担当者がリアルタイムな分析結果を取得できなかったり、レポートからは得られない気付きを見落としてしまうといった問題があったと指摘します。

そして、これらの問題を解決するための次世代セキュリティ DWH を構築するにあたり、日比野氏は既存のセキュリティ向けソリューションを導入するのではなく、BigQuery など、汎用のデータ分析ソリューションを組み合わせて実現することを選択しました。そこに多くのメリットがあると考えたからです。

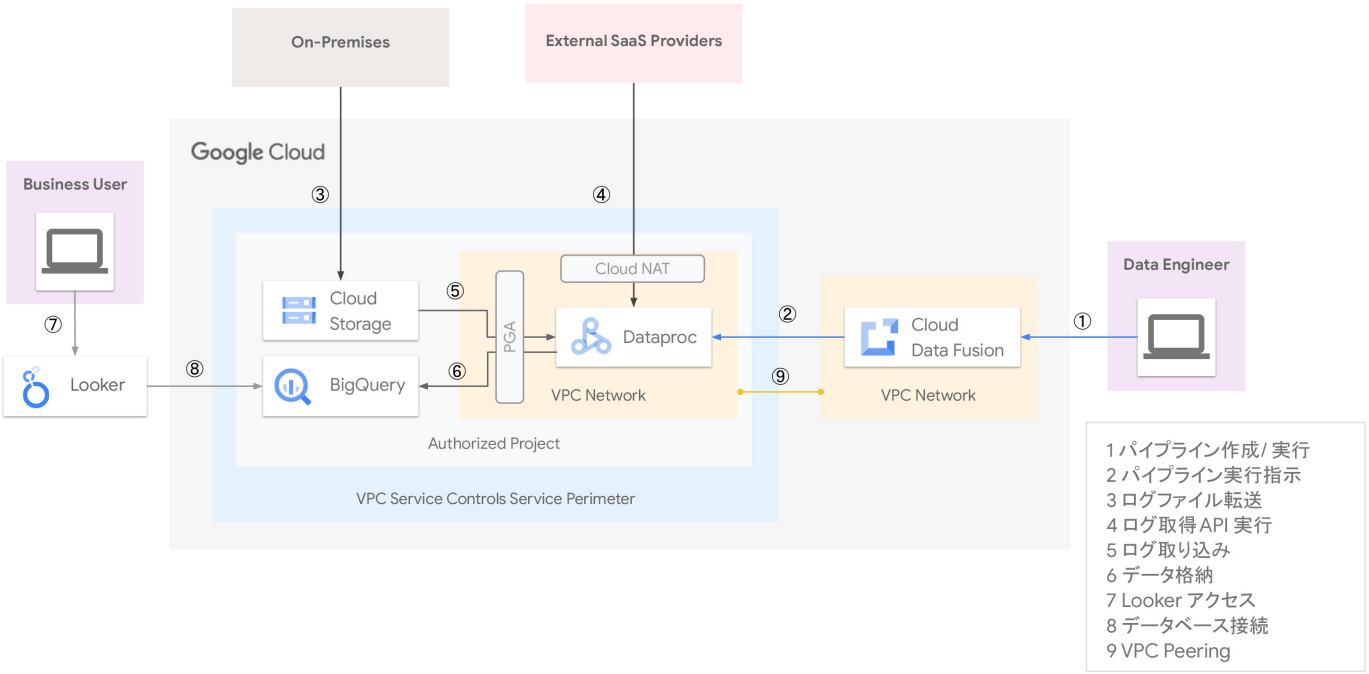
「第一にコストですね。セキュリティ向けのログ分析ソリューションは、取り込んだログ容量（GB）による課金やインスタンスの稼働時間（動いている間）で課金されてしまうライセンス体系でストレージも割高なため、コスト削減がしにくい問題がありました。ここにクエリ課金で使える DWH サービスを使えれば、ランニングコストを大きく削減できるのではないかと考えたのがこの取り組みのスタート地点です。そして、クエリ課金のプロダクトをパフォーマンス性、ストレージ コストの両面で評価し、BigQuery を選定しました。」（日比野氏）

