



Google Kubernetes Engine 最新情報

篠原 一徳

Google Cloud、カスタマー エンジニア

スピーカー自己紹介



篠原 一徳

Google Cloud
カスタマーエンジニア

主にゲーム業界のお客様向けに、
コンテナ関連サービスの提案、技術サポートを行っています。

趣味は子育て、Jリーグ観戦です。

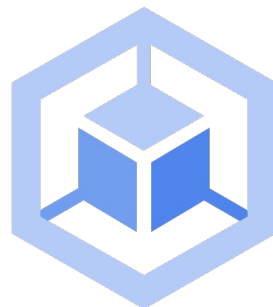


Google Kubernetes Engine Quick Recap

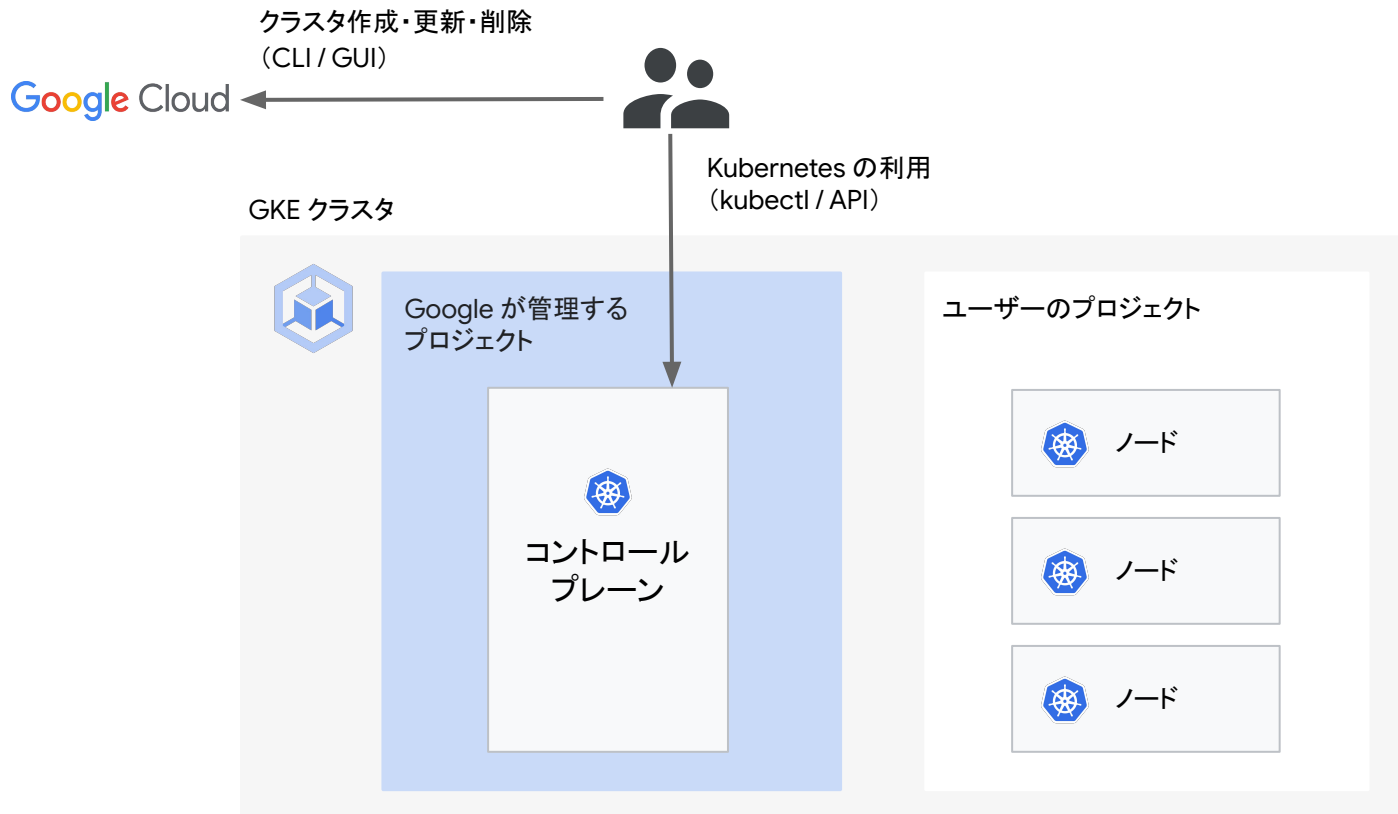
Google Kubernetes Engine (GKE) とは？

2015 年にリリースされた Kubernetes のマネージド サービス

- Kubernetes のコントロール プレーン は Google が管理
- 2 つのモード
 - **GKE Standard** : ノード はユーザー管理
 - **GKE Autopilot** : ノード も Google 管理
- Google Cloud の各種サービスとネイティブに連携



GKE の基本的なアーキテクチャ





Gateway 関連アップデート

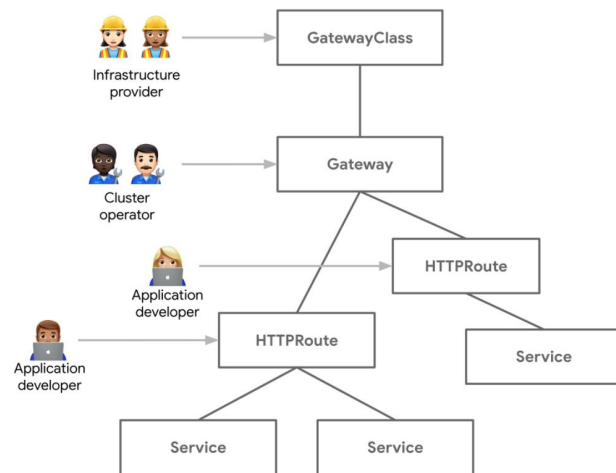
Recap: Gateway

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。

