

Google Cloud UPDATES: 2023 年第2回 Compute / Database 編

2023-2-27

Google Cloud

本日のスピーカー



Wataru Inoue
Customer Engineer,
Google Cloud



Minsoo Jun
Customer Engineer,
Google Cloud



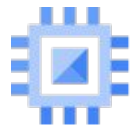
Yuya Ono
Customer Engineer,
Google Cloud



Yoshimasa Kataoka
Customer Engineer,
Google Cloud

01

インフラストラクチャ



新しいマシンファミリー (Tau T2A / M3)

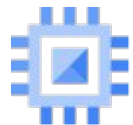
スケールアウト最適化: Tau T2A

- 初めての ARM アーキテクチャ ベースのマシン
- 最大 48 vCPU をサポート
- 最大帯域 32 Gbps をサポート

メモリ最適化: M3

- SAP HANA などのワークロードに最適化されたマシンタイプ
- 最大 128 vCPU / 1,952 GB メモリーをサポート
- 最大帯域 100 Gbps をサポート

ワークロード タイプ					
汎用のワークロード		最適化されたワークロード			
コスト最適化	バランス	スケールアウト最適化	メモリ最適化	コンピューティング最適化	アクセラレータ最適化
E2	N2、N2D、N1	Tau T2D、Tau T2A	M3、M2、M1	C2、C2D	A2
低コストで日々のコンピューティングを実現	幅広いマシンタイプにわたるバランスの取れた価格とパフォーマンス	スケールアウトワークロードに最適なパフォーマンスと費用	超高メモリ ワークロード	超高パフォーマンスでコンピューティング負荷の高いワークロードを実現	ハイパフォーマンスコンピューティングワークロード向けに最適化
<ul style="list-style-type: none"> ウェブサービス アプリの配信 バックオフィスアプリ 小規模データベース マイクローサービス 仮想デスクトップ 開発環境 	<ul style="list-style-type: none"> ウェブサービス アプリの配信 バックオフィスアプリ 中規模データベース キャッシュ メディア / ストリーミング 	<ul style="list-style-type: none"> スケールアウトワークロード ウェブサービス コンテナ化されたマイクロサービス メディアのコード変換 大規模 Java アプリケーション 	<ul style="list-style-type: none"> 中規模 OLAP およびインメモリ データベース (SAP HANA など) インメモリ データベースとインメモリ分析 Microsoft SQL Server などのデータベース ゲノム モデリング 電子設計自動化 	<ul style="list-style-type: none"> 計算依存型ワークロード 高パフォーマンスのウェブサービス ゲーム (AAA ゲームサーバー) 広告配信 ハイパフォーマンスコンピューティング (HPC) メディアのコード変換 AI / ML 	<ul style="list-style-type: none"> CUDA 対応の ML トレーニングと推論 HPC 超並列コンピューティング



A2 Ultra : アクセラレータ最適化の新しいマシン タイプ

A2-ultra GPU VM インスタンス

New Launch

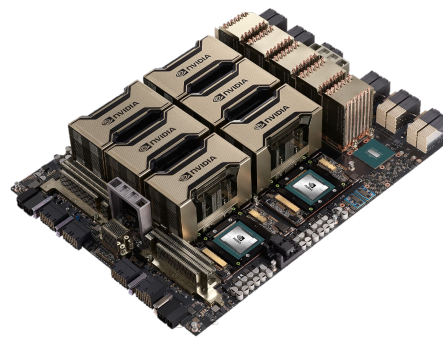
Google Cloud で最も速く大きいGPUメモリ

- 80 GB HBM2E*
- 2 TBps もの GPU メモリ帯域

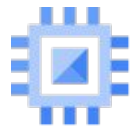
複雑で巨大なモデルに最適化

- 2倍の GPU メモリ - 80GB
- 最大3倍のスループット増加(DLRM)
- 最大25%のスループット増加(推論)
- 最大2倍高速な HPC シミュレーション






* high bandwidth memory

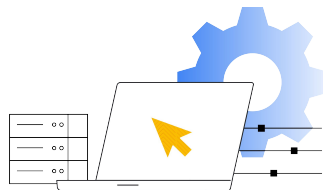


NVIDIA Ampere A100
Tensorcore GPU ベース



Flex CUD (フレキシブルな確定利用割引)

-  VM ファミリーやリージョンを **フレキシブル**に変更可能
-  **予測可能な利用**に最適
-  ほとんどのVMファミリー (N1, N2, N2D, E2, C2) の一時間ごとの累計に **割引を適用**
-  マシンサイズや OS に**限定されません**。また、GKE や Dataproc にも適用されます。
-  同じ課金アカウント下のプロジェクト間で、**デフォルトで共有**されます



1年コミットメント

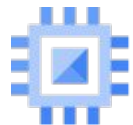
割引: **最大 28%**

通常の CUD (37%)

3年コミットメント

割引: **最大 46%**

通常の CUD (55%)



コミットメントの統合

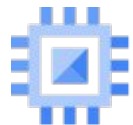
- マシン ファミリー、リージョン、期間などが同一である複数のコミットメントが統合可能
- 開始 / 終了日が統一されることでコミットメントの管理工数が削減

	Source Commit 1	Source Commit 2
Type	N2	N2
Region	asia-northeast1	asia-northeast1
Resource	<ul style="list-style-type: none">● vCPU : 100● RAM : 100 GB	<ul style="list-style-type: none">● vCPU : 200● RAM : 300 GB
Term	3 Year	3 Year
Start Date	2021 / 01 / 01	2021 / 12 / 01
End Date	2024 / 01 / 01	2024 / 12 / 01



Merge

	Target Commit
Type	N2
Region	asia-northeast1
Resource	<ul style="list-style-type: none">● vCPU : 300● RAM : 400 GB
Term	3 Year
Start Date	2023 / 02 / 27 (統合日)
End Date	2024 / 12 / 01



コミットメントの分割

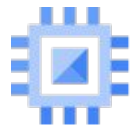
- 単一のコミットメントを複数のコミットメントに分割可能
- 小さい単位のコミットメントを作成することで細かい粒度での管理が可能（自動更新など）

	Source Commit
Type	N2
Region	asia-northeast1
Resource	<ul style="list-style-type: none">● vCPU : 200● RAM : 200 GB
Term	3 Year
Start Date	2021 / 01 / 01
End Date	2024 / 01 / 01



Split

	Source Commit	Split Commit
Type	N2	N2
Region	asia-northeast1	asia-northeast1
Resource	<ul style="list-style-type: none">● vCPU : 150● RAM : 100 GB	<ul style="list-style-type: none">● vCPU : 50● RAM : 100 GB
Term	3 Year	3 Year
Start Date	2021 / 01 / 01	2023 / 02 / 27 (分割翌日)
End Date	2024 / 01 / 01	2024 / 01 / 01



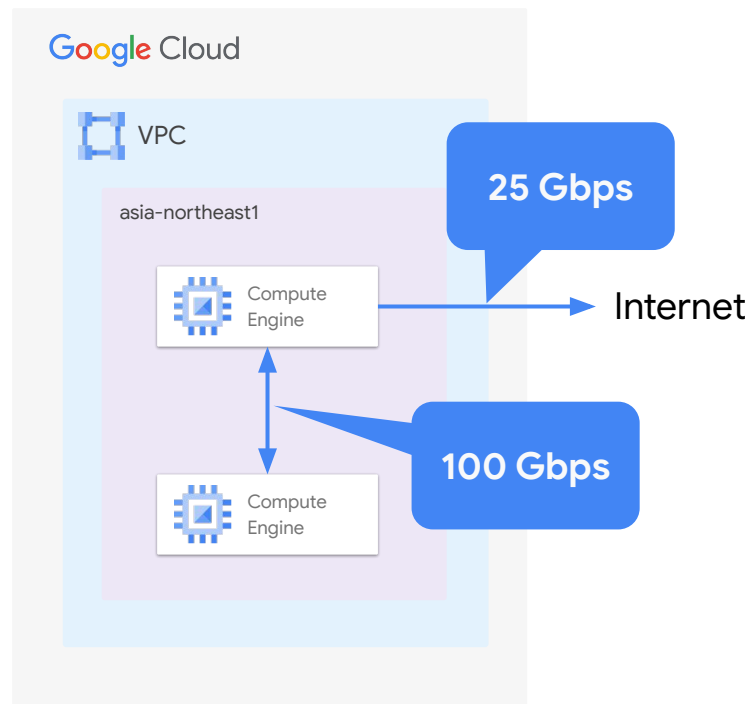
Tier_1 ネットワーキングの帯域幅上限の増加

Tier_1 ネットワーキングの定義

- 30 vCPU 以上を搭載した N2 / N2D / C2 / C2D マシン
- gVNIC ネットワークドライバを利用
- gVNIC 互換のイメージを利用

