



Session Report

クラウド利用コストを最適化し
エンジニア自身が管理できるようにする「FinOps」

Google Cloud FinOps ツールによるコスト管理

Google Cloud
パートナー エンジニア
梶沢 直樹

Google Cloud

セッションレポート概要

Google Cloud では、クラウド利用時のコストを管理するためのテクノロジーへの投資を継続的に行っています。ここでは、新しくリリースした FinOps ツールにより、お客様がクラウドのコストを理解し、割り当てたり、効率的にコスト管理を行ったりするための方法を紹介します。

プレゼンター紹介



Google Cloud パートナー エンジニア
枘沢 直樹

Google Cloud で、パートナー エンジニアとしてインフラ モダナイゼーション領域のパートナー支援、テクノロジー パートナーとのエンゲージを担当しています。

目次

- クラウド利用時のコスト管理を手助けする「Google Cloud FinOps」 3
- より細かい課金情報が分かる「可視化」 6
- 推奨事項をレコメンドする「最適化」 8
- エンジニア自身でプロジェクトのコストを管理できる「実行」 12

クラウド利用時のコスト管理を手助けする「Google Cloud FinOps」

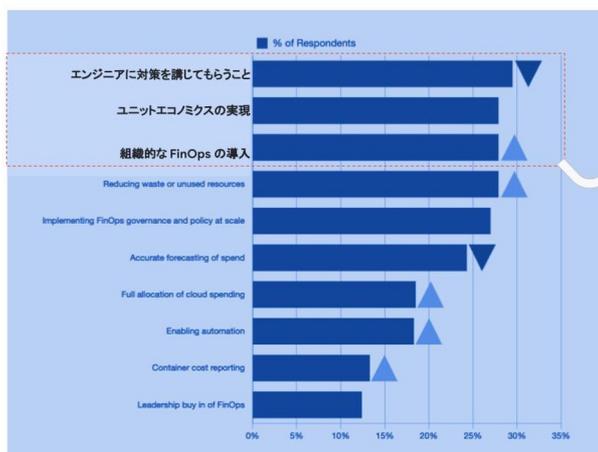
クラウドサービスを利用するときに、いかにコスト管理を行えばよいかと聞かれることがよくあります。例えば、「開発チームが想定していないリソースを使ってしまい、1カ月分の予算を3日で使ってしまった」とか、「ある一定の期間の利用料が想定以上の高額だった」といった話です。コスト管理の観点からクラウドを利用することを、不安に思っている開発担当者もいます。

FinOps の規律を推進する企業や、認定された実践者で構成される非営利団体の「FinOps Foundation」では、「財務の実務担当者が直面する課題」と「FinOps 機能の優先事項」に関する調査を行っています。これまでの開発では、財務部門や調達部門がコスト管理を行い、開発部門が予算申請をして、予算が通れば IT を調達し、エンジニアがリソースを使って開発するという形が一般的でした。つまり、エンジニア自身がコスト管理をする必要性が薄かったということです。

一方、クラウドを利用した開発では、開発中に追加のリソースが必要になると、オンデマンドでリソースを追加できるので、エンジニアもコストを把握した上で、リソースを適切に選択することが求められます。そうした中で Google Cloud では、クラウド利用におけるコスト管理の課題を解決できる FinOps ツールを提供しています。

クラウド利用におけるコスト管理は難しい？

財務の実務担当者が直面する課題



FinOps 機能の優先事項



出典: FinOps Foundation

Google Cloud Next '23

Proprietary

06

財務担当者が抱える課題と FinOps に求める機能の優先事項

FinOpsとは、データに基づいて支出を決定するときに、エンジニアリング、財務、テクノロジー、ビジネスチームが協力し、組織がビジネス価値を最大化できるようにするためのクラウド財務管理の規律であり実践でもあります。

そもそもクラウド利用時のコスト管理に関する課題を解決するには、財務部門が組織全体のクラウドコストの全体像を把握すること、エンジニアリング部門がリソースの調達を自由に行いながら、自身でコストを最適化できるようになることが重要です。

Google Cloud FinOps が目指しているのは、下図にある通り、コストの可視化、予算の設定、コストの最適化、利用料の予測、利用料の計画と実績の追跡の5つです。