Kubernetes の SIG-Network community で開発が進められている。

GKE では 2021 年 5 月より以下の GatewayClass が利用可能。

- External Gateway
- Internal Gateway
- External multi-cluster Gateway
- Internal multi-cluster Gateway



Gateway - Service capacity

Pod 単位の RPS を Service リソースに設定することで、

後述の

- Traffic-based load balancing
- Traffic-based autoscaling

を実現する。

デフォルト値は 100,000,000 RPS

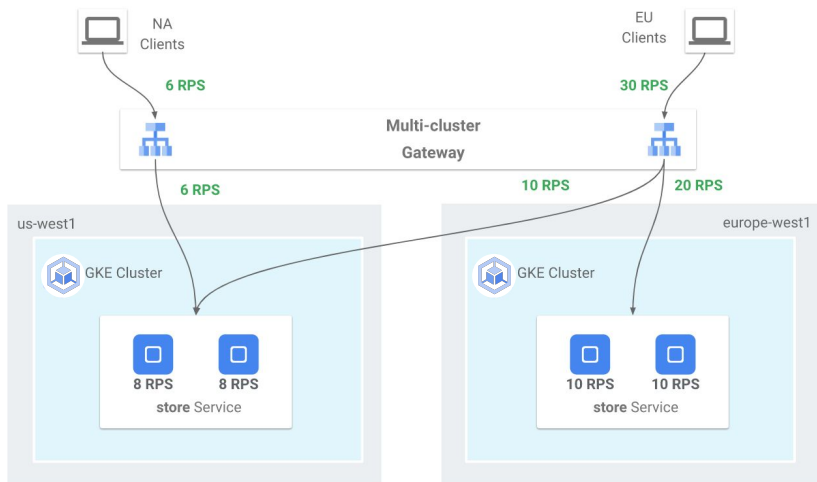
(全 GatewayClass 共通)

```
apiVersion: v1
kind: Service
metadata:
  name: store
  annotations:
    networking.gke.io/max-rate-per-endpoint: 5
spec:
  ports:
    - port: 8080
      targetPort: 8080
      name: http
  selector:
    app: store
  type: ClusterIP
```

