

UNIVERSIDAD
COMPLUTENSE
DE MADRID



Google Cloud Platform

Ismael Yuste
Strategic Cloud Engineer
NTIC / Marzo 2021



Agenda

- Dia 1
 - [La Nube Pública.](#)
 - Big Data en la Nube
 - Sistemas Batch - Hadoop - Dataproc.
 - Sistemas Streaming - Beam - PubSub y Dataflow
- Dia 2
 - Data Warehouse - BigQuery
 - Notebooks
 - Integración de Datos - CDAP - Data Fusion
 - Orquestación - Airflow - Composer
 - Metadatos - Data Catalog

1

La Nube Pública

Cómo las tecnologías que componen la nube pública nos permiten desarrollar el Big Data como nunca antes.



Why Google Cloud



WHY GOOGLE CLOUD

Choosing Google Cloud

Reasons why people choose us

Trust and security

Keep your data secure and compliant

Open cloud

Scale with open, flexible technology

Global infrastructure

Build on the same infrastructure Google uses





Periodo de Prueba de 90 días o 300 \$ -

<https://cloud.google.com/free/>

Google Data Centers

Los centros de datos de Google son la base de toda la plataforma de Google Cloud. Ofrecen poder computación, almacenamiento, memoria, GPUs para nuestras aplicaciones. Además, alberga el corazón de aplicaciones como Gmail, Youtube, Search...

- Rapidez
- Baja latencia
- Eficiencia de operaciones
- Eficiencia Energética
- Uso de Energías Renovables
- Cercanía al usuario
- Seguridad de la Información



Cloud Locations



REGIONES



ZONAS



UBICACIONES DE PERÍMETRO DE RED



PAÍSES Y TERRITORIOS

DISPONIBLE EN

PRÓXIMAMENTE Seguiremos ampliando la disponibilidad de Google Cloud a las regiones de Varsovia (Polonia), Doha (Qatar), Toronto (Canadá), Melbourne (Australia), Delhi (India), París (Francia), Milán (Italia), Santiago (Chile) y Madrid (España).

Google Datacenters - Cloud Regions

[GCPing](#)



Google Cloud
Platform

Regions

*Exception: region has 4 zones.