Google Cloud FinOps が目指すこと

- | | |
|---|--|
|  コストの可視化 | ● コストを正確に把握して、ビジネス価値の担保とガバナンスを確保する |
|  予算の設定 | ● 組織全体のコスト可視化とKPIに基づくクラウドリソースの最適化の指針を立てる |
|  コストの最適化 | ● コスト最適化の方法が推奨内容として提供される |
|  利用料の予測 | ● 予算編成、予測を行い、コスト効率の高い開発手法を選択できる |
|  利用料の計画と実績の追跡 | ● クラウドコストの管理と資産を行う統合ツールによる追跡ができる |

Google Cloud Next '23

Proprietary

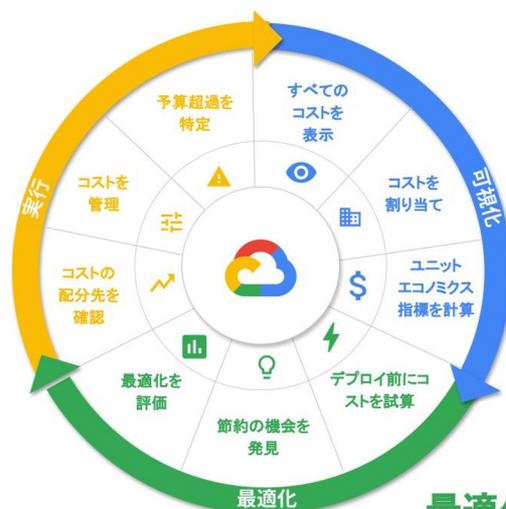
010

Google Cloud FinOps が実現を目指す5つのポイント

Google Cloud FinOps のフレームワークは、下図の通りで、「可視化」によりクラウドの全コストを把握して、より詳細なコスト管理を行い、「最適化」でコスト最適化のポイントを定期的に把握し、「実行」で将来のコストの「想定外」をなくします。可視化、最適化、実行の3つのサイクルを回すための各種ツールを提供しています。

Google Cloud FinOps のフレームワーク

実行
将来のコストの
「想定外」をなくす



可視化
クラウドの全コストを
把握して、より詳細な
コスト管理を行う

最適化
コスト最適化のポイントを定
常的に把握できる

「可視化」「最適化」「実行」のサイクルでビジネス価値を最大化

より細かい課金情報が分かる「可視化」

Google Cloud FinOps における「可視化」では、よりきめ細かな Billing エクスポートが可能になります。これまで Google Cloud を利用すると、請求先アカウントをプロジェクトに紐づけることで、プロジェクトで使っている、例えば「BigQuery」や「Compute Engine」、「Cloud Storage」などのサービスの利用料が課金されていました。

このとき、例えば「BigQuery でいくら」など、サービス単位では課金情報が可視化されていましたが、そのほかにも「このジョブでいくら」など、より細かい課金情報を知りたいという要望がありました。そこで現在では、よりきめ細かい Billing エクスポートができるようになっていました。この機能は BigQuery 以外にも、「Cloud Spanner」や「Cloud Firestore」、Compute Engine で利用できます。最新の情報はリリースノートで公開されるので、ぜひ参照してください。

よりきめ細かい Billing エクスポート

NEW

例 BigQuery への詳細な Billing エクスポート

リソースレベルでのアクティビティをより詳細にコスト情報として反映

ユニットエコノミクスと全コストの割り当て

2、7

FinOps のベストプラクティスを導入するための課題上位

出典: FinOps Foundation

コストの割り当てとユニットコストの測定

1、11

FinOps 機能の優先事項上位

Google Cloud Next '23

以前 - Billing エクスポート

時間	サービス	SKU	プロジェクト	使用状況	コスト
12:00 ~ 13:00	BigQuery	Analytics	my-project-id	1000 バイト	\$0.10

現在 - よりきめ細かい Billing エクスポート

時間	サービス	SKU	プロジェクト	リソース	使用状況	コスト
12:00 ~ 13:00	BigQuery	Analytics	my-project-id	global_name: //bigquery.googleapis.com/projects/my-project-id/jobs/job-1	200 バイト	\$0.02
12:00 ~ 13:00	BigQuery	Analytics	my-project-id	global_name: //bigquery.googleapis.com/projects/my-project-id/jobs/job-2	300 バイト	\$0.03
12:00 ~ 13:00	BigQuery	Analytics	my-project-id	global_name: //bigquery.googleapis.com/projects/my-project-id/jobs/job-3	500 バイト	\$0.05