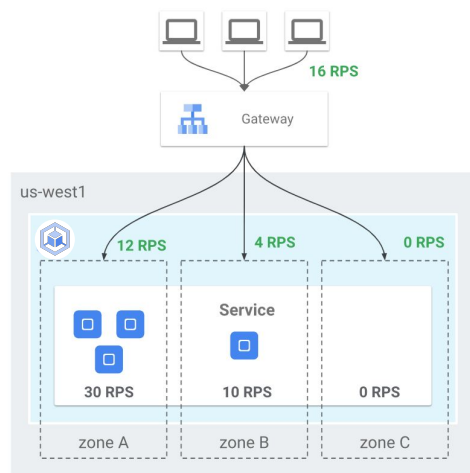

Gateway - Traffic-based load balancing

Service capacity で指定した Pod 単位の RPS を元に、ロード バランシングを行う。

Multi region の例



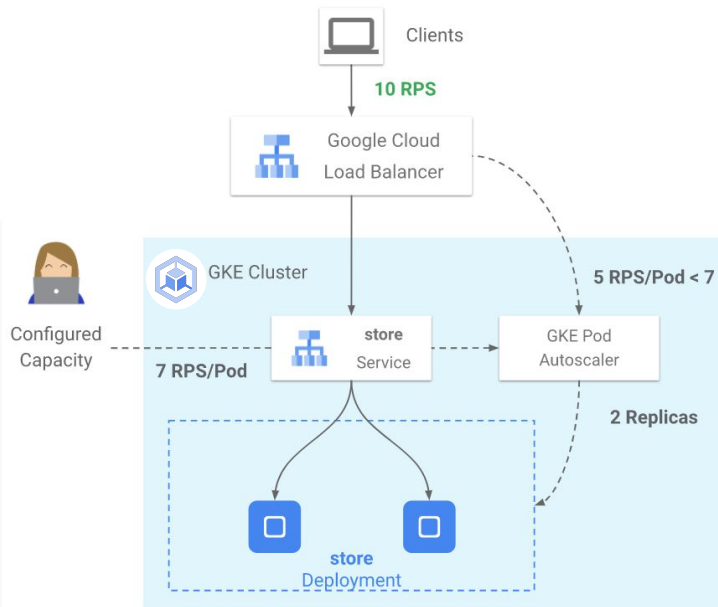
Single region の例



Gateway - Traffic-based autoscaling

実際のトラフィック流量と、
サービスのキャパシティをベースに Pod の
水平スケーリングを行う。

```
kind: HorizontalPodAutoscaler
spec:
  metrics:
    - type: Object
      object:
        metric:
          name: "autoscaling.googleapis.com|gclb-capacity-utilization"
        target:
          averageValue: 70
          type: AverageValue
```