Google Cloud

Google Datacenters - Cloud Regions

Google Cloud Platform

Europe

6


Regions


18


Zones

56

PoPs

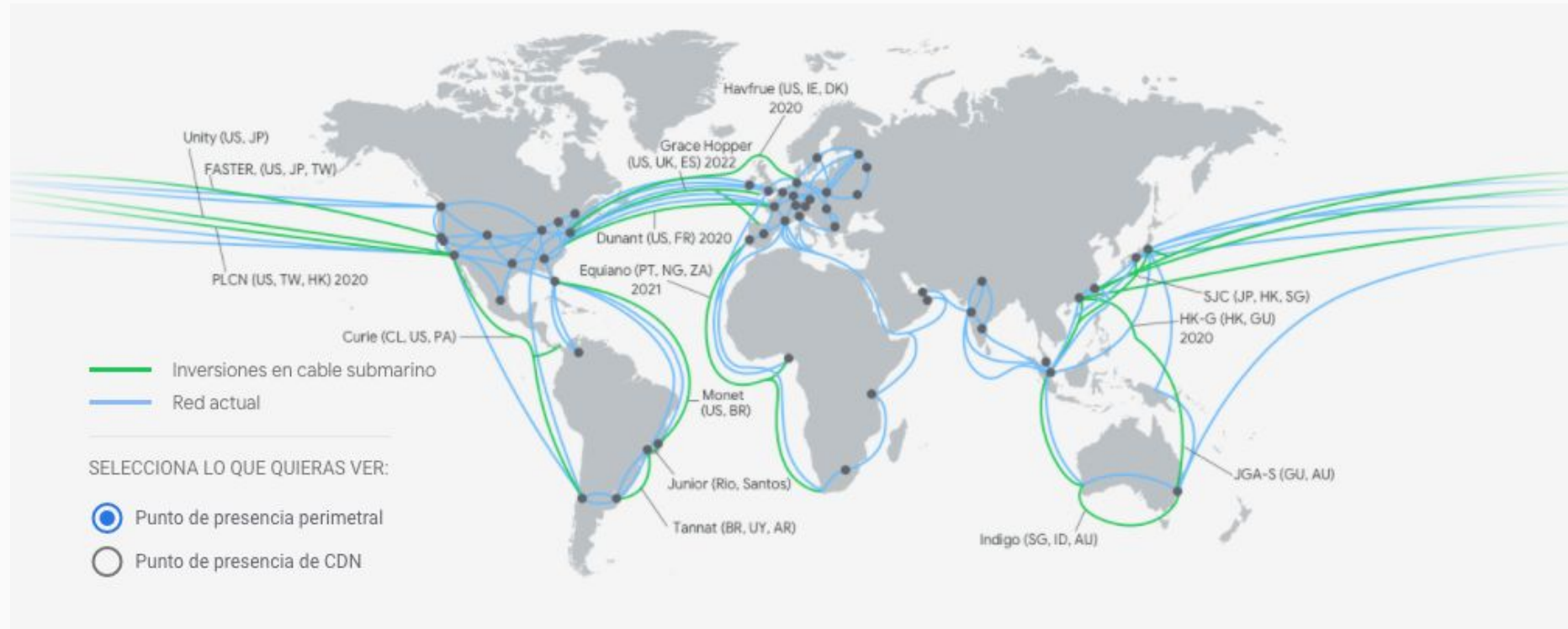
 Current region with 3 zones

 Future region with 3 zones

 Edge point of presence

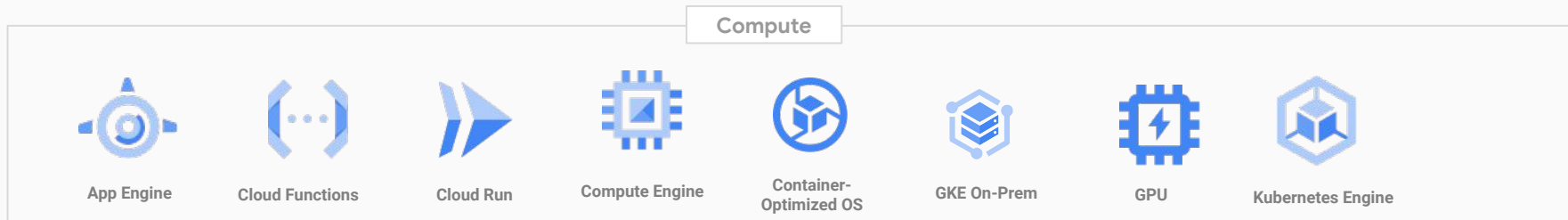


Google Datacenters - Network



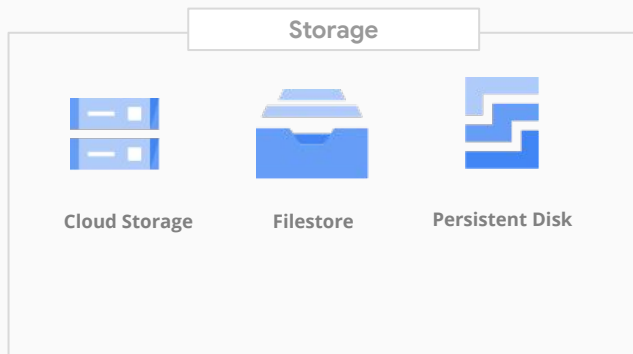
Compute

- ✓ Cost-effective, scalable virtual machines running on Google infrastructure
- ✓ Choose between predefined and custom virtual machines
- ✓ Everything you need for your computing infrastructure



Storage

- ✓ Advanced features like geo-redundancy, edge caching, and WORM compliance built in
- ✓ Easy control over disaster recovery, performance, location, and price
- ✓ Integrate with analytics and ML solutions to get more from your data storage



Google Cloud Storage (GCS)

High-performance object storage

High frequency access



Standard

Most projects start with our Standard class of storage, which is optimized for performance and high frequency access.

Backup and archival storage

Less frequent access



Nearline

Our Nearline class of storage is fast, highly durable storage for data accessed less than once a month.

Low frequency access



Coldline

Our Coldline class of storage is fast, highly durable storage for data accessed less than once a quarter.

Lowest frequency access



Archive

Our Archive class of storage is designed for cost-effective, long-term preservation of data accessed less than once a year.