外部 IP を利用した Egress 通信の帯域幅上限が増加

- 今までの 7 Gbps から **25 Gbps** へ増加
- Tier_1 ネットワーキングが利用可能なマシンファミリーの中で **N2 / N2D** が対象





永続ディスクのパフォーマンス ベースライン

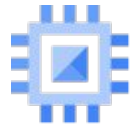
永続ディスク (Persistent Disk) のパフォーマンスは次の要素に依存して決定

- ディスク サイズ
- ディスクがアタッチされた VM の vCPU 数

Balanced / SSD 永続ディスクでパフォーマンスのベースラインが策定され、容量の小さいディスクでも一定のパフォーマンスを得ることが可能

実際のパフォーマンスは (ベースライン + ディスク サイズや vCPU に応じた値) となる

インスタンスあたりの ベースライン	Balanced (Zone)	SSD (Zone)
IOPS	3,000	6,000
スループット (Mbps)	140	240



インスタンスの稼働時間上限の設定

Compute Engine インスタンスの稼働時間に制限を設けることでインスタンスが自動的に停止、または削除される

稼働時間の上限を設定することでコストを最小限に抑えられ、またクォータの開放が可能になるため、一時的なワークロードに最適 (Spot VM でも利用可)

可用性ポリシー

VM プロビジョニング モデル

標準

[スポット] を選択すると、プリエンプティブル VM を割引料金で利用できます。それ以外の場合は [標準] を使用します。 [Learn more](#)

VM の制限時間を設定する ?

制限時間の種類
時間

時間を入力*
12

VM の終了時

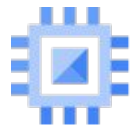
停止

Choose what happens to your VM when it's preempted or reaches its time limit

ホスト メンテナンス時

VM インスタンスを移行 (推奨)

Compute Engine が定期的なインフラストラクチャ メンテナンスを実施する際、ダウンタイムを発生させずに VM インスタンスを他のハードウェアに移行できます



インスタンスの名前変更

Compute Engine インスタンスの作成後にもインスタンス名を変更することが可能
(変更時にはインスタンスが停止している必要があります)

ユースケース

- ワークロードをより適切に表す
- 組織の命名規則を遵守する

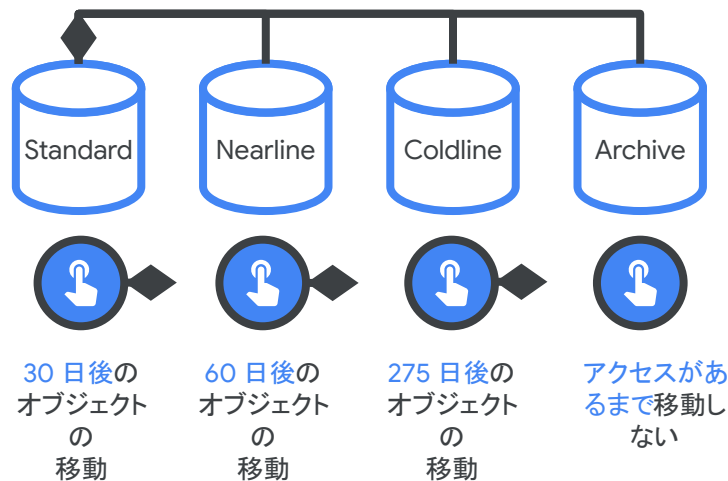
```
# インスタンスの停止
$ gcloud compute instances stop <インスタンス名>

# インスタンス名の変更
$ gcloud compute instances set-name <インスタンス名> \
  --new-name=<新しいインスタンス名>
```



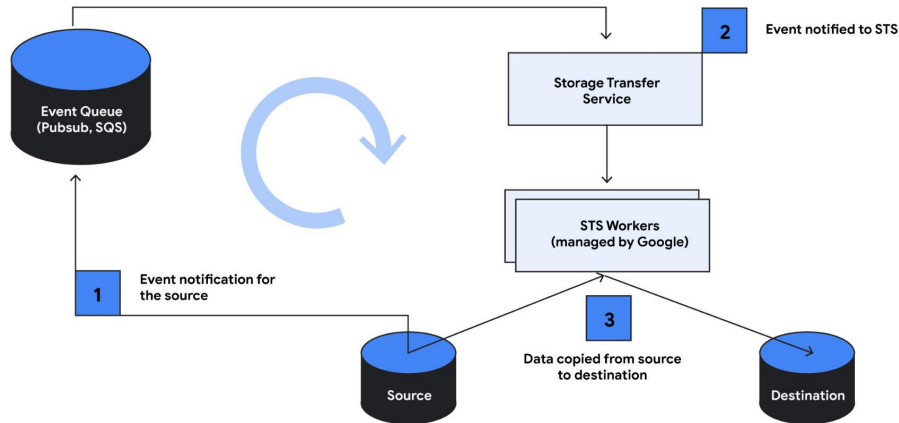
Cloud Storage の Autoclass 機能

- Cloud Storage に保存されたオブジェクトへのアクセスパターンに基づいて自動で適切な Storage Class へ移動
- 各オブジェクトを適切な Storage Class へ自動で移動することで管理コスト、利用コストの削減が可能



Storage Transfer Service のイベントドリブンのデータ転送

- Cloud Storage や S3 から Cloud Storage へのデータ転送がイベントドリブンで実施可能
- 今までは最短でも 1 時間に 1 度のバッチ処理だったためにリアルタイムでのデータ転送は不可だった
- イベント (データの変更) 検知は PubSub (Cloud Storage) / SQS (S3) 経由で実施



02

ネットワーク



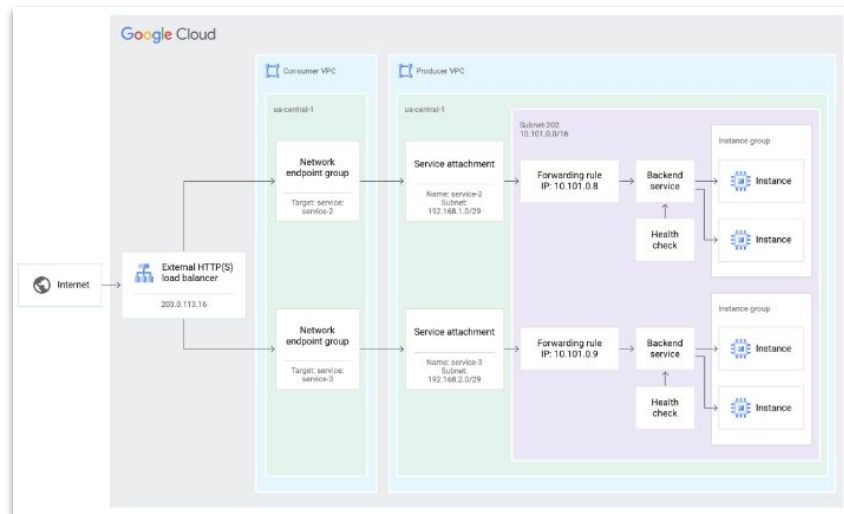
コンシューマ HTTP(S) サービス制御での PSC アクセスが GA

Private Service Connect によるサービスの利用方法

- Private Service Connect エンドポイント
- **コンシューマ HTTP(S) サービス制御 (GA)**
 - グローバル外部 HTTP(S) LB の利用のみ GA
 - リージョン外部 / 内部 HTTP(S) LB の利用はプレビュー