運用系機能 アップデート

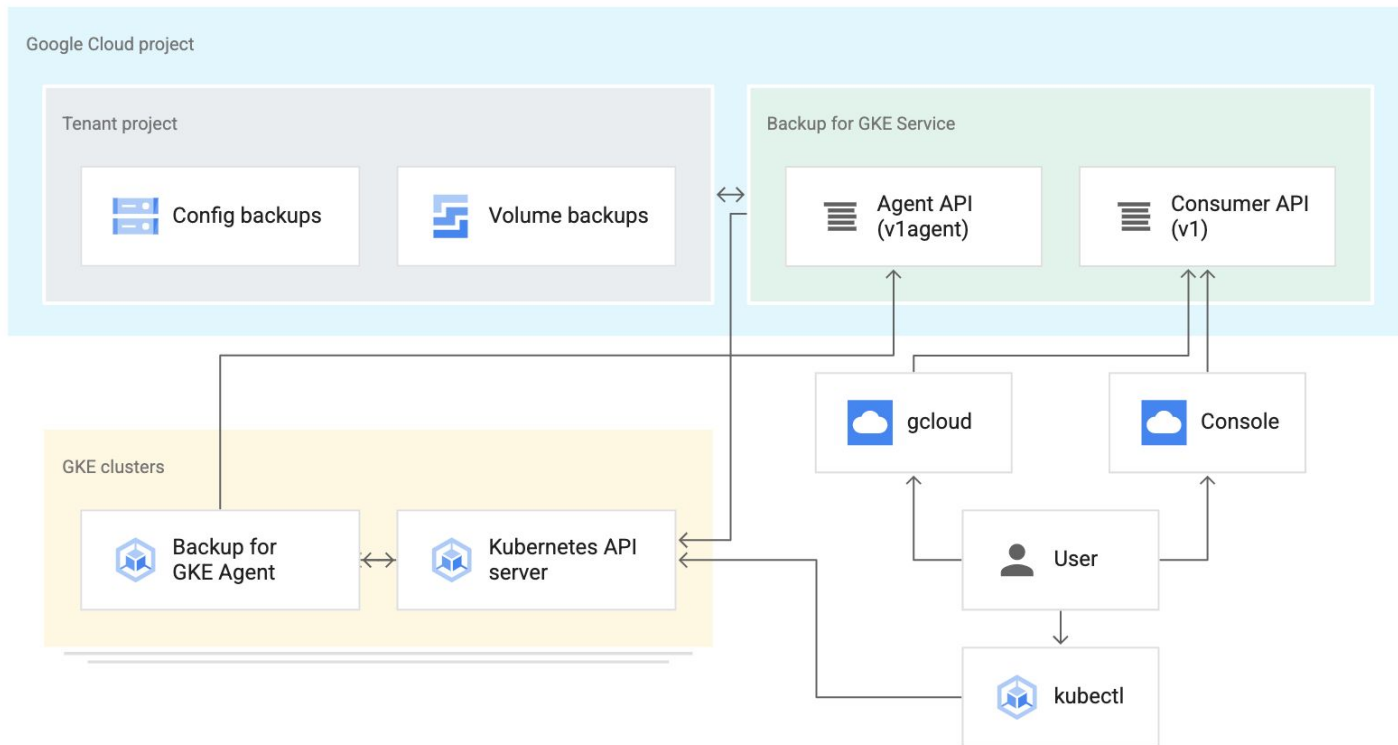
Backup for GKE

GKE 上のワークロードのバックアップ及びリストアを行う機能。

主に Stateful なワークロードを対象とし、以下のデータを扱う。

- Kubernetes リソース
 - kube-apiserver から取得出来る情報、
Backup for GKE では **Config backup** と呼ぶ。
- GCE Persistent Disk (PD) の Volume snapshot
 - PVC (Config backup される) に対応した PD の snapshot
Backup for GKE では **Volume backup** と呼ぶ。

Backup for GKE のアーキテクチャ



Backup for GKE 利用シナリオ

Disaster Recovery 対応や CI/CD のパイプライン、ワークロードをクローンしたり、クラスタ アップグレード時のバックアップ用途に。

バックアップ、リストアの範囲は以下の通り選択可能。

- クラスタ全体のバックアップ / クラスタ全体のリストア
- クラスタ全体のバックアップ / 部分的なリストア
- Namespace 単位のバックアップ / リストア
- アプリケーション 単位のバックアップ / リストア

Recap: GKE のバージョン

- GKE ではコントロール プレーン 及び ノード のバージョンをそれぞれ分けて管理
- コントロール プレーンは自動でアップグレードされる(無効化不可)
- 各バージョンは以下の通り、3 つのコンポーネントで構成されている