A single API for all storage classes

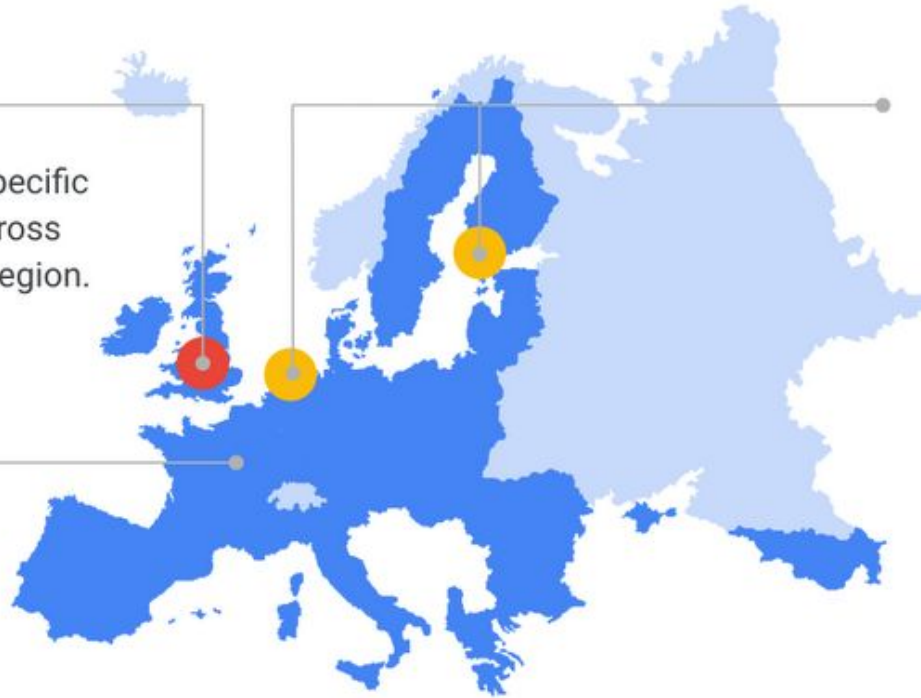
Google Cloud Storage (GCS)

Regional

Your data is stored in a specific region with replication across availability zones in that region.

Multi-regional

Your data is distributed redundantly across US, EU or Asia.



Dual-regional

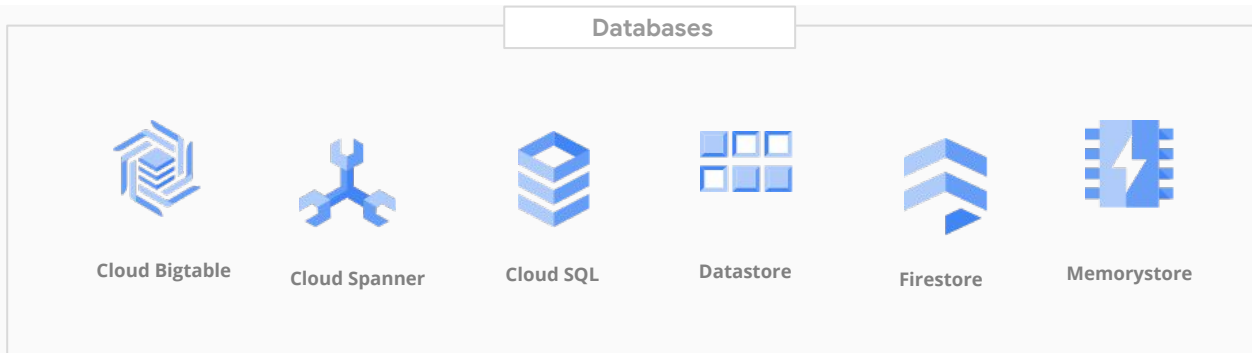
Your data is replicated across a specific pair of regions.

Databases

✓ Flexible performance
and scale

✓ Broad open-source
compatibility

✓ Do more with your data
with analytics and ML/AI
integrations



Databases

DATABASE TYPE	COMMON USES	GCP PRODUCT
Relational	Compatibility Transactions Complex queries Joins	Cloud SQL Cloud Spanner
NoSQL / Nonrelational	Time series Streaming Mobile Web IoT Offline sync Caching Low latency	Cloud Bigtable Cloud Firestore Firebase Realtime Database Cloud Memorystore

Databases

	Relational		NoSQL / Nonrelational			
	Cloud Spanner	Cloud SQL	Cloud Bigtable	Cloud Firestore	Firebase Realtime Database	Cloud Memorystore
Massive scalability	✓		✓	✓		
Data distribution	global	zonal	global	global	zonal	zonal
OSS compatibility		✓	✓			✓
Replica consistency	strong	strong	eventual	strong	n/a	eventual
Multi-primary	✓		✓	✓		
Transactions (strong consistency)	✓	✓		✓		
Joins and complex queries	✓	✓				
Ultra low latency (microsecond or single-digit ms)		✓	✓			✓
Serverless				✓	✓	
Realtime sync to clients				✓	✓	
In-memory						✓

Networking

✓ Protect users, resources, and environment with defense-in-depth network security

