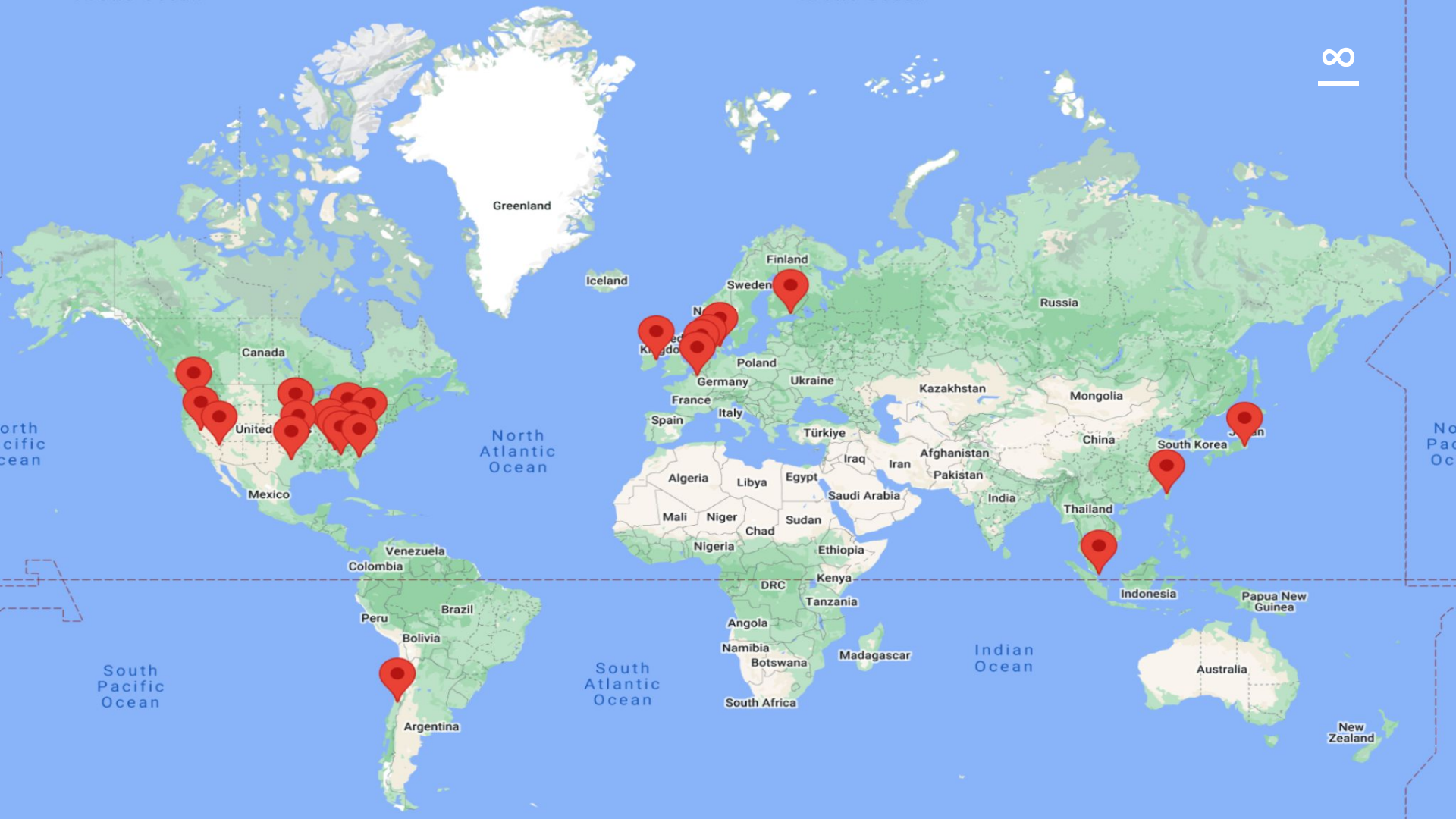
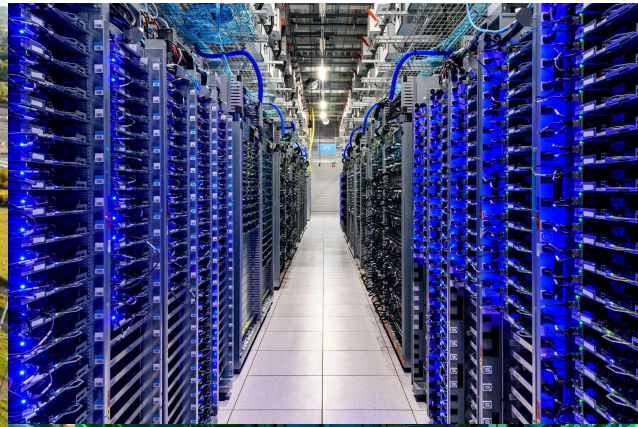
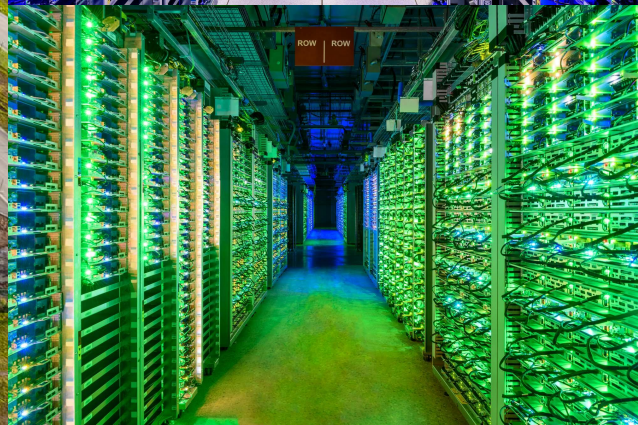
Google Cloud Platform

High Level Overview


Greenland
Iceland
Finland
Sweden
Russia
Canada
United
Mexico
Poland
Germany
Ukraine
France
Spain
Italy
Kazakhstan
Mongolia
China
South Korea
Türkiye
Iraq
Iran
Afghanistan
Pakistan
India
Thailand
Algeria
Libya
Egypt
Saudi Arabia
Mali
Niger
Chad
Sudan
Nigeria
Ethiopia
Kenya
DRC
Tanzania
Angola
Namibia
Botswana
Madagascar
South Africa
Indonesia
Papua New Guinea
Australia
New Zealand
Venezuela
Colombia
Peru
Brazil
Bolivia
Argentina
North Atlantic Ocean
South Atlantic Ocean
South Pacific Ocean
Indian Ocean