コンシューマ HTTP(S) サービス制御の特徴

- ロードバランサー経由でマネージド サービスへアクセス可能
- パスベース (v1 / v2 など) でのアクセス制御やカスタムドメインによるアクセスが可能



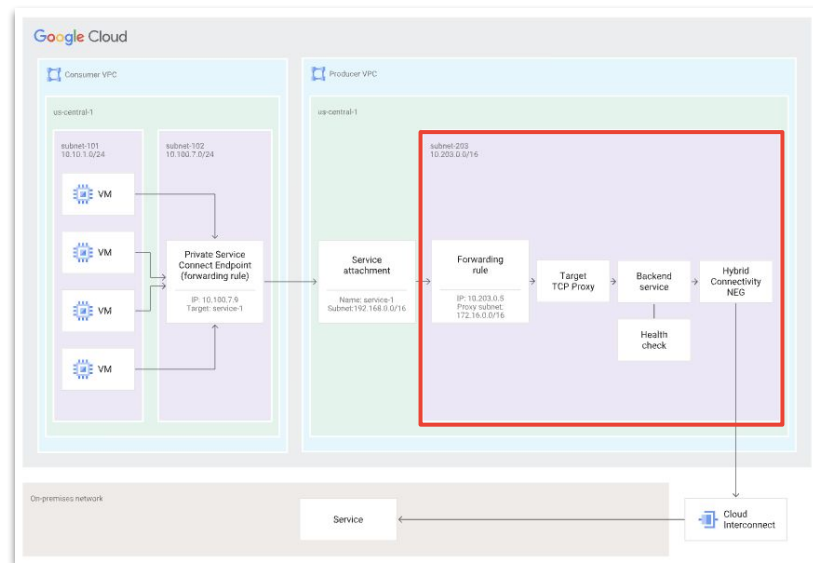
<https://cloud.google.com/vpc/docs/configure-private-service-connect-services-controls>



Private Service Connect の内部リージョン TCP LB サポート

Private Service Connect のサービス アタッチメントターゲットとしてリージョン内部 TCP プロキシロードバランサーの利用が GA

ハイブリッド NEG 経由で VPC 内の VM やオンプレミス環境、もしくは他クラウド上でホストされたサービスに対してアクセス可能



<https://cloud.google.com/load-balancing/docs/tcp/set-up-int-tcp-proxy-hybrid#publish>

ファイアウォール ポリシー

ファイアウォール ポリシー ルールで以下の条件に
基づいたトラフィックの許可 / 拒否が可能

- 位置情報オブジェクト : 特定の国を対象とした通信制御 (ISO 3166-alpha-2 国コードを利用)
- Threat Intelligence: 悪意のある IP アドレスやクラウド サービスで利用される IP アドレスを対象とした通信制御
- アドレス グループ: 指定した IP アドレスグループを対象とした通信制御
- FQDN オブジェクト: FQDN を指定した通信制御

The screenshot shows the '送信元' (Source) configuration section of a Firewall Policy rule. It includes the following fields and options:

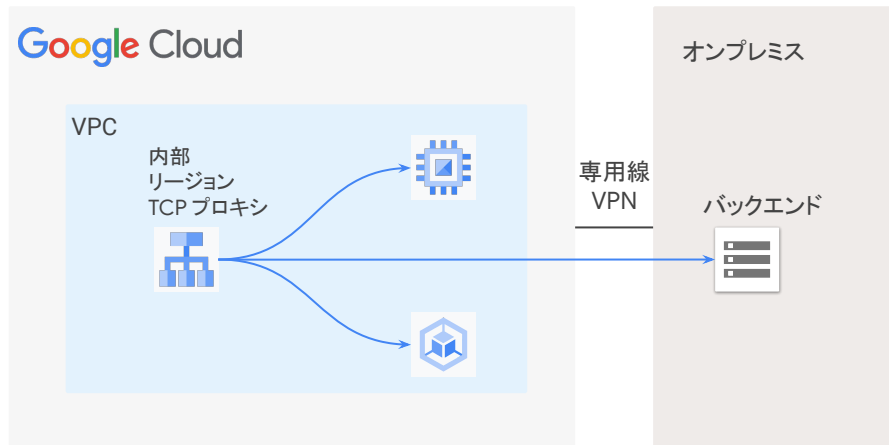
- 送信元** (Source):
 - IP タイプ** (IP Type): A dropdown menu currently set to 'IPv4'.
 - IP 範囲** (IP Range): A text input field containing '0.0.0.0/0' with a clear button (X) and a help icon (?).
 - FQDN**: An empty text input field.
 - 位置情報** (Location): A dropdown menu currently set to '日本 (JP)'.
 - ネットワークの脅威インテリジェンス** (Network Threat Intelligence): A dropdown menu currently set to '既知の悪意のある IP、TOR の exit ノード'.



内部リージョン TCP プロキシ ロードバランサーが GA

Envoy プロキシ ベースのロード バランサーである内部リージョン TCP プロキシ ロード バランサーが GA

同じ VPC ネットワーク内のクライアント、または VPC ネットワークに接続されたクライアントのみから、内部 IP アドレスを利用して Google Cloud、オンプレミス、他クラウド環境でホストされたバックエンド サービスにアクセス可能





SSL ポリシーのサポート

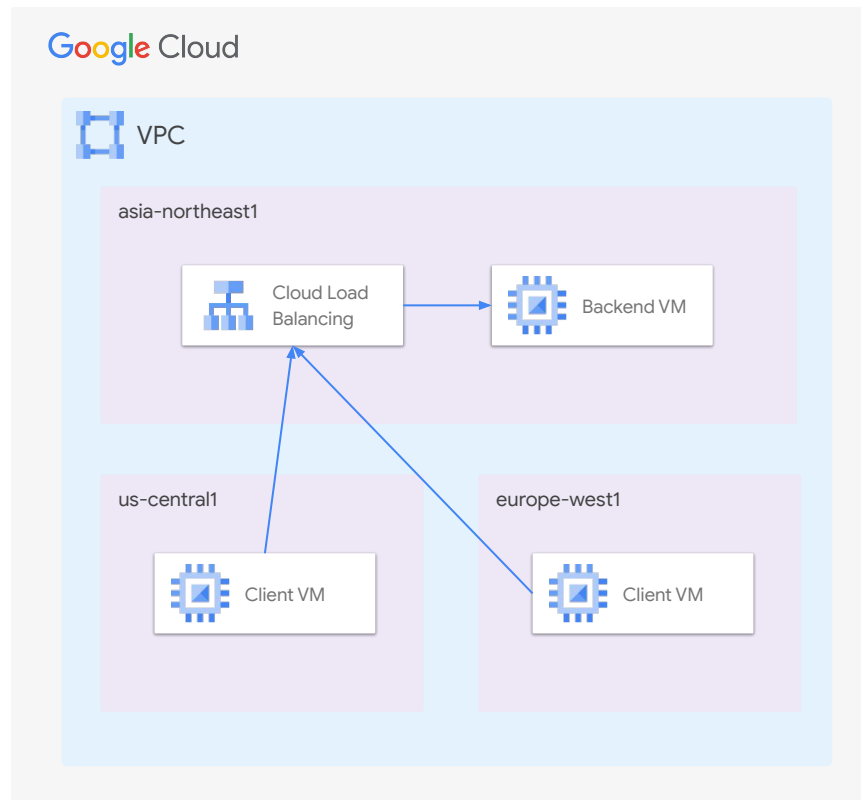
- リージョナル外部 / 内部 HTTP(S) ロードバランサーでリージョナル SSL ポリシーをサポート
- SSL ポリシーを利用することでクライアント - ロードバランサー間で利用される暗号スイートが制御可能
- SSL ポリシーは以下の 4 つのプロファイルが設定可能
 - COMPATIBLE: 最も多くの暗号スイートが含まれ、古いクライアントに適しています
 - MODERN: 幅広い暗号スイートが含まれ、最新のクライアントに適しています
 - RESTRICTED: 最も厳しい設定であり、厳しいコンプライアンス要件を満たします
 - CUSTOM: 利用する暗号スイートを個別に選択できます



内部ロードバランサーへのグローバル アクセス

内部 HTTP(S) ロード バランサー、および内部 TCP プロキシ ロードバランサーに対して別リージョンにデプロイされたクライアントからアクセスが可能

グローバル アクセスの機能は、クライアントがロード バランサーと同じ VPC ネットワーク、もしくはロード バランサーが属する VPC ネットワークと VPC ピアリングで接続された VPC に存在する必要がある