Recap: リリースチャンネルとは？

バージョニングとアップグレードを行う際のベスト プラクティスを提供する仕組み。

リリース チャンネルにクラスタを登録すると、

コントロールプレーン及びノードのアップグレードが自動的に行われる。

利用できる機能と更新頻度の異なる、以下 3 つのチャンネルがある。



Maintenance Exclusion の強化

新たに **Scope** という設定を導入。

指定する Scope 次第では、**アップグレードを半年に一度にするなど、より柔軟な運用が可能に。**

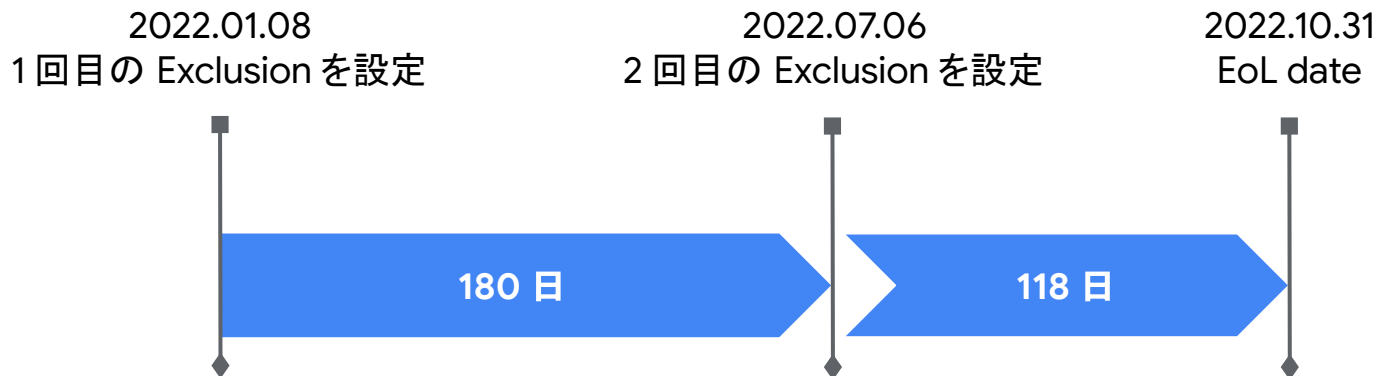
Scope	Control plane			Node pools			
	Minor upgrade	Patch upgrade	VM disruption due to GKE maintenance	Minor upgrade	Patch upgrade	VM disruption due to GKE maintenance	
No upgrades (default)	No	No	No	No	No	No	最大 30 日
No minor upgrades	No	Yes	Yes	No	Yes	Yes	最大 180 日
No minor or node upgrades	No	Yes	Yes	No	No	No	最大 180 日

No minor upgrades と No minor or node upgrades scope は

リリースチャネル登録済みクラスタのみ利用可能。

Maintenance Exclusion を 180 日以上設定する例

- Version は 1.21.5-gke.1302 を利用 (1.21 の EoL は 2022 年 10 月 予定)
- Scope には No minor or node upgrades scope を選択
- Maintenance Exclusion 設定時点から 180 日を超えて設定出来ないが、
EoL の範囲内であれば再度 180 日後に Maintenance Exclusion を設定可能



Cluster notification

任意の Pub/Sub トピックに対し、
アップグレード及びセキュリティに関する
通知メッセージを送信。

- **UpgradeAvailableEvent**

- 新バージョンが利用可能になった時に通知
 - マイナー バージョン: 2 - 4 週間前
 - パッチ バージョン: 1 週間前

- **UpgradeEvent**

- アップグレードが開始されると通知

- **SecurityBulletinEvent**

- セキュリティ脆弱性に関する情報の通知

New master version "1.19.9-gke.1400" is available for upgrade in the RAPID channel.

```
cluster_location: asia-northeast2
cluster_name: rapid-autopilot-an2
payload: {"version": "1.19.9-gke.1400", "resourceType": "MASTER", "releaseChannel": {"channel": "RAPID"}}
project_id: 605899591260
type_url: type.googleapis.com/google.container.v1beta1.UpgradeAvailableEvent
```

New master version "1.18.16-gke.2100" is available for upgrade in the REGULAR channel.

```
cluster_location: asia-northeast2
cluster_name: regular-autopilot-an2
payload: {"version": "1.18.16-gke.2100", "resourceType": "MASTER", "releaseChannel": {"channel": "REGULAR"}}
project_id: 605899591260
type_url: type.googleapis.com/google.container.v1beta1.UpgradeAvailableEvent
```

New node version "1.18.17-gke.1200" is available for upgrade.

```
cluster_location: asia-northeast2
cluster_name: static-standard-an2
payload: {"version": "1.18.17-gke.1200", "resourceType": "NODE_POOL", "resource": "projects/kzs-sandbox/locati"}
project_id: 605899591260
type_url: type.googleapis.com/google.container.v1beta1.UpgradeAvailableEvent
```

New master version "1.18.16-gke.2100" is available for upgrade in the STABLE channel.

```
cluster_location: asia-northeast2
cluster_name: stable-standard-an2
payload: {"version": "1.18.16-gke.2100", "resourceType": "MASTER", "releaseChannel": {"channel": "STABLE"}}
project_id: 605899591260
type_url: type.googleapis.com/google.container.v1beta1.UpgradeAvailableEvent
```