St. Ghislain, Belgium





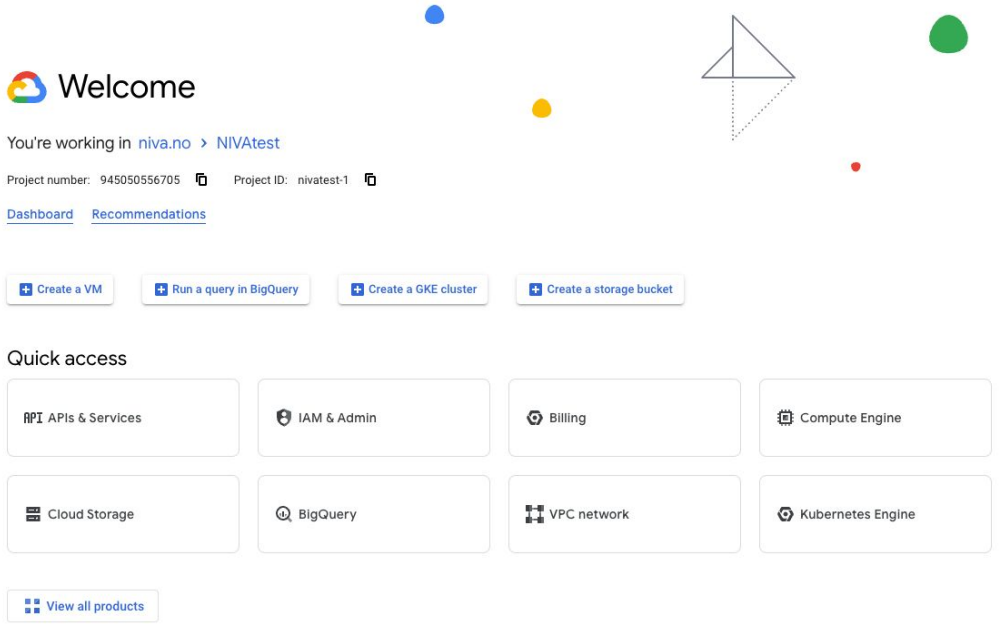
GCP Geographical Hierarchy

- “Region“
  - An independent data center in a specific geographical location, e.g. St. Ghislain, Belgium
  - Regions are referenced by codes
    - Example: europe-west1 (= St. Ghislain)
- “Zone”
  - An independent part of the Region that can fail independently without affecting the other Zones
  - Example: europe-west1-**c**

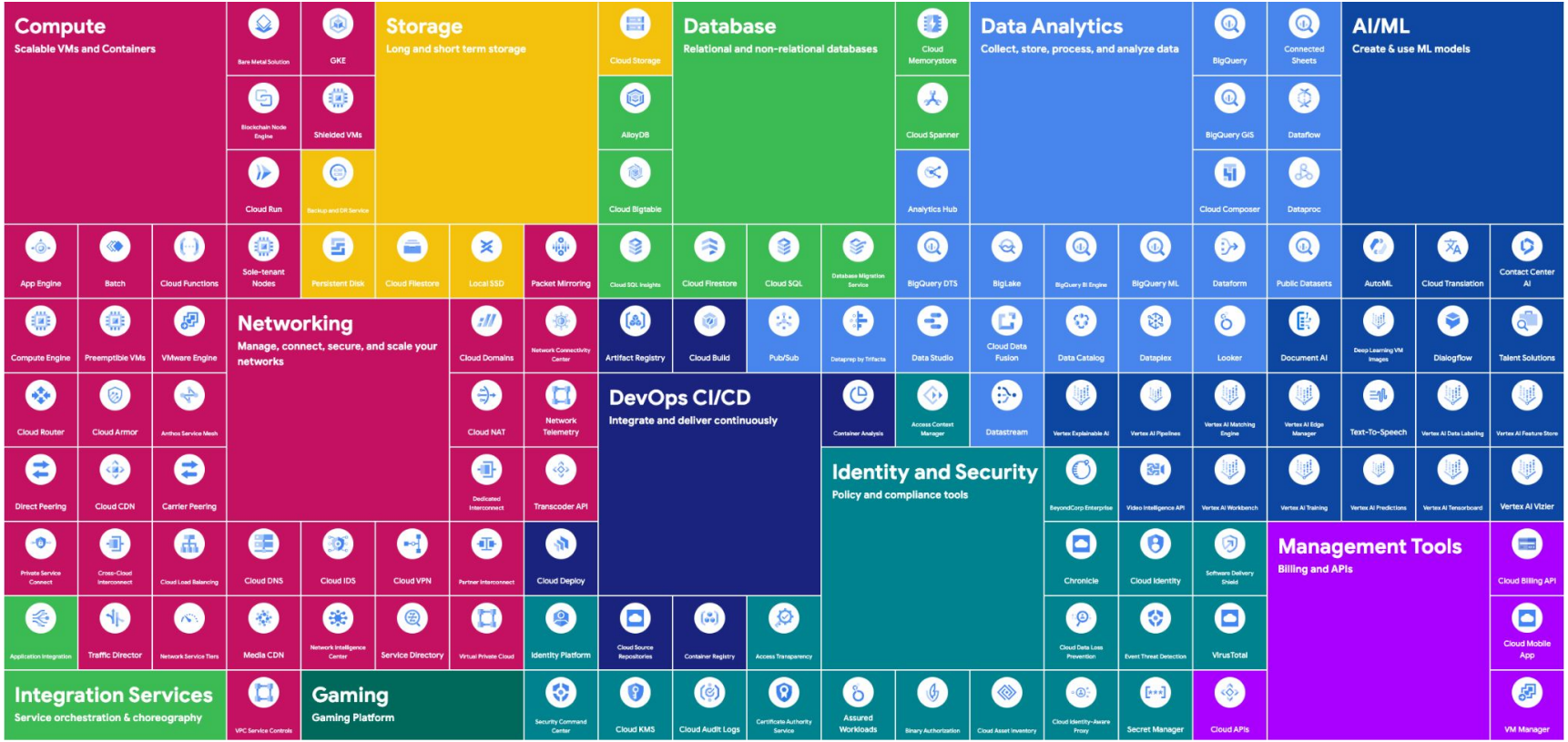
Resources in different zones/regions reduce the risk of outage at the risk of at the expense of increased latency

How to interact with GCP

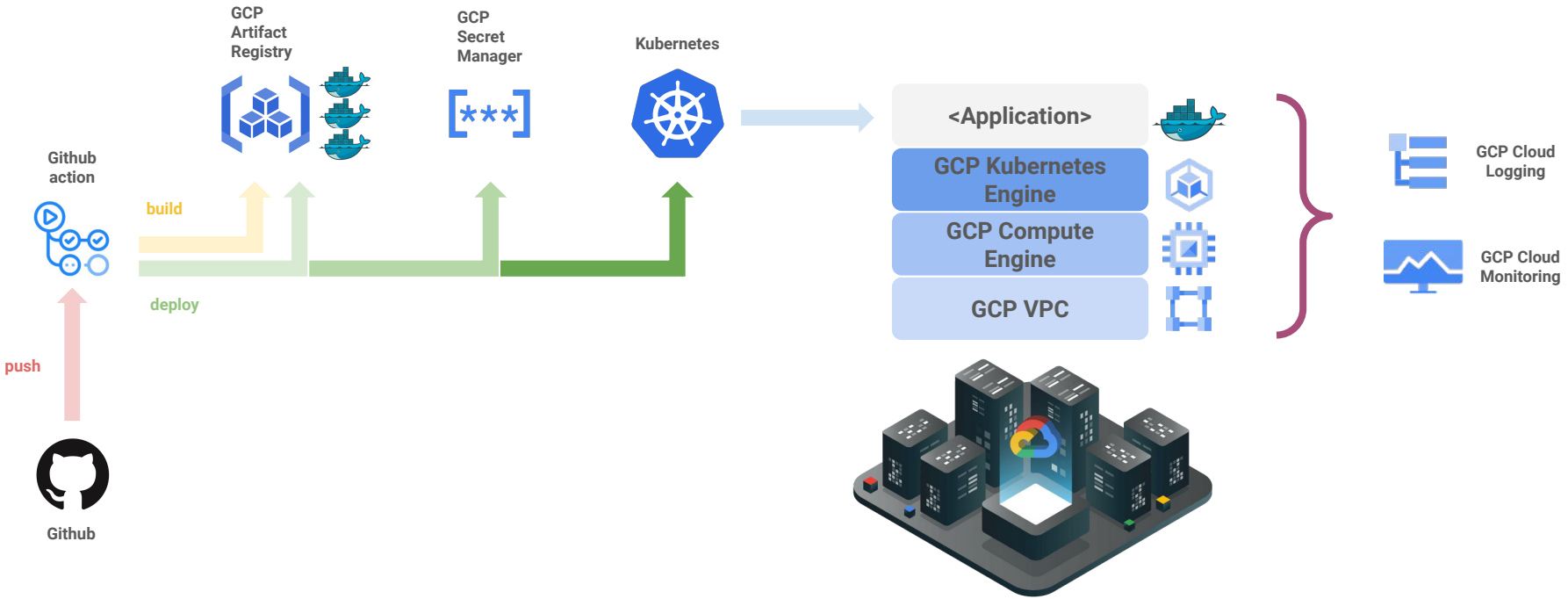
1. Command line gcloud
   a. Examples:
      i. *gcloud config list*
      ii. gcloud storage buckets create gs://my-bucket
2. Cloud Console
3. API