✓ Ensure reliability with industry-leading SLAs and reduced lag

✓ Innovate on Google Cloud using protocols like QUIC, BBR, and gRPC

Networking



Cloud Armor



Cloud CDN



Cloud DNS



Cloud External IP Addresses



Cloud Firewall Rules



Cloud Interconnect



Cloud Load Balancing



Cloud NAT



Cloud Network



Cloud Routes



Cloud Router



Cloud VPN



Partner Interconnect



Premium Network Tier



Standard Network Tier



Traffic Director



Virtual Private Cloud

<https://cloud.google.com/products/networking>

Serverless Computing



Cloud Functions



App Engine



Cloud Run

Developer Tools

✓ Quickly write, deploy, and debug cloud-native applications

✓ Build software and standardize CI/CD across all languages

✓ Create pipelines, automate deployments, and get fast feedback

Developer Tools



Cloud Build



Cloud Code



Cloud Code for
IntelliJ



Cloud Scheduler



Cloud SDK



Cloud Source
Repositories



Cloud Tasks



Cloud Test Lab



Cloud Tools for
Eclipse



Container
Registry



Gradle App Engine
Plugin



IDE plugins



Maven App Engine
Plugin



Tools
for PowerShell



Tools for
Visual Studio

<https://cloud.google.com/products/tools>

Data Analytics - Big Data

Big data products

Efficiently capture, process, and analyze data with Google Cloud data analytics products.

Fully managed serverless approach

Seamlessly unlock business insights with Google Cloud Platform's fully managed, proven, end-to-end data analytics products. Remove operational complexities of data analytics with a serverless approach and make important business decisions quickly and efficiently.



<https://cloud.google.com/products/big-data>

Big Data

Data Analytics



BigQuery



Cloud Composer



Cloud Data
Fusion



Data Catalog



Dataflow



Datalab



Dataprep



Dataproc



Genomics



Pub/Sub

AI and Machine Learning

- ✓ **AI Hub**, our hosted repository of plug-and-play AI components, encourages experimentation and collaboration within your organization.
- ✓ **AI building blocks** make it easy for developers to add sight, language, conversation, and structured data to their applications.
- ✓ **AI Platform**, our code-based data science development environment, lets ML developers and data scientists quickly take projects from ideation to deployment.

AI and Machine Learning

AI and Machine Learning



AI Platform



AI Hub



Advanced Solutions
Lab



AutoML



AutoML Natural
Language



AutoML Tables



AutoML
Translation



AutoML Video
Intelligence



AutoML Vision



Cloud Inference
API



Cloud
Jobs API



Cloud Natural
Language API



Cloud TPU



Cloud Translation
API



Cloud Vision API



Data Labeling



Dialog Flow
Enterprise Edition



Recommendations
AI



Speech-to-Text



Text-to-Speech



Video Intelligence
API



Google Codelabs

Google Developers Codelabs provide a guided, tutorial, hands-on coding experience. Most codelabs will step you through the process of building a small application, or adding a new feature to an existing application. They cover a wide range of topics such as Android Wear, Google Compute Engine, Project Tango, and Google APIs on iOS.

[Codelab tools on GitHub](#)

[A-Z](#) [Recent](#)

Choose an event ▾

Cloud ▾

Access files in Cloud Storage with the Spring Resource abstraction

13 minutes Updated December 4, 2020

Learn to access files in Cloud Storage with the Spring Resource abstraction.



Start

Add advanced features to Actions for Google Assistant

Updated October 12, 2020

Learn to develop Actions with advanced features for Assistant, including rich visual responses, custom reprompts and exits, and persistent user storage across...



Start

AI Platform (Unified): Building a fraud detection model with AutoML

32 minutes Updated November 20, 2020

In this lab, you will use AI Platform (Unified) to train and serve a model with tabular data. This unified offering is the newest AI product offering on Google...



Start

AI Platform (Unified): Training and serving a custom model

39 minutes Updated November 23, 2020

In this lab, you will use AI Platform (Unified) to train and serve a TensorFlow model using code in a custom container. This unified offering is the newest AI...



Start

Analyze Clinical Data using BigQuery and AI Platform Notebooks

Updated October 7, 2020

In this codelab, we demonstrate a solution to access and analyze clinical data in GCP using BigQuery and AI Platform Notebooks.



Start

Analyze production performance with Cloud Profiler

Updated October 22, 2020

In this codelab, you'll learn about Cloud Profiler, a tool for continuous collection and analysis of the performance data from production environments with low...



Start

[Google Cloud Platform Blog](#)

[Google Cloud Platform Web](#)

[GCP Twitter](#)

[GCP Podcast](#)

[GCP Youtube](#)

[Hablemos en Cloud Playlist](#)

UNIVERSIDAD
COMPLUTENSE
DE MADRID