その上で日比野氏は、BigQuery のような DWH サービスの採用には費用面以外にも大きなメリットがあると言います。

「セキュリティに特化したソリューションではログデータから分析に必要な情報を検索、抽出する際に、専用言語でクエリを書くことを求められるため、その学習コストが人材育成、確保において大きなハードルになっていました。そこで、次世代セキュリティ DWH では、多くのエンジニアが長年慣れ親しんできた SQL を中心に技術スタックを構成し、人材育成・採用の敷居を下げることを目指しています。これは今回のプロジェクトが始まる前から温めていたアイデアなのですが、当時と比べても SQL の利用シーンは大きく拡大しており、読みが当たったなと思っています。」

# ローコードな Cloud Data Fusion で分析業務の敷居を下げる

## Architecture: Recruit セキュリティ DWH




上図は、昨年 10 月から実運用が開始された次世代セキュリティ DWH のシステム構成図。BigQuery を中心にログを蓄積する Cloud Storage や、データを可視化するための Looker などを連携させることで分析業務に必要な機能を実現しています。なお、パイプラインの実行環境については Dataproc を利用。VPC Service Controls を駆使することで BigQuery や Cloud Storage とセキュアにデータをやり取りできるように工夫しました。

「プロジェクトが具体的に動きだす以前から BigQuery や Looker に関しては検証していたため、その導入に大きな苦労はありませんでした。しかし、ログデータを BigQuery に蓄積していく際の ETL（Extract = 抽出 / Transform = 変換 / Load = 格納）についてはかなりの試行錯誤がありました。当初は主にコストの観点から Cloud Composer を使って Python ベースで処理していくことを想定していたのですが、それだと開発や保守に高度な専門知識と経験が求められるようになり、本来の目的であった分析業務の敷居を下げられなくなってしまいます。そこで、フルマネージドのデータ統合サービスである Cloud Data Fusion を使うことでローコードな ETL ツールを実現。Cloud Storage にアップロードされたオンプレミス環境のログを Cloud Data Fusion で ETL 処理しているほか、SaaS や IaaS など、外部サービスのログについても API 経由でクローリングして Cloud Data Fusion 経由で BigQuery に入れる処理を行っています。」（日比野氏）

Cloud Data Fusion の導入に際しては、いくつかの想定外の苦労があり、手直ししながら安定させていく必要があったと当時を振り返る日比野氏。

「その際、Google Cloud の皆さんから適切なサポートを得られたことがありがたかったですね。一緒になって解決策を検討していただいたり、パイプラインの設計に関してアイデアをいただけたりしたことにとても感謝しています。おかげで現在はかなり安定したデータ処理が行えるようになりました。」

現在の次世代セキュリティ DWH の活用状況について、日比野氏は次のように語ります。

「次世代セキュリティ DWH の活用を促すため、利用している現場部署の定例会などに参加させてもらっているのですが、従来は SOC に頼まなければ出せなかったような分析データを、自らその場で Looker を活用して出せるようになったインパクトは大きく、これまでの業務フローを変える動きも起き始めています。また、そうした分析を通じて新たなインサイトが得られること、その気付きをその場でスピーディにかたちにできるというところが現場に新たな価値を生み出していることを実感しています。まだ運用開始から 1 年程度ではありますが、想像以上の手応えを感じているところです。」

もちろん、今後も次世代セキュリティ DWH はさらなる進化を遂げていく予定。増え続けるパイプラインをより効率的に運用していくために、現在は GUI ベースとなっているインターフェースをコードベースで管理できるようにするなど、大規模化を意識した機能改善を行っていくとのことです。

「Google Cloud の魅力は、エンジニア、デベロッパー向けの分析基盤やツールが充実しているところ。今回の取り組みを通じ、そうした部分がわれわれのビジネスにも有用だということが確認できました。昨今、働く人々の IT 環境が大きく変わりつつある中、利便性や生産性を損なわずにセキュリティを高めていくためにどのようなことができるのかが大きなテーマとなっています。Google Cloud は以前から『BeyondCorp Enterprise』というゼロトラスト ソリューションを出すなど、私個人としても注目しているところです。今後はそれらをわれわれの業務においてどのように活用していけるのかを深掘りしていきたいですね。また、時代の変化という点ではオンプレミスに持っているデータを、より電源効率の良いクラウド上に移行することでカーボン ニュートラルにつなげていけるのかといった検討もしていかねばならないと感じています。そうしたサステナビリティの観点でも Google Cloud には期待をしています。」（日比野氏）

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、イノ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から
https://goo.gl/CCZL78

Google Cloud

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。