GCP Services Overview

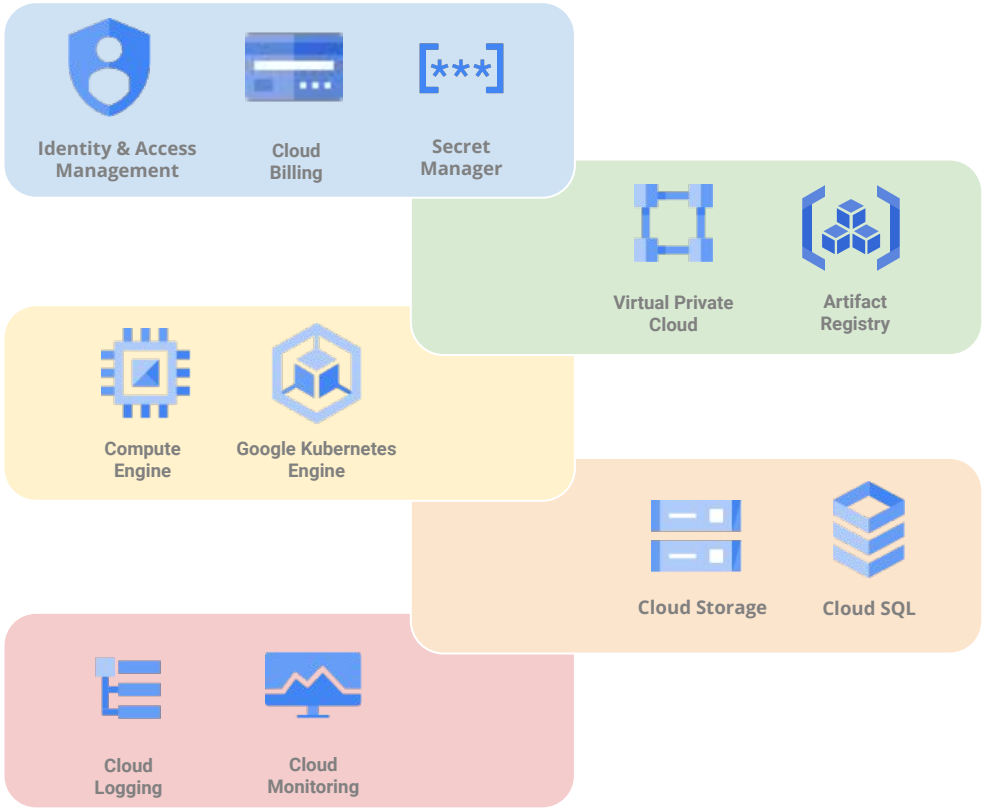


The typical CI/CD pipeline at Niva

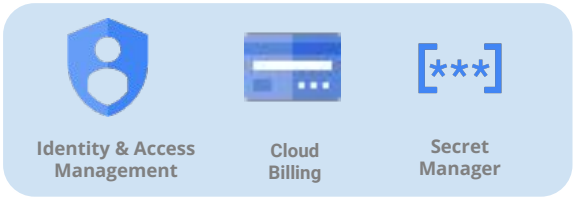


Index

1. Management
    a. IAM
    b. Billing
    c. Secret Manager
2. CI/CD & Networking
    a. Artifact Registry
    b. VPC
3. Computation
    a. Compute Engine
    b. Kubernetes Engine
4. Storage
    a. Cloud Storage
    b. SQL
5. Operations
    a. Cloud Monitoring
    b. Cloud Logging



1. Management



[1a] Identity and Access Management  ∞

IAM manages Accounts, Roles and Permissions

- The concept of Account is associated to “**Principals**”, the WHO in GCP user management
  - **Principals** are identified by an email, and a **Principal** can be:
    - User
    - Service Account
    - Group
- **Roles** are collections of **permissions**. A Policy assigns **Roles** to **Principals** to give **Principals** specific **permission**, the WHAT in GCP user management
  - There are three different typology of of **Roles**:
    - Basic (legacy and not recommended)
    - Predefined (GCP defined, good in most cases)
    - Custom (user-defined)

Resource Hierarchy

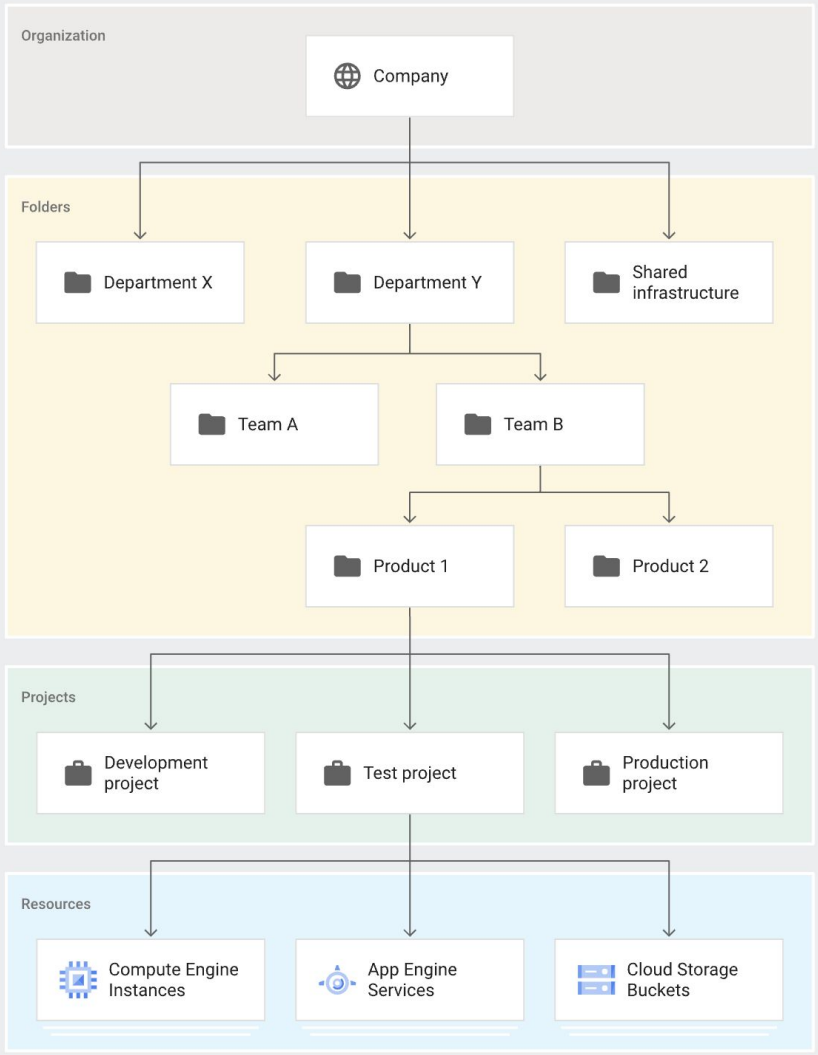
- **Principals**/**Roles** are inherited top-to-bottom.
- Bottom elements can add additional **Roles** (but they cannot remove a **Role**)