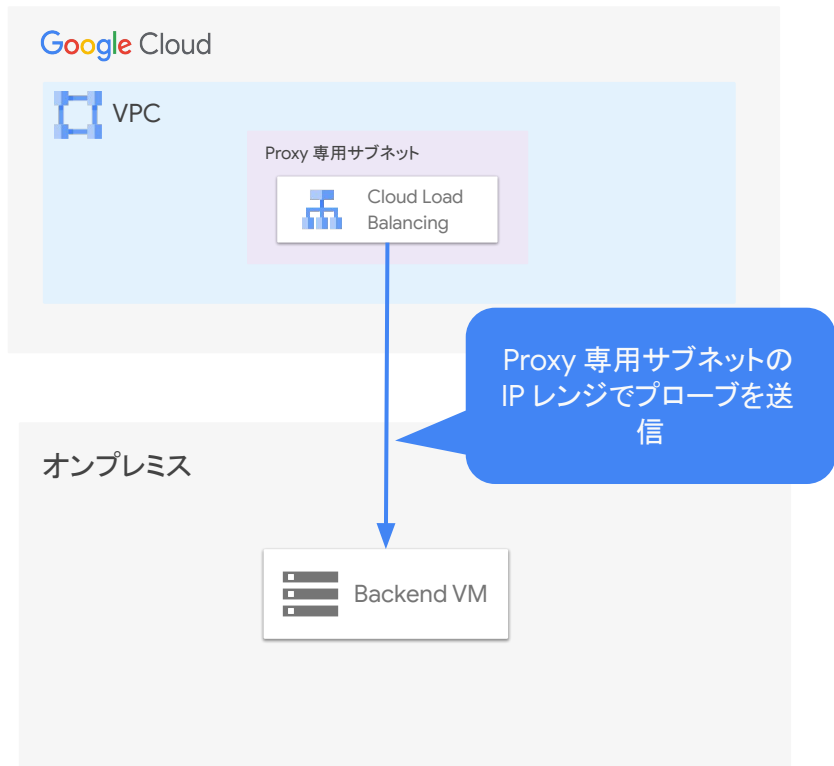




プライベート IP アドレスからのヘルスチェック

ハイブリッド NEG を利用した場合のヘルスチェック プロブを Google のヘルスチェック範囲ではなく独自のプライベート IP アドレスから発信可能 (分散 Envoy ヘルスチェック)

オンプレミス環境で Google のヘルスチェック範囲が許可できない場合などに有効 (許可リストに登録されたプロジェクトのみで利用可能な限定的 GA)





Cloud Armor のレート制限

レート制限を適用するキーとして以下の 3 つのキーが利用可能になりました

- HTTP-PATH: HTTP リクエストの URL パス
- SNI: HTTPS リクエストの TLS セッションでの Server Name Indication
- REGION-CODE: リクエスト送信元の国 / リージョン

元々は以下のキーが利用できていました

- IP: クライアントの送信元 IP アドレス
- HTTP-HEADER: 名前が構成された一意の HTTP ヘッダー
- XFF-IP: クライアントのオリジナルの送信元 IP アドレス (X-Forwarded-For)
- HTTP-COOKIE: 名前が構成された各 HTTP Cookie 値

パフォーマンス ダッシュボード

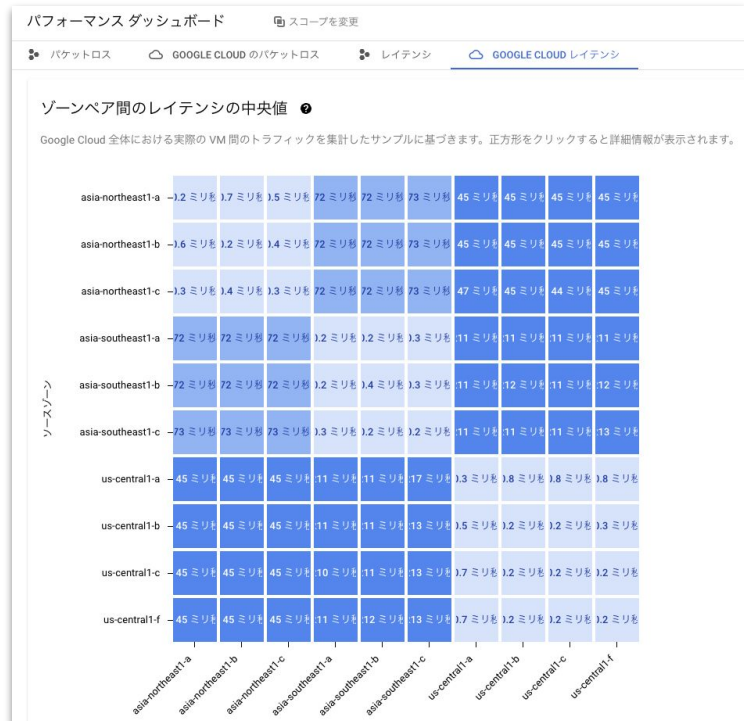
パフォーマンス ダッシュボードによって以下 2 つのシナリオのパフォーマンス情報が可視化可能

- **VM インスタンス間**

Compute Engine インスタンスが存在するゾーン間のパケットロスとレイテンシ

- **Google Cloud とインターネット ロケーション**

Compute Engine インスタンスが存在するリージョンと各地域のインターネット エンドポイントにおけるレイテンシ



03

アプリ開発基盤



Cloud Run(第2世代)の一般提供を開始

システムコールのエミュレーションではなく、Linux との完全な互換性を持ち、以下を実現

- CPU パフォーマンスの高速化
- ネットワーク パフォーマンスの高速化 (特にパケットロスがある場合)
- すべてのシステムコール、名前空間、cgroup のサポートを含む、Linux との完全な互換性
- ネットワークファイル システムのサポート



その他多くの機能追加

- 1つ未満の CPU の使用
- 最大 32 GiB memory, 8 CPU のサポート
- Cloud Run Jobs
- セッションアフィニティ
- [起動プローブ](#)
- Cloud Deploy 対応
- [起動時 CPU ブースト](#)
- [HTTP ライブネスプローブ](#) のサポート

Preview 2022-02-25

GA 2022-04-22

Preview 2022-03-11

GA 2022-06-08

GA 2022-12-09

Preview 2022-09-13

Preview 2022-09-14

GA 2022-12-09

Cloud Run の統合機能

Integrations では、シンプルな Google Cloud コンソール UI と Google Cloud CLI が用意されています。特定のインテグレーションに必要なリソースとサービスを作成して構成できるので、面倒な作業が不要になります。

現在対応しているサービス

- Redis - Google Cloud Memorystore
- Custom Domains - Google Cloud Load Balancing

https://cloud.google.com/run/docs/integrate/redis-memorystore#command-line_1

Add integration

Redis - Google Cloud Memorystore
Fully managed in-memory data store service to build application caches that provides sub-millisecond data access. This integration uses services and manages resources for you. [Learn more](#)

Name *
redis-3u8e
It cannot be changed later.

Capacity (GB)
1
Provision up to 300 GB of space

Resources ⌵

The missing resources will be deployed with this integration.

Resource type	Status	Description
Memorystore Redis Instance	Pending	Fully managed Redis service
Serverless VPC Connector	Pending	Connects serverless environment to a VPC network

Add integration

Custom Domains - Google Cloud Load Balancing
Host your service(s) behind a custom domain with Google Cloud Load Balancers across multiple regions. When this integration is deployed, it will enable services and manage resources for you. [Learn more](#)

Routes

The domain to configure for your Cloud Run service. This must be a domain you can configure DNS for

Domain 1 * Path 1 * Service 1 *

[+ ADD ITEM](#)

Resources ⌵

The missing resources will be deployed with this integration.