Proprietary

014

リソースレベルでのコスト情報をより詳細に反映

Google Cloud とそのほかのパブリック クラウドを併用している場合、サービスごとに課金体系が異なります。そこで FinOps Foundation では、FOCUS (FinOps Open Cost & Usage Specification) と呼ばれるフレームワークを提供しています。FOCUS は、Google Cloud だけでなく、各クラウド サービスを利用するときの請求管理を標準化するためのフレームワークです。

最大の目的は、課金データのオープン化です。まだ発展途上ではあるものの、今後クラウドサービスの横断利用における共通言語をつくることにより、より容易に課金を可視化することを目指しています。いわゆるコストデータのためのオープン規格として FOCUS があり、Google Cloud は FinOps Foundation の推進メンバーにも参画しています。

また Google Cloud では、Billing において「Duet AI」が使えます。Duet AI は、Google Cloud の各サービスと連携して、サービスを使う上での FAQ を提供してくれます。生成 AI について、Google Cloud は以前からも取り組んできており、大規模言語モデル (LLM) 自体にビジネスと技術的な知見を橋渡ししてきました。この仕組みを課金管理でも使えるようにするのが Duet AI との連携です。

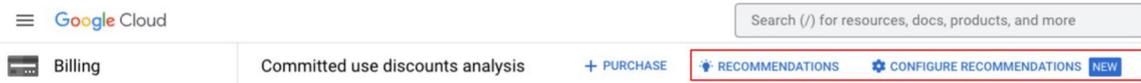
Duet AI の AI アシスタント機能は、Google Cloud コンソールに組み込まれています。Duet AI に問い合わせをすることで、課金管理を最大限に活用できる最適な答えを返してくれます。その精度は、FinOps Foundation が実施しているベンチマーク テストにも合格するほどです。Billing では正確性が求められますが、利便性の向上や疑問の解消に Duet AI が有効です。

推奨事項をレコメンドする「最適化」

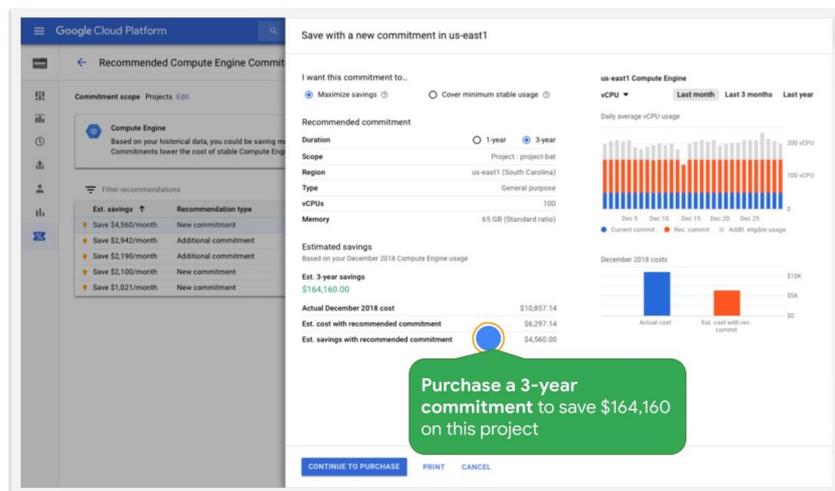
次に Google Cloud FinOps における「最適化」では、確約利用割引の条件を設定する機能である CUD (Committed Use Discounts) Recommender が 2023 年 8 月初めにパブリックレビューとして公開されています。

CUD は、特定の期間（現時点で 1 年間 または 3 年間）、リソースや一定金額の利用をコミット（確約）することで割引を受けられる仕組みです。割引率のコミットする期間と対象のサービスによっては最大 50% を超える割引率を提供するサービスもあり、効率的に運用することで大幅に総コストを引き下げることが期待できます。

CUD Recommender



- 定常分析に基づく購入量のご案内
- 安定した使用法と最適な推奨事項により、予想される効果 (割引額) を定量化します。
- **vCPU とメモリで利用可能。**
SSD、GPU、ライセンスは対象外
- **過去 30 日間**の使用履歴に基づくデータから推奨事項をご案内