In our infrastructure, you will in most cases deal with resources in the Projects:

- ***nivatest-1***
- ***nivaprod-1***

And perhaps:

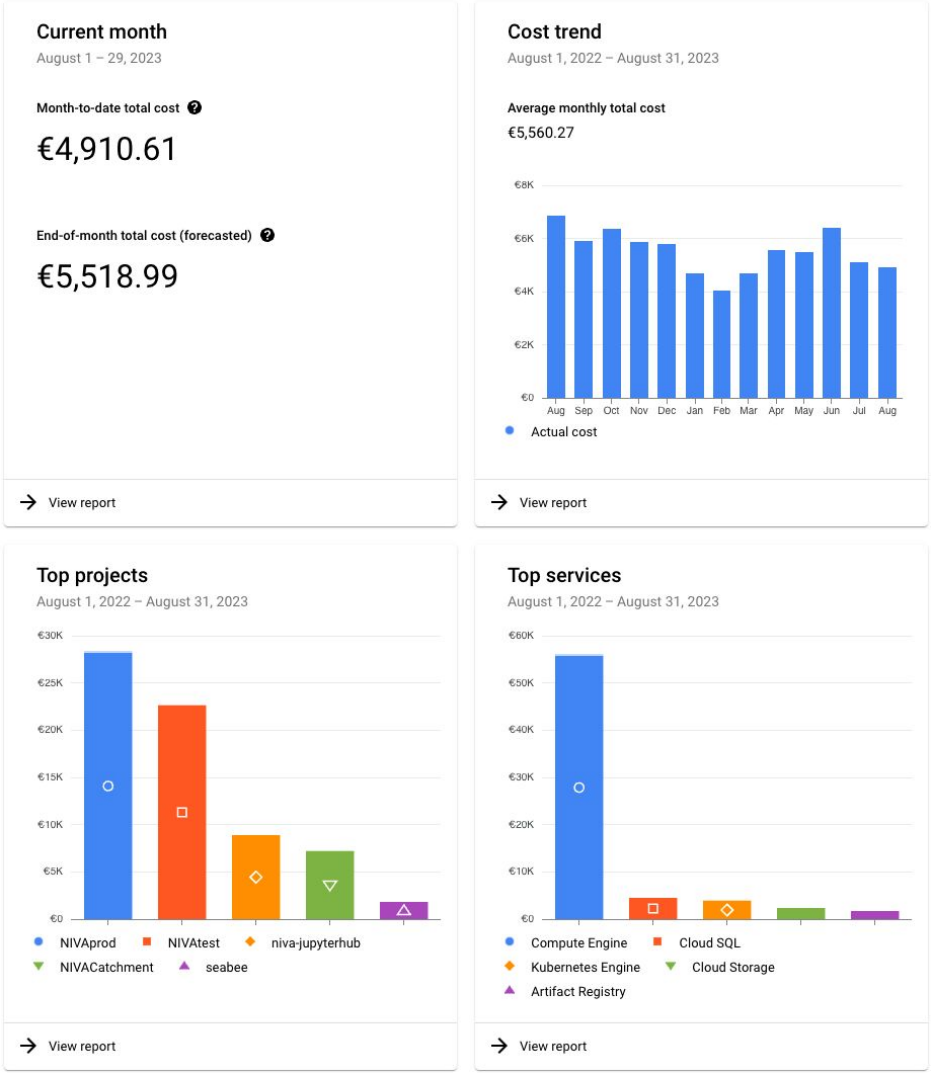
- ***niva-cd***



[1b] Billing

Collection of tools to track GCP spending and optimize costs

Contains both historical information and prediction of future cost based on the current state

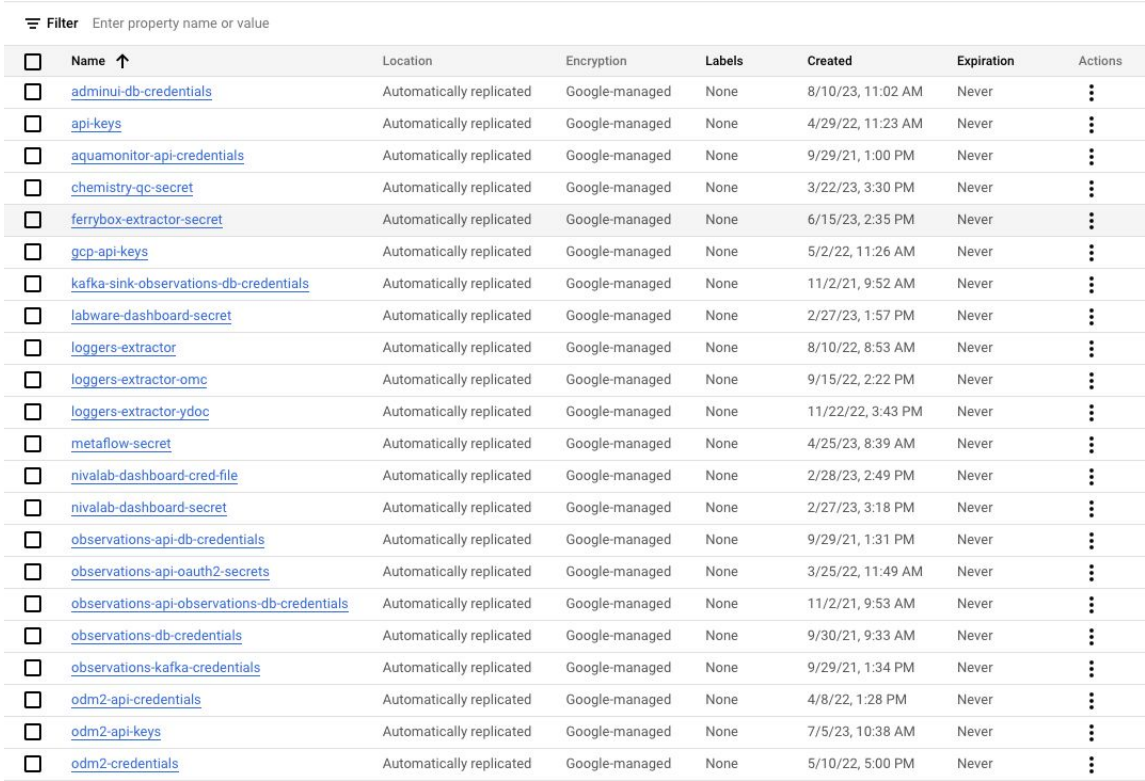


[1c] Secret Manager

The Secret Manager safely stores and manages secrets as binary blobs or text strings.

With the appropriate permissions, it is possible to view the contents of the secret.