Resource type	Status	Description
URL Map	Pending	Routes requests to the correct backend service.
SSL Certificate	Pending	Google-managed SSL certificate
Target HTTP Proxy	Pending	Routes incoming requests to correct URL map
Global Forwarding Rule	Pending	Specifies IP address and frontend configuration
Serverless Network Endpoint Group	Pending	A network endpoint that resolves to a serverless service
Backend Service	Pending	Defines how traffic is distributed



Cloud Functions(第2世代)の一般提供を開始

- 1つのインスタンスで最大 1,000 件の同時リクエストを処理
- 最大 1 時間の実行時間
- 利用可能なインスタンスサイズの増加
(最大 16 GB の RAM と 4 つの vCPU)
- 事前にウォームアップ



Cloud Functions

機能	Cloud Functions (第 1 世代)	Cloud Functions (第 2 世代)
イメージ レジストリ	Container Registry または Artifact Registry	Artifact Registry のみ
リクエストのタイムアウト	最大 9 分	HTTP でトリガーされる関数の場合は 最大 60 分 イベントによってトリガーされる関数の場合は最大 9 分
インスタンスのサイズ	最大 8 GB の RAM (2 vCPU)	最大 16 GiB RAM (4 vCPU) (Preview で 32 GiB RAM)
同時実行	関数インスタンスごとに 1 件の同時リクエスト	関数インスタンスあたり 最大 1,000 件の同時リクエスト
トラフィック分割	非対応	サポート対象
イベントタイプ	7 つのソースからのイベントの直接サポート	Eventarc でサポートされているすべてのイベントタイプをサポート (Cloud Audit Logs を介した 90 以上のイベントソースを含む)
CloudEvents	Ruby、.NET、PHP のランタイムでのみサポート	すべての言語ランタイムのサポート



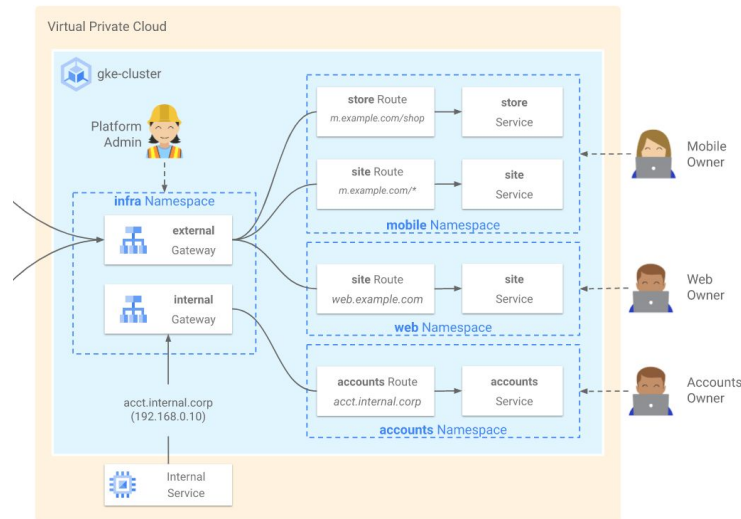
Gateway API

サービスを外部公開する際に用いられる新しい API リソース

Ingress との主な違い:

- Gateway や HTTP Route など複数のリソースから構成
 - クラスタ管理者とアプリケーション開発者で利用リソースを分割
- 複数 Namespace 間で共有可能
- より多くの HTTP(S) ロード バランシング 機能が利用できる
 - Header based routing など
- Internal LB を利用したマルチクラスタ負荷分散サポート

GKE 1.24 以降で**シングルクラスタ用 Gateway API が GA に**
(マルチクラスタ Gateway は Preview)

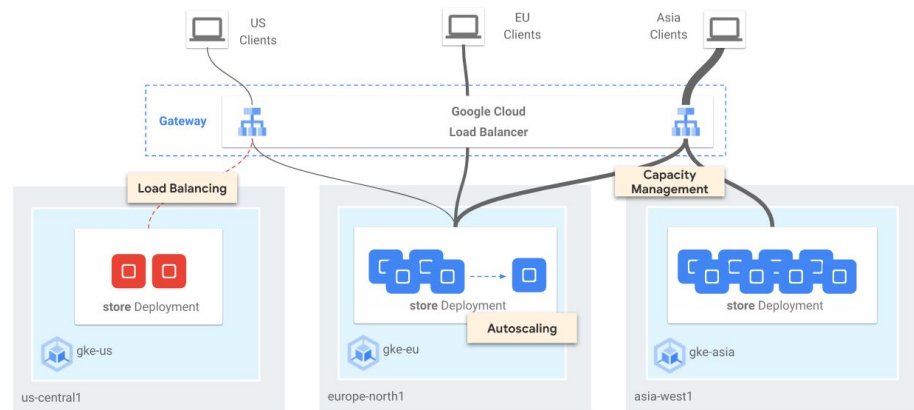




Gateway Traffic Management

Gateway API で Service Capacity の定義や
Capacity Based Routing / Autoscaling が
可能に

定義した Service Capacity を超えた量のトラフィック
がくると別のクラスタにトラフィックを逃すような構
成や RPS ベースのオートスケールを実現





Gateway API の Envoy ベースの外部 HTTP(S) LB サポート

GKE 1.24.2-gke.300 以降で Gateway API で Envoy ベースの

[外部 HTTP\(S\) LB](#) をサポート

ヘッダーベースのルーティングや重み付けトラフィック分割など、[高](#)

[度なトラフィック制御](#)を外部 LB でも利用可能に

Envoy ベースの外部 HTTP(S) LB を利用する場合は以下の

Gateway Class を選択する

- `gke-l7-global-external-managed` (シングルクラスタ)
- `gke-l7-global-external-managed-mc` (マルチクラスタ)



GKE Autopilot の GPU サポート

GKE Autopilot 1.24.2-gke.1800 以降で GPU をサポート
機械学習やレンダリング等のワークロードでも
GKE Autopilot が活用可能に

GPU を使った Pod をデプロイする場合は
コンピューティングクラス同様、`nodeSelector` の条件として
`cloud.google.com/gke-accelerator` ラベルを指定
(利用者側での GPU ドライバのインストール不要)

Autopilot ではタイムシェアリング GPU やマルチインスタ
ンス GPU は GKE Autopilot でサポートされていない

spec:

nodeSelector:

[cloud.google.com/gke-accelerator: nvidia-tesla-a100](https://cloud.google.com/gke-accelerator)



Log Analytics の提供

- Cloud Logging にて BigQuery エンジンを使用し、ログデータにSQL クエリを実行
- 分析のための事前プロビジョニングやETL は不要
- BigQuery 内の他のデータとJOINして分析が可能

The screenshot shows the Cloud Logging Log Analytics interface. At the top, there's a 'Log Analytics' header with a 'PREVIEW' tab and a 'SHARE LINK' button. Below that, there's a 'Query' section with a 'Recent (69)' list and buttons for 'Run in BigQuery', 'Run query', 'Format', 'Clear', and 'SQL reference'. The SQL query is as follows:

```
1 SELECT timestamp, resource.type, severity, json_payload
2 FROM logs_next22_US...AllLogs
3 WHERE timestamp > TIMESTAMP_SUB(CURRENT_TIMESTAMP(), INTERVAL 1 HOUR)
4 AND json_payload IS NOT NULL
5 AND JSON_VALUE(json_payload.message) = "request complete"
6 AND JSON_VALUE(resource.labels.pod_name) LIKE "frontends"
7 LIMIT 50
```

Below the query, there's a 'Log views' section with a 'Results (50)' list and a 'Download' button. The results table shows 12 rows of log entries. Each row contains a timestamp, resource type, severity, and a message. The messages are HTTP requests, with the 6th row expanded to show the full JSON payload.

Log ID	Timestamp	Resource Type	Severity	Message
4	2022-09-22 03:42:48.643 UTC	k8s_container	DEBUG	{http.req.id: "9828d176-7679-4f91-8f8f-e435a38cd4e2", http.req.method: "GET"
5	2022-09-22 03:42:57.183 UTC	k8s_container	DEBUG	{http.req.id: "d66fb2d4-17e9-428a-9add-719d53647162", http.req.method: "GET"
6	2022-09-22 03:42:54.704 UTC	k8s_container	DEBUG	{ http.req.id: "215788e4-aa8-411a-b48c-91773a79ac28" http.req.method: "GET" http.req.path: "/product/OLJCESPC72" http.resp.bytes: 8893 http.resp.status: 200 http.resp.took_ms: 13 message: "request complete" session: "3902dac7-9088-4281-9c98-778d7628d85e" timestamp: "2022-09-22T03:42:54.704588483Z"
7	2022-09-22 03:42:56.954 UTC	k8s_container	DEBUG	{http.req.id: "1e879944-3ec3-48a8-a6b1-fb48ef92aa49", http.req.method: "GET"
8	2022-09-22 03:42:54.176 UTC	k8s_container	DEBUG	{http.req.id: "01b947aa-9969-4cb6-911f-9ac298d8a695", http.req.method: "POST"
9	2022-09-22 03:42:56.994 UTC	k8s_container	DEBUG	{http.req.id: "79c251d1-bb3a-450c-85a3-74847106aded", http.req.method: "GET"
10	2022-09-22 03:42:56.968 UTC	k8s_container	DEBUG	{http.req.id: "260c9886-4fbc-4e76-9f74-30d65279eb8b", http.req.method: "GET"
11	2022-09-22 03:42:56.211 UTC	k8s_container	DEBUG	{http.req.id: "42a43aec-094a-4b16-8d2d-809c939281cd", http.req.method: "GET"
12	2022-09-22 03:42:53.007 UTC	k8s_container	DEBUG	{http.req.id: "d559781c-2f4b-4f37-8b8c-154c8ded254a", http.req.method: "GET"