Google Cloud Next '23

Proprietary

018

分析結果から購入量をレコメンド

CUD では、Compute Engine がリソースベースの CUD と支払ベースの CUD を利用できます。Compute Engine、「Cloud SQL」、Cloud Spanner、「VMware Engine」、「Cloud Run」、「Google Kubernetes Engine (GKE)」、「AlloyDB for PostgreSQL」が支払ベースの CUD を利用できます。

例えば、Compute Engine のリソースベースの CUD ではリージョンと利用するマシンタイプを選択し、CPU コアとメモリに対して使用量をコミットします。割引率は1年（最大 37%）、または3年（最大 55%）というコミット期間によって変わります。割引の適用範囲は、デフォルトはプロジェクト単位ですが、割引の共有をすれば請求先アカウント単位に変えられます。複数のプロジェクトを1つの請求先アカウントで共有することもできます。

一方、支払いベースの CUD は、時間あたりの利用料なので、リソースベースよりも柔軟性があります。割引率は1年（28%）または3年（46%）。割引の適用範囲は請求先アカウントのみです。CUD はあくまでも利用料の割引であり、リソースの確約ではありません。CUD には今後も、さまざまな新機能が追加される予定です。

GCE の Resource-Based CUD と Flexible CUD

	GCE Resource-Based CUD	GCE Flexible CUD (Spend-based)
コミット対象	リージョン + Machine Type (N1, N2, N2D など) ごとのリソース※ ※現時点では GCE の CPU Core と Memory のリソースのみ	時間当たりの利用料金 [\$ /h]※ ※現時点では GCE の N1、N2、N2D、E2、C2、C2D の CPU Core、Memory の費用のみ
コミット期間 (割引率)	1 Year (up to 37 %) or 3 Years (up to 55 %)	1 Year (28 %) or 3 Years (46 %)
スコープ	Project (Default) or 請求先アカウント ※ Cloud 請求先アカウントから「割引の共有」を有効にして、複数のプロジェクトで割引を共有する (ScopeをBilling Accountに切り替える) ことが可能だが、購入している CUD すべての単位でしか切り替えができない (購入した CUD 個別での切り替えは不可)	請求先アカウント

Google Cloud Next '23

Proprietary

020

Compute Engine のリソースベース/支払いベースの CUD 比較

参考までに、Google Cloud には、継続利用割引(SUD)という自動的に適用される割引もあります。Google Cloud では、オンデマンドでの利用時に、請求月内の期間の 25%を超えて使用し、ほかの割引が適用されていないリソースに対して継続利用割引 (SUD) が自動的に適用されます。割引を最大限に享受するには、下図の順序で購入の検討をすることが理想的です。

ディスカウントの適用される順序

割引を最大限享受するには、#1 -> #2 の順で購入検討をすることが理想的



理想的な購入方法で割引を最大限に享受できる

「コミットの期間をどのように切り替えれば、どれくらい効果を得られるのか」を確認するためのベースとなる設定をするのが CUD Recommender Configuration です。コミットの期間、コミットするしきい値を設定することで、どれだけコストに差が出るかを画面で確認できます。

また Google Cloud コンソール上には、新たに FinOps Hub のタブがプレビューで追加されました。FinOps Hub は、Google Cloud の最適化を一覧できる画面で、特定のサービスや推奨事項を迅速に把握できます。これによって、放置されているプロジェクトやコストに影響する可能性のあるプロジェクトを見つけることができます。

FinOps Hub NEW

FinOps 専用の Google Cloud 全体の最適化を一元管理できるコンソール

- 取り組む優先順位
- 効率的な最適化

利用されていないリソースの削減

4

FinOps のベストプラクティスを導入するための課題上位

FinOps 文化の確立とリソース利用とサイズの最適化

2, 3

FinOps 機能の優先事項上位 2 項目

出典: FinOps Foundation
Google Cloud Next '23

Potential savings/ month	Service	Description
Save upto \$432.33/ month	Compute Engine	Purchase a 3 year new Compute Engine Flexible CUD
Save upto \$202.17/ month	Cloud Spanner	Purchase a new 3 year or 1 year CUD for Cloud Spanner
Save upto \$432.33/ month	Compute Engine	Purchase a 3 year new standard CUD for Gen purpose A2 mem...
Save upto \$432.33/ month	Cloud Run	Purchase an additional 3 year CUD for Cloud Run
Save upto \$432.33/ month	Kubernetes Engine	Rightsize instance 'optimal-cud-test-9lp' to the machine type e...
Save upto \$432.33/ month	Cloud SQL	Shut down idle instance 'org-policy-test'