Secrets are scoped at the project level



Filter Enter property name or value

Name	Location	Encryption	Labels	Created	Expiration	Actions
adminui-db-credentials	Automatically replicated	Google-managed	None	8/10/23, 11:02 AM	Never	
api-keys	Automatically replicated	Google-managed	None	4/29/22, 11:23 AM	Never	
aquamonitor-api-credentials	Automatically replicated	Google-managed	None	9/29/21, 1:00 PM	Never	
chemistry-qc-secret	Automatically replicated	Google-managed	None	3/22/23, 3:30 PM	Never	
ferrybox-extractor-secret	Automatically replicated	Google-managed	None	6/15/23, 2:35 PM	Never	
gcp-api-keys	Automatically replicated	Google-managed	None	5/2/22, 11:26 AM	Never	
kafka-sink-observations-db-credentials	Automatically replicated	Google-managed	None	11/2/21, 9:52 AM	Never	
labware-dashboard-secret	Automatically replicated	Google-managed	None	2/27/23, 1:57 PM	Never	
loggers-extractor	Automatically replicated	Google-managed	None	8/10/22, 8:53 AM	Never	
loggers-extractor-omc	Automatically replicated	Google-managed	None	9/15/22, 2:22 PM	Never	
loggers-extractor-ydoc	Automatically replicated	Google-managed	None	11/22/22, 3:43 PM	Never	
metaflow-secret	Automatically replicated	Google-managed	None	4/25/23, 8:39 AM	Never	
nivalab-dashboard-cred-file	Automatically replicated	Google-managed	None	2/28/23, 2:49 PM	Never	
nivalab-dashboard-secret	Automatically replicated	Google-managed	None	2/27/23, 3:18 PM	Never	
observations-api-db-credentials	Automatically replicated	Google-managed	None	9/29/21, 1:31 PM	Never	
observations-api-oauth2-secrets	Automatically replicated	Google-managed	None	3/25/22, 11:49 AM	Never	
observations-api-observations-db-credentials	Automatically replicated	Google-managed	None	11/2/21, 9:53 AM	Never	
observations-db-credentials	Automatically replicated	Google-managed	None	9/30/21, 9:33 AM	Never	
observations-kafka-credentials	Automatically replicated	Google-managed	None	9/29/21, 1:34 PM	Never	
odm2-api-credentials	Automatically replicated	Google-managed	None	4/8/22, 1:28 PM	Never	
odm2-api-keys	Automatically replicated	Google-managed	None	7/5/23, 10:38 AM	Never	
odm2-credentials	Automatically replicated	Google-managed	None	5/10/22, 5:00 PM	Never	

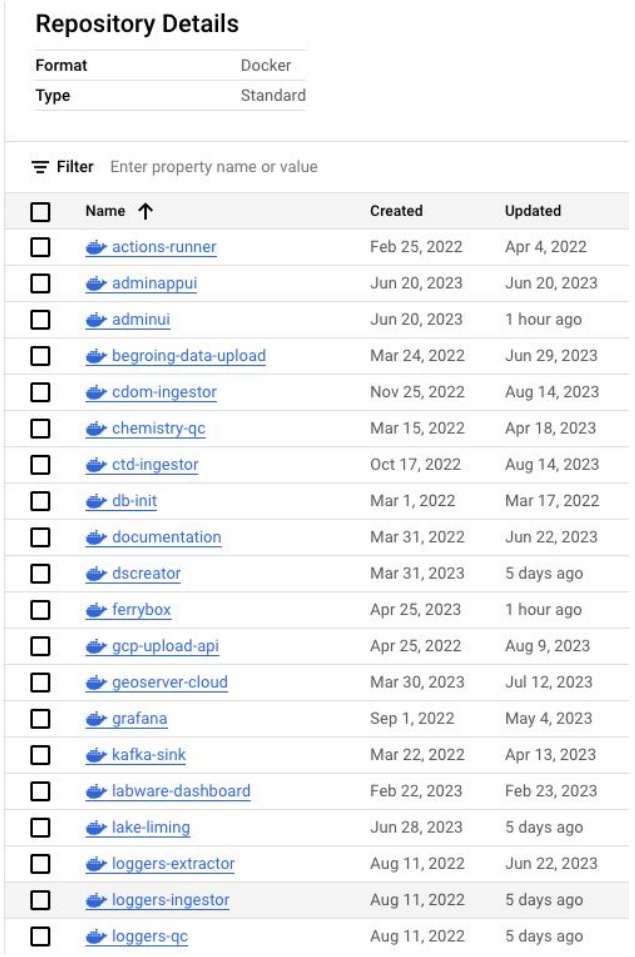
2. CI/CD & Networking

[2b] Artifact Registry

Central storage of artifacts, similar to a cloud bucket

It is deployed in ***niva-cd***, which is reachable from both ***nivatest-1*** and ***nivaprod-1***

Big majority of the artifact is a Docker image embedded in our deployment pipeline



Repository Details

Format	Docker
Type	Standard

Filter Enter property name or value

Name	Created	Updated
actions-runner	Feb 25, 2022	Apr 4, 2022
adminappui	Jun 20, 2023	Jun 20, 2023
adminui	Jun 20, 2023	1 hour ago
begroing-data-upload	Mar 24, 2022	Jun 29, 2023
cdom-ingestor	Nov 25, 2022	Aug 14, 2023
chemistry-qc	Mar 15, 2022	Apr 18, 2023
ctd-ingestor	Oct 17, 2022	Aug 14, 2023
db-init	Mar 1, 2022	Mar 17, 2022
documentation	Mar 31, 2022	Jun 22, 2023
dscreator	Mar 31, 2023	5 days ago
ferrybox	Apr 25, 2023	1 hour ago
gcp-upload-api	Apr 25, 2022	Aug 9, 2023
geoserver-cloud	Mar 30, 2023	Jul 12, 2023
grafana	Sep 1, 2022	May 4, 2023
kafka-sink	Mar 22, 2022	Apr 13, 2023
labware-dashboard	Feb 22, 2023	Feb 23, 2023
lake-liming	Jun 28, 2023	5 days ago
loggers-extractor	Aug 11, 2022	Jun 22, 2023
loggers-ingestor	Aug 11, 2022	5 days ago
loggers-qc	Aug 11, 2022	5 days ago

[2a] Virtual Private Cloud

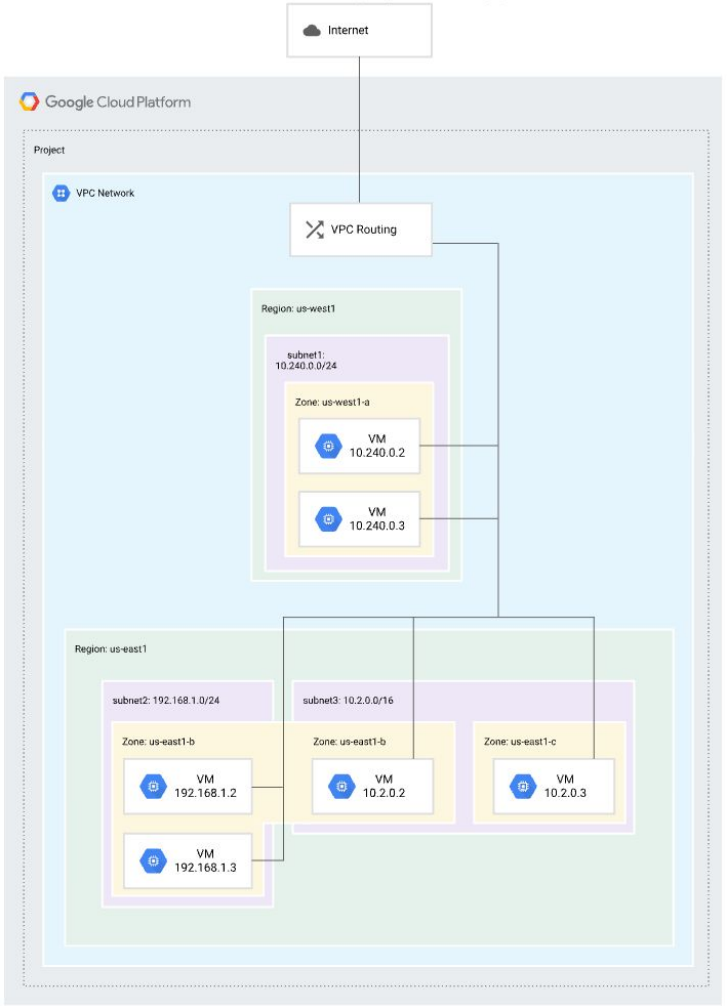
Virtual Private Clouds are network abstractions to isolate different cloud tenants. By default, GCP VPCs are global