Cloud Build - ステップ Fail 時の継続実行

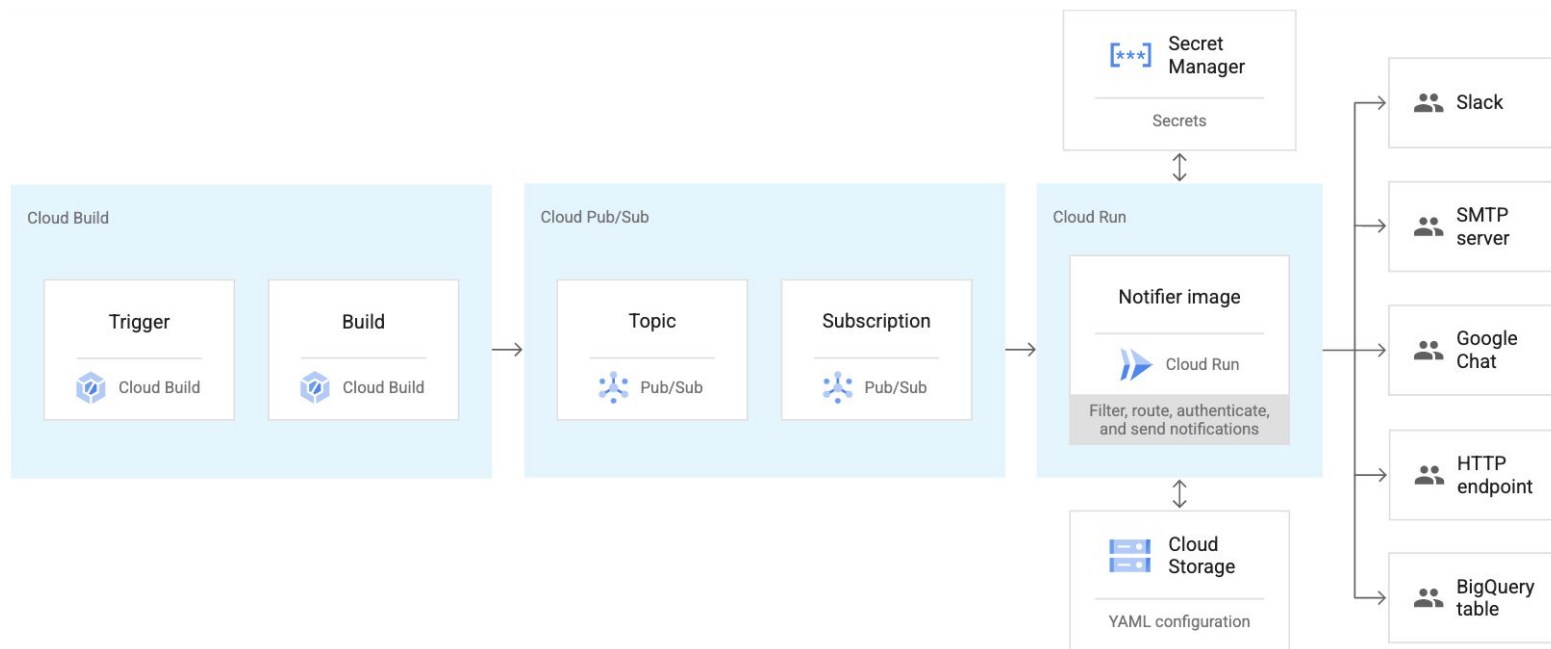
特定のステップが Fail した場合もビルドを継続実行することが可能に

- allowFailure
 - true に設定すると、そのステップが失敗しても後続のステップが成功していればビルド全体が成功となる
- allowExitCodes
 - そのステップで許容可能な Exit code を指定する

```
steps:  
- name: 'ubuntu'  
  args: ['-c', 'exit 1']  
  allowFailure: true  
steps:  
- name: 'ubuntu'  
  args: ['-c', 'exit 1']  
  allowExitCodes: [1]  
steps:  
- name: 'ubuntu'  
  args: ['echo', 'Hello World']
```



Cloud Build - Notifier テンプレートを使用して、メール、BigQuery、Webhook ベースの通知をカスタマイズ可能に





マネージドでセキュアな開発環境

- 事前定義された開発環境により、開発者のオンボーディングを高速化
- VPC 内で実行可能
- 複数の IDE (JetBrains IDEs, VSCode) のサポート
- 再現可能なコンテナベースの環境設定
- ステージング環境で直接コードをテスト / 実行

```
EXPLORER
├── .idea
├── .readmes
├── .vscode
├── img
├── src
├── backend
├── frontend
├── kubernetes-manifests
├── public
├── utils
├── views
├── JS app.js
└── Dockerfile

src > frontend > JS app.js > ...
22 // Application will fail if environment variables are not set
23 if(!process.env.PORT) {
24   const errMsg = "PORT environment variable is not defined"
25   console.error(errMsg)
26   throw new Error(errMsg)
27 }
28
29 if(!process.env.GUESTBOOK_API_ADDR) {
30   const errMsg = "GUESTBOOK_API_ADDR environment variable is not
31   console.error(errMsg)
32   throw new Error(errMsg)
33 }
34
35 // Starts an http server on the $PORT environment variable
36 const PORT = process.env.PORT;
```

04

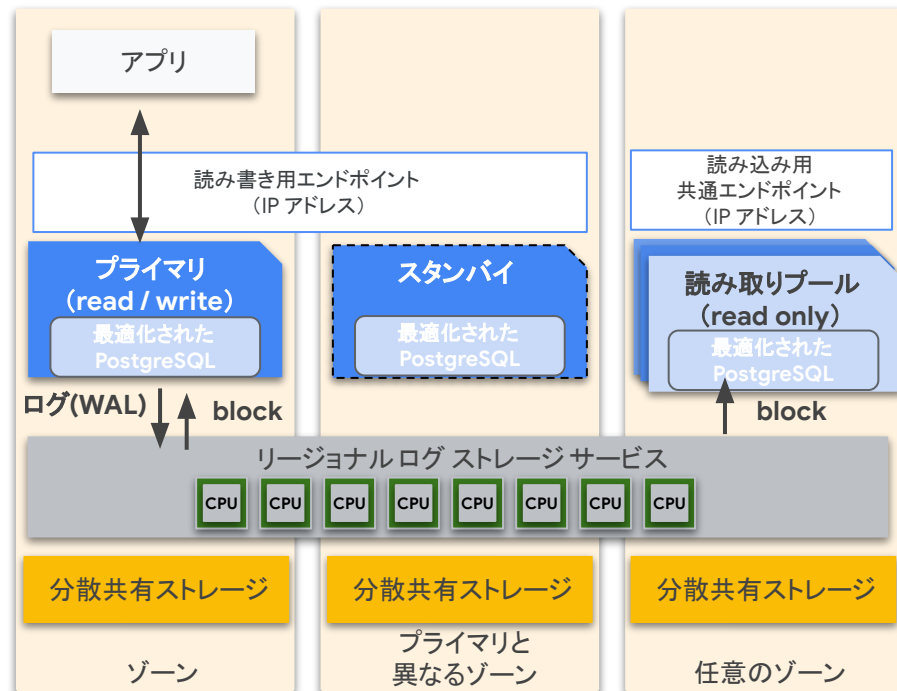
データベース



PostgreSQL に独自のログストレージサービスを
導入した新しい DB サービス

- PostgreSQL との完全な互換性
- OLTP で標準の PostgreSQL の最大 4 倍高速*
- OLAP で標準の PostgreSQL の最大 100 倍高速*
- SLA としてメンテナンス含め稼働率 99.99%
- リージョン間レプリケーションによる
DR 構成にも対応