Google Cloud 全体の最適化を一元管理できる

FinOps Score では、Google Cloud ツールを利用して、AI によるコストのモニタリングと最適化が継続的に行われているかの評価が可能です。ピア ベンチマークは、同業他社に対する指標による相対的な評価を提示します。指標は、業界、セグメント、地域、Google Cloud の使用状況などの要素に基づき算出されます。ピア ベンチマークは、デフォルトでオプトインされていますが、いつでもオプトアウトできます。

Recommendations は、Google Cloud の AI を利用した、コスト推奨に関する機能の多くを FinOps Hub へ集約しています。例えば、放置されているプロジェクトや Compute Engine のサイズを最適化したり、Compute Engine のアイドル状態のインスタンス、Cloud SQL のアイドル状態、およびオーバープロビジョニングされているインスタンス、Cloud Run の CPU 割り当てなどを可視化したりできます。これにより、コストの最適化が可能になります。

CUD Recommender によるトレードオフ分析では、CPU のカバー率やコミットの選択によって、どれくらいコストを最適化できるかを可視化することで、どの手段が最もコスト効果が高いかを判断できます。最適化のサマリでは、FinOps Hub が最適化 ROI の追跡を継続し、適用した最適化に基づいて節約総額を表示できます。

エンジニア自身でプロジェクトのコストを管理できる「実行」

最後に、Google Cloud FinOps の「実行」に関して紹介します。請求先アカウントにアクセスするとき、Google Cloud のプロジェクト ユーザーは追加の権限を求められることなく予算を作成し、クラウド支出を把握できます。これにより、エンジニア自身によるプロジェクトのコスト管理が可能になります。この機能もパブリックレビューですが、ぜひ使ってみてほしい機能です。

エンジニア (プロジェクト ユーザー) 向け予算

NEW

Google Cloud のプロジェクト ユーザーは、Cloud 請求先アカウントにアクセスするための追加の権限を必要とせずに予算を作成してクラウド支出の把握が可能

- エンジニアによる自身のプロジェクトのコスト管理が可能になる

プロジェクトオーナー/プロジェクト編集者

🔧 コンソール内 (プロジェクトレベルの [Cost Reports] ページ) または API で単一のプロジェクトの予算を作成/編集可能

プロジェクト閲覧者

🔄 個々のプロジェクトに限定された予算を表示

プロジェクトオーナー

📧 単一のプロジェクトの予算通知を受信

The screenshot shows the Google Cloud Billing console interface. On the left, there's a navigation menu with 'Billing account', 'Overview', 'Cost management', 'Reports', 'Budgets & alerts', 'Cost optimization', 'Cost estimation', 'Billing management', and 'Account management'. The 'Budgets & alerts' section is highlighted. The main content area shows a table of budgets and alerts. A modal window is open, displaying a 'Billing Budget Alert' for 'fighting piracy staging' with a message: '25% of budget expected to be reached'. The alert details include the budget name, project name, project ID, budget amount (\$100.00), and budget period (May 1, 2023 - May 31, 2023).

Google Cloud Next '23

Proprietary

エンジニア自身がプロジェクトのコストを管理できるように

Google Cloud FinOps ツールによるコスト管理の課題解決のポイントは、財務部門で組織全体のクラウドコストの全体像を把握するだけでなく、エンジニアリング部門がリソースの調達を自由に行いながらコストを最適化できることです。

FinOps を採用することの意義は、「コスト最適化=コスト削減」ではなく、いかに効率的にビジネスに直結するサービス リソースを活用するか、効率的に利用するかです。データに基づいて、エンジニアと管理部門が協力し、一緒にビジネスを創出していくためにも FinOps の各種ツールを活用していただきたいと思います。

今後も新しい機能やサービスが順次拡張・追加される計画なので、ぜひ注目してください。

参照リンク

1. [Cloud FinOps の概要](#)
2. [Google Cloud FinOps ツールによるコスト管理 アーカイブ動画視聴ページ](#)

製品、サービスに関するお問い合わせ



goo.gl/CCZL78

Google Cloud の詳細については、上記 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 2024 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。