VPC networks manage the IP addresses and implement configurable virtual firewalls

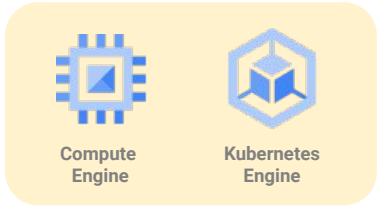
VPCs provide the underlying networking functionality for computing resources (more later)

<Computing>

VPC



3. Computation



[3a] Compute Engine

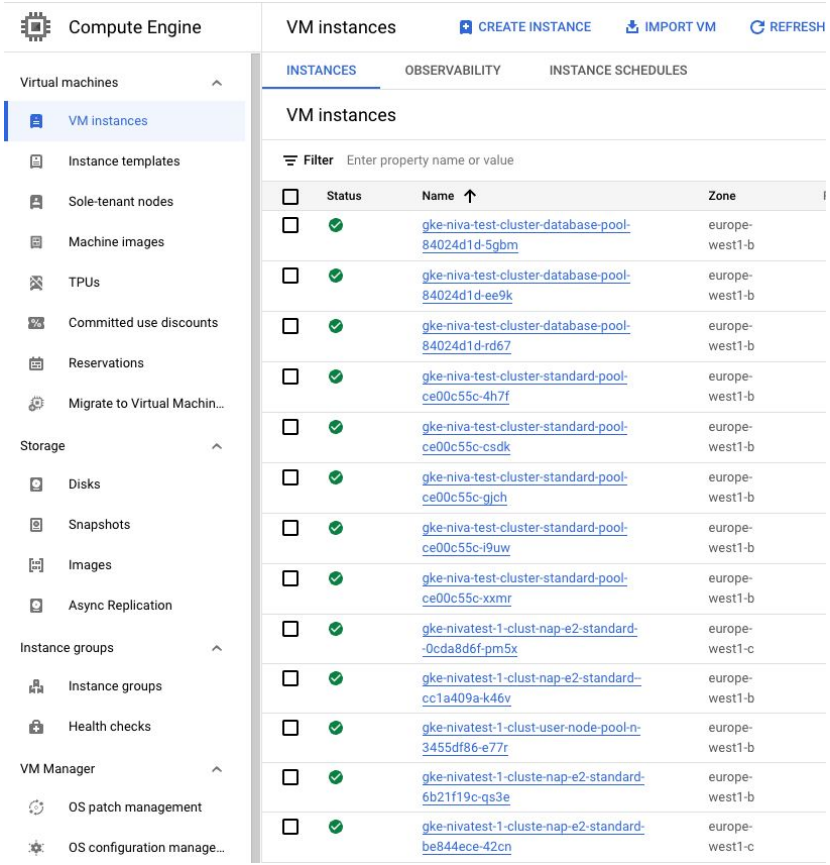
GCP service for managing and deploying VMs

- Availability models:
  - Standard vs spot machines
- Machine families:
  - General purpose
  - Compute optimized
  - Memory optimized
  - …

<Application>

Compute Engine

VPC

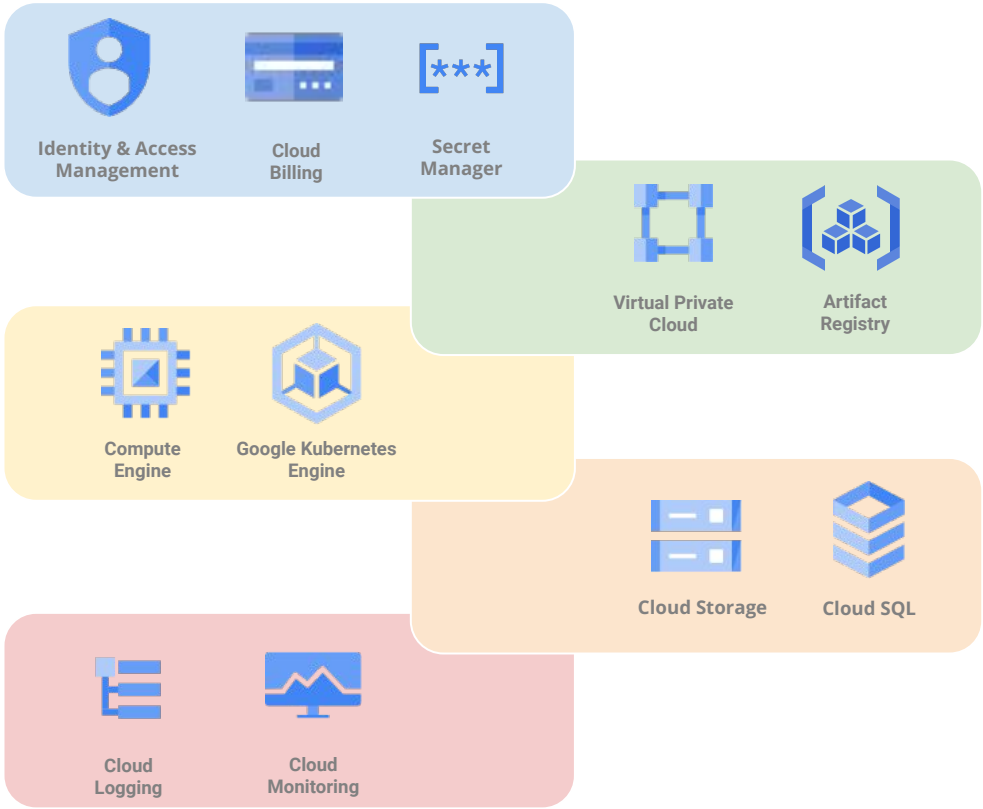


Google Cloud Platform

High Level Overview
Part 2

Index

1. Management
    a. IAM
    b. Billing
    c. Secret Manager
2. CI/CD & Networking
    a. Artifact Registry
    b. VPC
3. Computation
    a. Compute Engine
    b. Kubernetes Engine
4. Storage
    a. Cloud Storage
    b. SQL
5. Operations
    a. Cloud Monitoring
    b. Cloud Logging



[3b] Kubernetes Engine

Google KE is a managed Kubernetes service. Kubernetes is a complex and mature framework to orchestrate the deployment of containerized applications