その他 アップデート

Image Streaming

コンテナイメージを Pull する際に、イメージのデータをストリーミングすることで、

- 自動スケーリングの高速化
- イメージを pull する際のレイテンシの短縮
- Pod の起動の高速化

を実現する。

コンテナのイメージサイズが大きく、起動時間に時間が掛かっている場合に有効。
(e.g. 機械学習の学習済みモデルを含むコンテナイメージ)

アプリケーションの起動時間が 3 倍改善された事例も。

Image Streaming の仕組み

リモートファイルシステムをネットワークマウントし、起動するコンテナのルートファイルシステムとして利用しコンテナを起動する。

並行して全体のコンテナイメージを Node にダウンロードし、ダウンロードが完了したらローカルディスクの cache 利用に変更する。

Image streaming がない場合

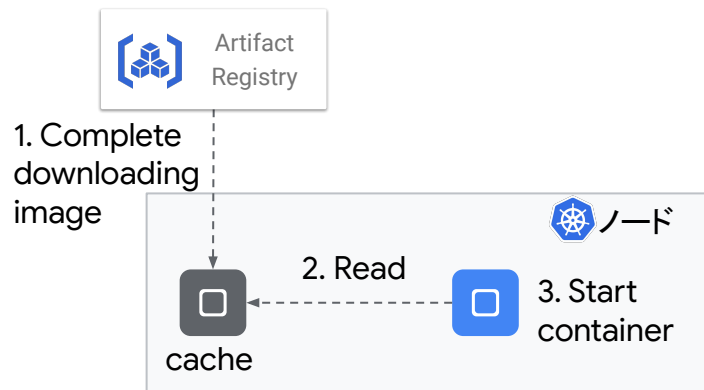
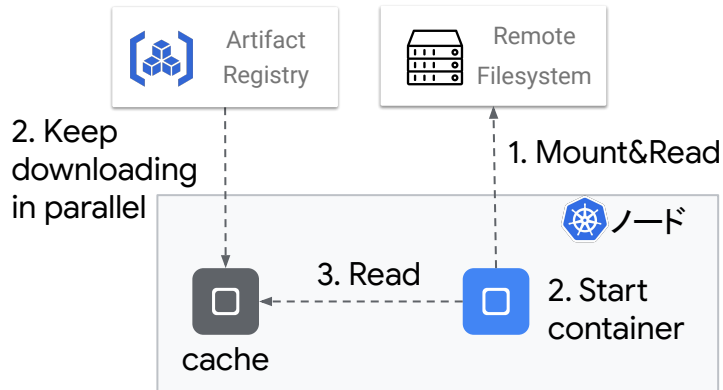


Image streaming がある場合



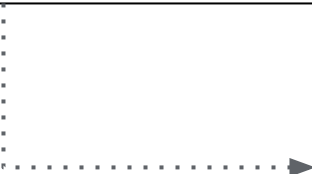
Spot VM / Pod

プリエンティブル VM と違い **24 時間の稼働時間制限がない**

Spot VM を GKE Standard の Node として利用可能。

Autopilot mode の場合は、Spot Pod として割安な価格で利用可能。

```
gcloud container node-pools create test-pool \  
  --cluster=test-cluster \  
  --spot
```