* Google Cloud の 2022 年 3 月のパフォーマンス テストに基づく。





Cloud SQL for MySQL

GA

Cloud SQL for MySQL が、IAM データベース認証の自動認証に対応しました。

- データベースの組み込み認証: ユーザー名とパスワードを使用してデータベース内の認証
- IAM データベース認証: Google Cloud の IAM に基づいてデータベース内の認証
 - **自動認証: Cloud SQL Auth Proxy などにアクセストークンのリクエストと管理を任せる**
 - 手動認証: 自分でアクセストークンを取得して、接続時に渡す

※ Cloud SQL for PostgreSQL は以前から IAM の自動認証に対応済み

Cloud IAM

既存の IAM プリンシパルをこのユーザー アカウントに関連付けます。接続するには、インスタンス レベルのアクセスを提供するロールが割り当てられている必要があります。

プリンシパル*
wataruinoue@example.com

Cloud IAM 認証を使用してユーザー アカウントを作成した後、データベース権限は付与されないため、必要に応じて権限を付与してください。 [詳細](#)

追加

キャンセル



Cloud SQL for SQL Server

GA

Cloud SQL for SQL Server で以下の機能が一般提供になりました。

- ポイントインタイム リカバリ
- カスタムタイムゾーンの指定

ポイントインタイム リカバリを有効にする

ログ先行書き込みのアーカイブを使用して、特定の時点をもとにデータを復元できます。

現時点では、クロスリージョンレプリケーションとポイントインタイムリカバリを同じインスタンスで設定することはできません。 [詳細](#)

ログを保持する日数を選択する

同時に保存するトランザクションログの日数を決定する保持ポリシーを設定できます。 [詳細](#)

ログの日数 *

7

注目のフラグの設定

フラグによって有効になっている、またはパラメータによって定義されている、特定の機能の設定をカスタマイズします

デフォルトの照合

Japanese_CI_AS

設定は後から変更できません。このインスタンスのデータベースにデフォルトの照合を指定します。 [詳細](#)

タイムゾーン

Tokyo Standard Time

設定は後から変更できません。SQL Server インスタンスのローカルタイムゾーンを指定します - 基盤となる VM は UTC で実行されます。 [詳細](#)



Cloud SQL Auth プロキシの v2 がリリースされました。v2 では以下の新機能がサポートされます。

- Cloud Monitoring と Cloud Trace による指標とトレース
- Prometheus のサポート
- 指定したサービス アカウントへの Impersonation (権限借用)
- Cloud SQL Go Connector としてリリースされた別の Dialer 機能
- 環境変数による設定
- 完全に POSIX 準拠したフラグ

v1 から v2 の移行については、[移行ガイド](#)を参照



- [有効期限\(TTL\)機能](#) で PostgreSQL 言語データベースがサポートされました
 - 指定した Timestamp 型のフィールドに基づいて、Cloud Spanner 側でデータを自動削除
 - 指定したタイムスタンプから何日後に削除するかを指定
 - ユースケース
 - 過去の不要なデータの削除でコスト削減
 - コンプライアンス要件への対応
- PostgreSQL 言語データベースで [JSONB データ型](#) がサポートされました
 - JSONB データ型を使うと、Spanner の PostgreSQL 言語で JSON 型の半構造化データを保持できます
 - [一部の PostgreSQL JSONB 機能](#) は Spanner JSONB ではサポートされていません

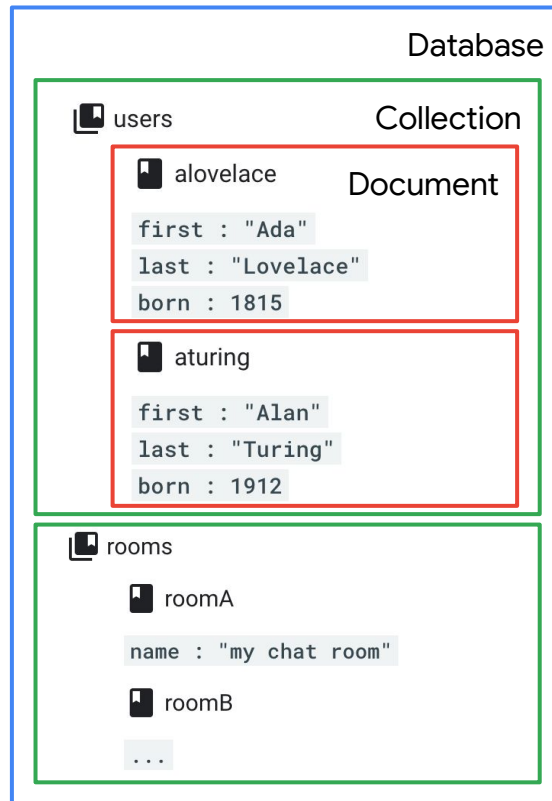


Firestore(ネイティブモード)で以下の制限が撤廃されました。

- 10,000 書込み/秒/Database
- 100 万同時接続/Database

注意事項

- 引き続き、約 1 書込み/秒/Document の制限は存在します
- 引き続き、大規模なトラフィックが発生する場合は [「500/50/5」ルール](#) に従ってトラフィックを徐々に増やす必要があります
 - 「500/50/5」ルール: 毎秒 500 回から、5 分ごとにトラフィックを 50% ずつ増やしていく





Time-to-Live (TTL) ポリシーが一般提供になりました。

- 指定した Timestamp 型のフィールドに基づいて、Firestore 側でドキュメントを自動削除
- 期限切れ後 72 時間以内のどこかで削除
- ユースケース
 - 過去の不要なデータの削除でコスト削減
 - コンプライアンス要件への対応

TTL ポリシーの作成

コレクショングループ 1 つに対し、TTL ポリシーを 1 つ設定できます。 [詳細](#)

タイムスタンプフィールドを入力します

指定したコレクショングループのタイムスタンプフィールドに基づく削除のスケジュールがドキュメントに設定されています。データは通常、有効期限が切れた後 72 時間以内に削除されず。

i 単一フィールドのインデックスは、TTL を受け取るフィールドでは問題の原因になる場合があります。 [詳細](#)

コレクショングループ*

users

タイムスタンプフィールド*

expiredAt



Firestore で `count()` クエリがプレビューで利用できるようになりました。

- ドキュメントを取得してローカルでカウントするのではなく、Firestore 側でクエリの結果件数を集計できるようになりました
- SDK では `getCountFromServer()` 関数で利用
- `count()` の条件に一致した件数に応じて料金が発生し、1,000 件につき 1 回のドキュメント読み取り相当の料金が課金されます
(東京の場合: \$0.038/10 万読み取り)

```
const collection_ = collection(firestore, 'employees');

(async () => {
  // employees コレクションの全件カウント
  const totalCount =
    await getCountFromServer(query(collection_));

  // employees の中で、role = dev の件数をカウント
  const devCount =
    await getCountFromServer(query(collection_,
    where('role', '==', 'dev')));

  console.log('totalCount', totalCount.data());
  // -> {count: "100"}
  console.log('devCount', devCount.data());
  // -> {count: "10"}
})();
```

05

セキュリティ



Policy Intelligence

[11/18] Policy Analyzer は、[組織のポリシー分析](#)を提供するようになりました。Policy Analyzer は、組織のポリシーの制約によって影響を受けるリソースに関する詳細情報を取得するのに役立ちます。この機能はプレビューで利用できます。

Google Cloud Retail-Search-Demo

スラッシュ (/) を使用してリソース、ドキュメント、プロダクトなどを検索

ポリシー アナライザ

Policy Analyzer は、「誰がこのリソースにアクセスでき、何ができるのか」を調べるうえで役立ちます。IAM ポリシーの設定を調べ、各プリンシパルとリソースにどのように影響するかを示します。 [詳細](#)

ポリシーの分析

いずれかのテンプレートからクエリを作成し、どのユーザーがどのリソースにアクセスできる状態を確認します。

カスタムクエリ
プリンシパルが Google Cloud リソースに対してどのようなアクセス権を持っているかを確認するには、カスタムクエリを作成します。

カスタムクエリを作成

サービスアカウントの権限を借用できるユーザー
サービスアカウントの権限借用を使用すると、プリンシパルは、特定のサービスアカウントがアクセスできるリソースに間接的にアクセスできます。 [Learn more](#)