Clusters: groups of machines acting as Kubernetes workers

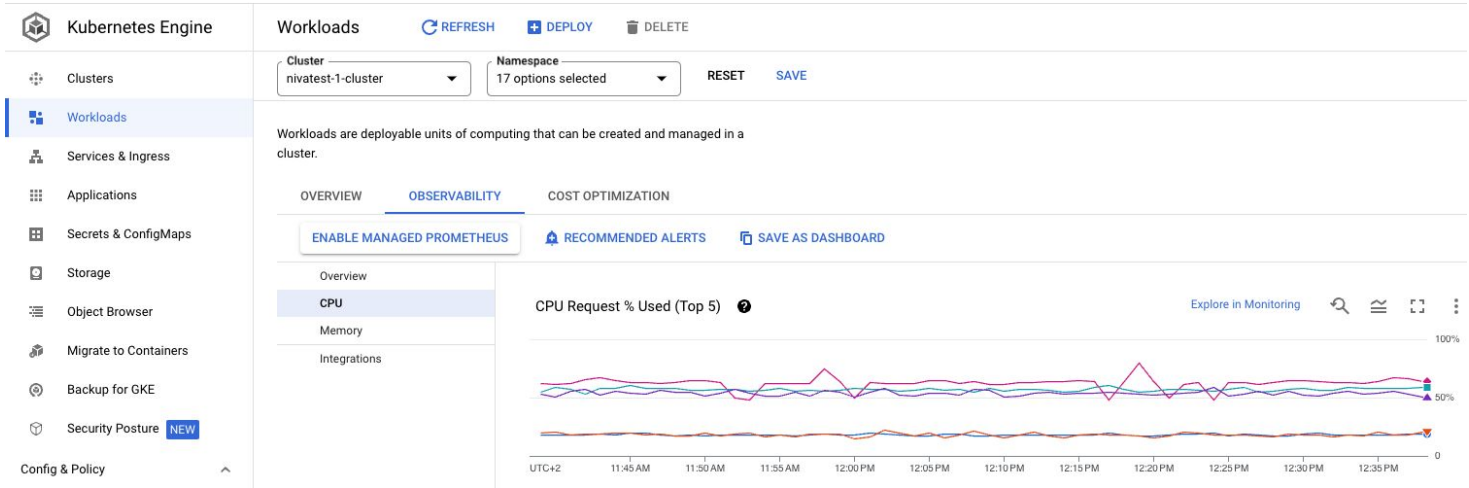
Workloads: collection of Kubernetes components running on the cluster

<Application>

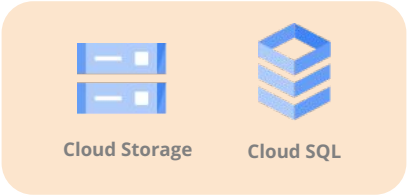
Kubernetes Engine

Compute Engine

VPC



4. Storage



[4a] Cloud Storage

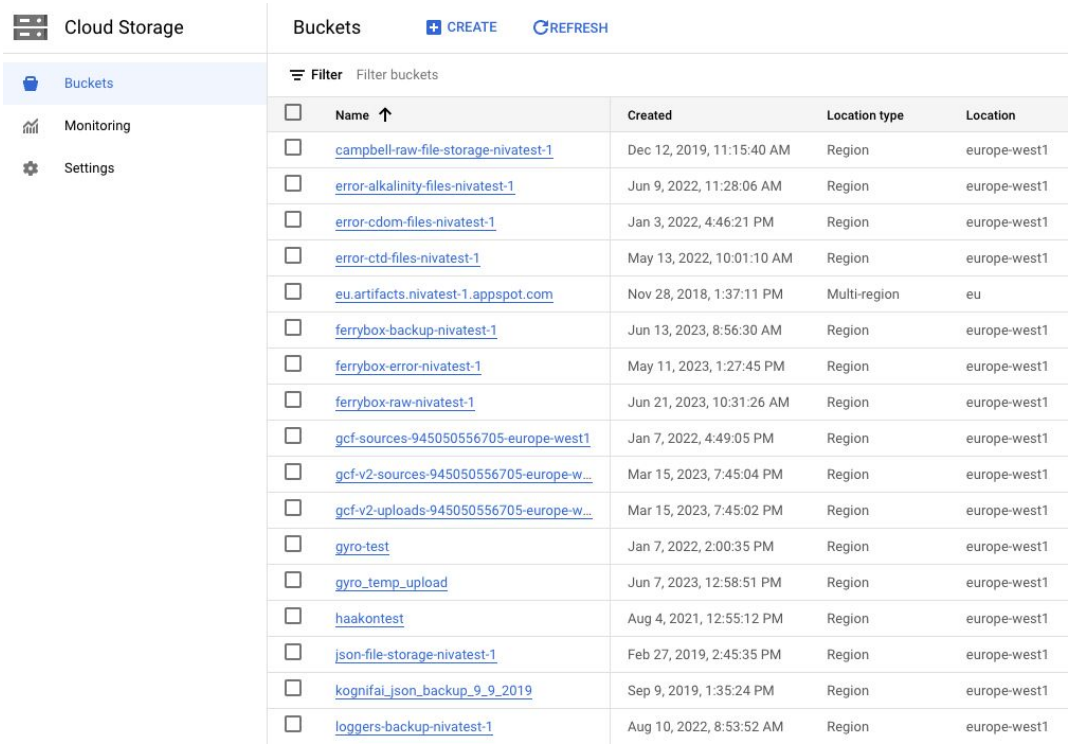
Service for storing unstructured data: files, images, videos etc.

Buckets have a geographic location

Based on the predicted number of accesses, it is possible to select different storage classes:

- Standard
- Nearline
- Coldline
- Archive

Buckets can have fine-grained access control with Access Control Lists (ACLs)



[4b] SQL

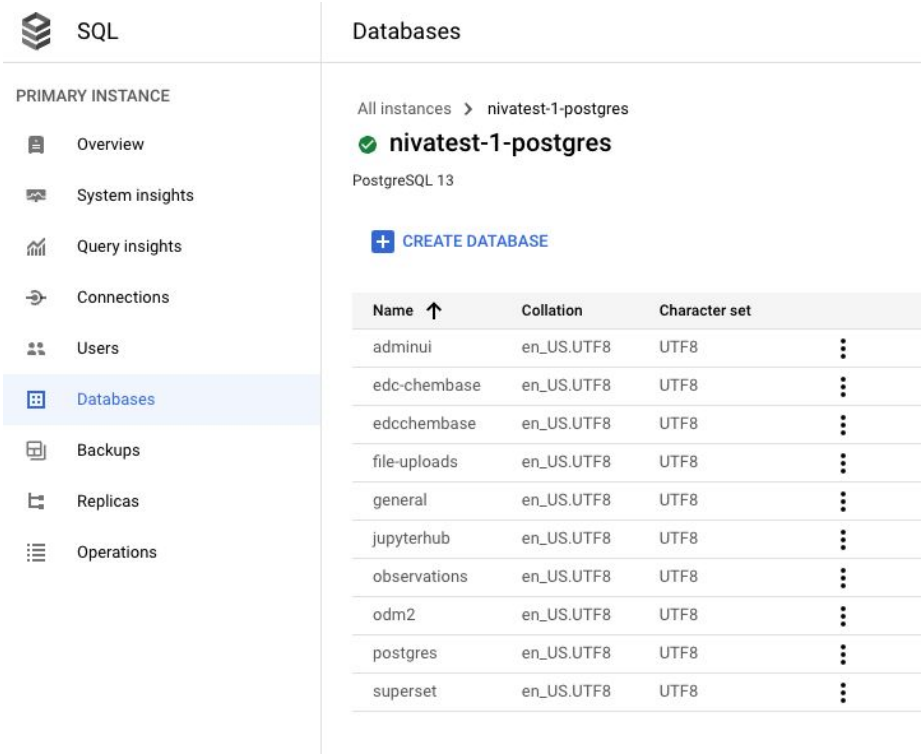
Google SQL is a managed relational DB service for:

- Postgres
- MySQL
- SQL Server

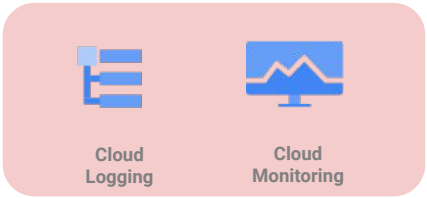
The number of available extension is restricted by GCP

- Timescale for time series handling ✗
- PostGIS for geographic data ✓

Data backup and replication is automated



5. Operations



[5a] Logging

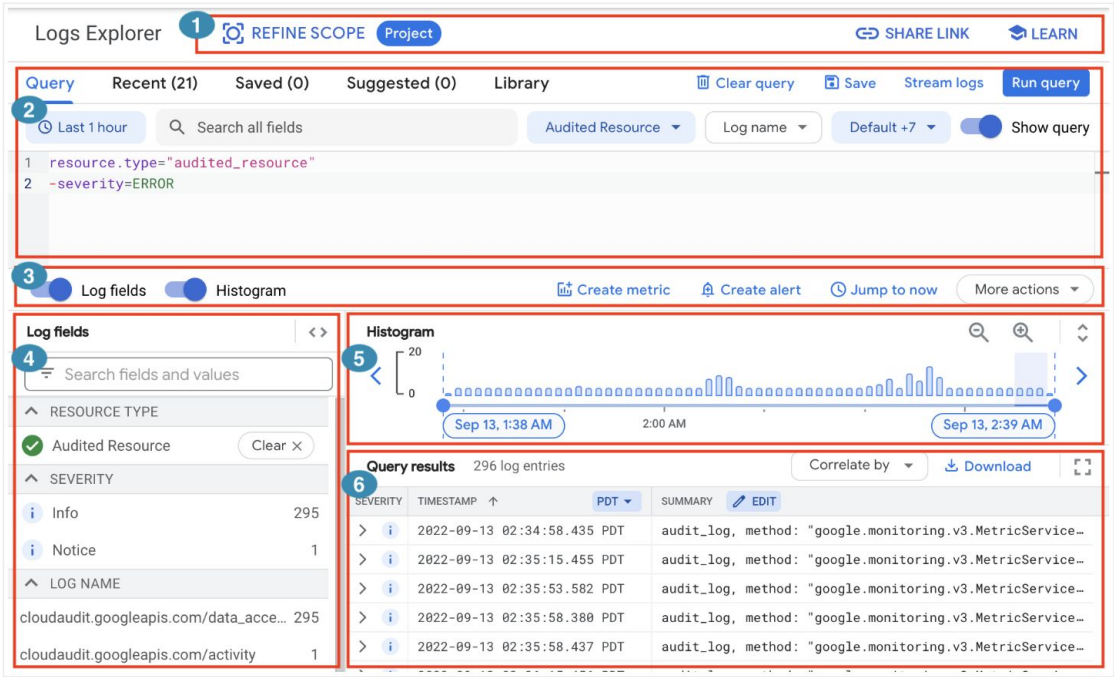
Centralized dashboards and tools to retrieve, view, and analyze log data

Log navigation is done through a simplified query language

Selected features:

- Shareable link
- Filtering capabilities
- Custom alerts

Logs a kept for a limited time known as the retention period

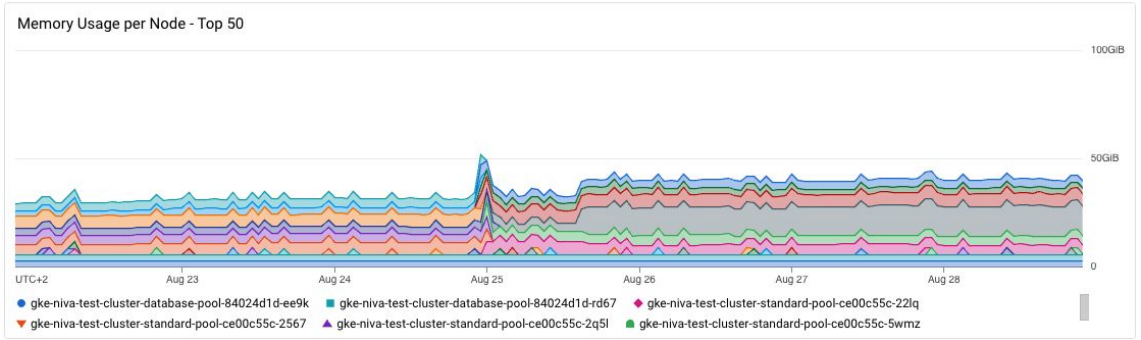


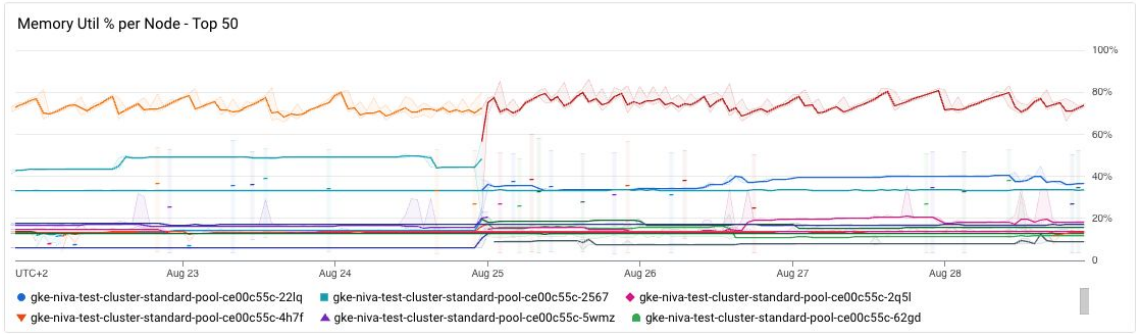
[5b] Monitoring

Centralized dashboards and tools to retrieve and visualize system metrics. Monitoring features are integrated in individual services as well

Features:

- Custom alerts






Thank you

Extra

Example: gcloud ∞

Goal: create a bucket, make it public and access the content over the internet

```
gcloud storage buckets create gs://BUCKET_NAME/ --uniform-bucket-level-access --project=nivatest-1 --location=europe-west1

gcloud storage buckets add-iam-policy-binding gs://BUCKET_NAME/ --member=allUsers --role=roles/storage.objectViewer

gcloud storage cp squirrel.jpeg gs://BUCKET_NAME
```

Example: Pulumi ∞

Goal: create a bucket, make it public and access the content over the internet

```typescript
import * as pulumi from "@pulumi/pulumi";
import * as gcp from "@pulumi/gcp";

// Create a GCP resource (Storage Bucket)
const bucket : Bucket  = new gcp.storage.Bucket("create-bucket", {
    location: "europe-west1",
    name: "pulumi-gen-bucket"
});

// Make the bucket public by adjusting IAM policy
const publicReadBinding : BucketIAMMember  = new gcp.storage.BucketIAMMember("public-read", {
    bucket: bucket.name,
    role: "roles/storage.objectViewer",
    member: "allUsers",
});

const file : BucketObject  = new gcp.storage.BucketObject("upload-file", {
    bucket: bucket.name,
    // replace this with the path to your source file
    source: new pulumi.asset.FileAsset( path: "./dog.jpg"),
});

// Export the DNS name of the bucket
no usages  Andrea Merlina
export const bucketName : pulumi.Output<string>  = bucket.url;
no usages  Andrea Merlina
export const fileUrl : pulumi.Output<string>  = file.selfLink;
```