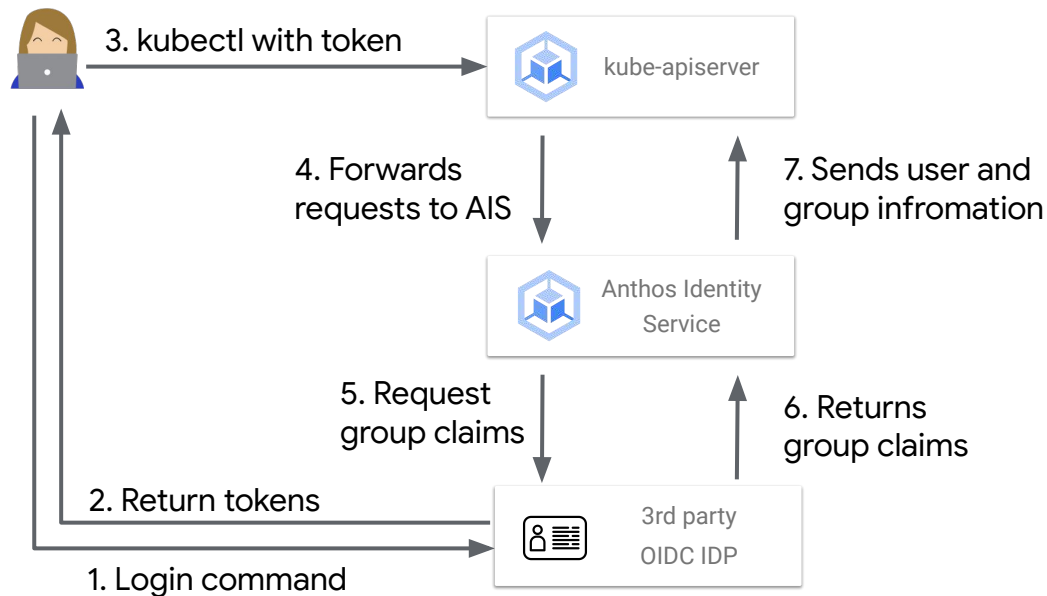


```
apiVersion: v1  
kind: Pod  
spec:  
  nodeSelector:  
    cloud.google.com/gke-spot: "true"
```

Identity Service for GKE

GKE クラスターのユーザー認証に、
**3rd Party の IDP を利用した
OIDC 認証**が可能に。

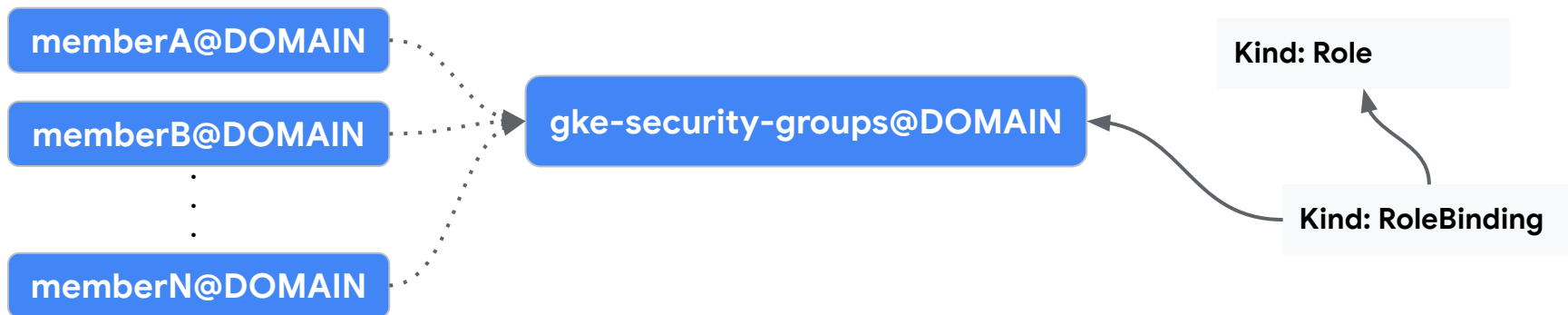
Anthos Identity Service と
同じ仕組みを利用。



Google Groups for RBAC

Google グループのメンバーに対して、一括で GKE の RBAC 権限を割り当てることが可能に。

ユーザーアカウント個別の権限管理をせずに、プロジェクトメンバーの入退職時の処理などをシンプルに運用できる。



まとめ

- **Ingress 以上に機能が充実していきそうな Gateway**
 - GA はもう少しお待ち下さい mm
- 運用系の機能が拡充
 - **Backup for GKE** : Stateful なワークロードがより運用しやすく
 - **Maintenance Exclusion / Cluster notification** を使いこなし、快適な GKE アップグレードライフを
- 細かいですが色々アップデートがあります。
 - **Image streaming** でコールドスタートタイムを改善
 - **Spot VM / Pod** を使ってリーズナブルに GKE を運用

Thank you.