クエリを作成

プロジェクトのファイアウォール ルールを変更できるユーザー
ファイアウォール ルールは、さまざまなネットワークからリソースにどのユーザーがアクセスできるかを制御します。 [Learn more](#)

クエリを作成

従業員（または退職した従業員）がどのようなアクセス権を持っているか
対象の従業員が現在アクセスできるリソースを確認できます。

クエリを作成

▼ その他のテンプレート

最近のアクティビティの分析 [プレビュー](#)

いずれかのテンプレートからクエリを作成し、どのユーザーがどのリソースにアクセスしたかを確認できるようにします。

このサービス アカウントが最後に使用された日時
プロジェクトでサービス アカウントが最後に使用された日時を確認できます。

クエリを作成

このサービス アカウント キーが最後に使用された日時
プロジェクトでサービス アカウント キーが最後に認証に使用された日時を確認できます。

クエリを作成

06

オペレーション



Cloud Billing

[12/19] Billing の BigQuery へのエクスポートで Cloud SQL の費用を表示する **Cloud Billing** の詳細な費用エクスポートで、詳細な Cloud SQL インスタンスの費用データを表示できるようになりました。エクスポートの **resource.global_name** フィールドを使用して、Cloud SQL インスタンス データを表示およびフィルタリングします。

[public] Google Cloud Billboard v2.29 - bit.ly/billboard-template

リセット 共有

- Billing overview
- Compute Engine
- Compute Engine - resource ...
- GKE
- Cloud SQL**
- Cloud Storage
- Analytics
- Serverless
- Data Transfer
- other services / adhoc analy...
- Version history

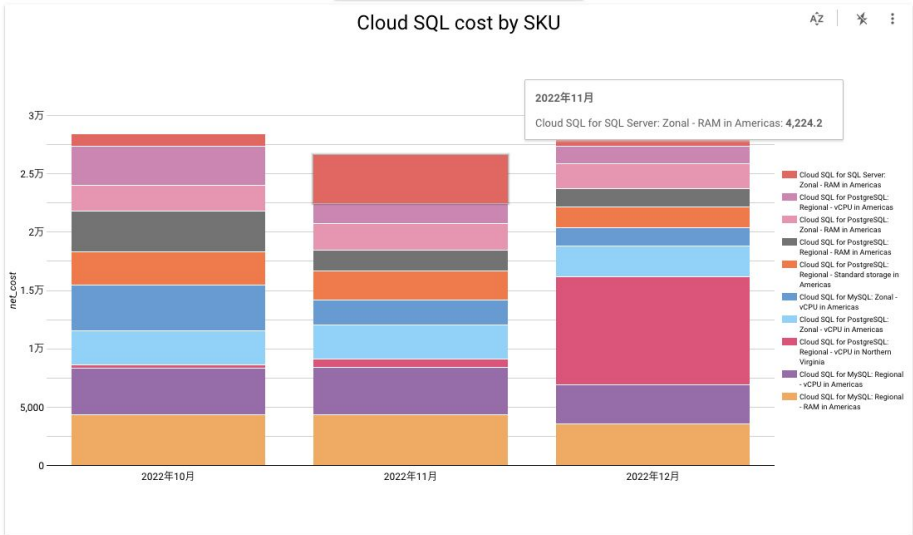


GCP BILLboard

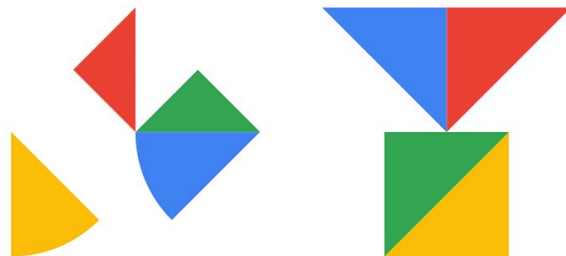
Billing and usage insights

Select Project

Cloud SQL cost by SKU



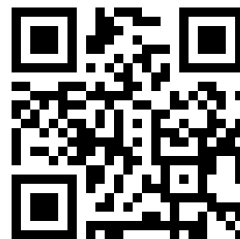
Google Cloud Day '23 Tour



- 📍 東京 (オンライン) 5月23日(火)～25日(木)
- 📍 大阪 (ハイブリッド) 6月2日(金)
- 📍 名古屋 (ハイブリッド) 6月22日(木)
- 📍 福岡 (ハイブリッド) 6月30日(金)

企業のDXを加速する、そのヒントを4都市からお届けします。

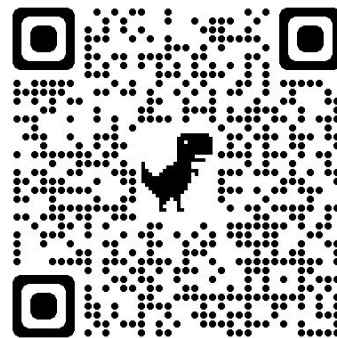
今すぐ登録 goo.gle/gcd23_1p



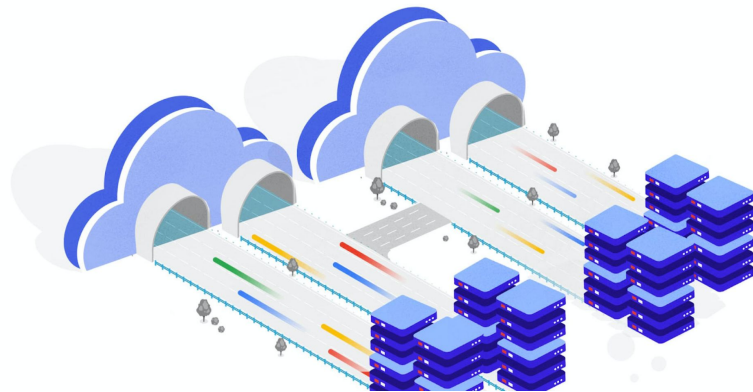
3月7日、8日開催

Infra Onair

- Google Cloud への移行に役立つ Migrate to VM のご紹介
- 実例解説！ Google Cloud マイグレーションにおけるポイント
- SAP を Google Cloud へマイグレーションして お手軽に SAP データを可視化
- メディアサービス向け CDN「Media CDN」概要とオブザーバビリティの実践
- Google Cloud の最新のネットワーク セキュリティ機能のご紹介



<https://goo.gle/InfraOnAir>

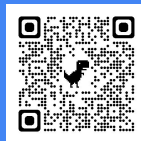
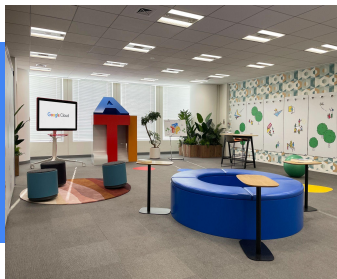
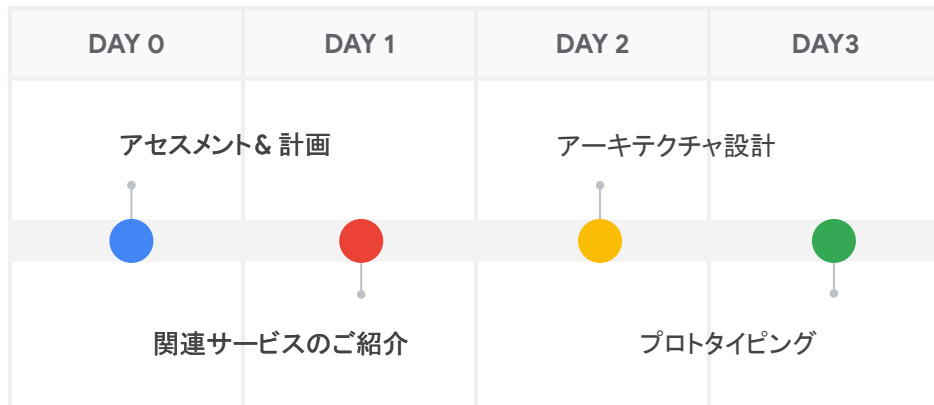


Google Cloud エンジニアによる、内製化支援プログラム

Tech Acceleration Program (TAP)

3 日間のワークショップ形式で学ぶ！
モダン アプリケーションの開発

- アセスメント、プランニング
- 関連サービスの紹介
- アーキテクチャ設計
- プロトタイピング
- CI / CD パイプライン構築



<https://goo.gle/tap>