

FIME

Ingeniería en Computación Inteligente

(ICI)

Materia: Computo en la Nube

6°B

Estudiante: Jonathan Guadalupe Alvarado  
Vargas

Maestro: Oswaldo Carrillo

Colima

# Cloud data fusion

## **Concepto:**

Cloud Data Fusion es un servicio gestionado ofrecido por Google Cloud Platform para el procesamiento e integración de datos. Es la versión gestionada de Google de la plataforma open source CDAP.

CDAP es una plataforma de aplicaciones para crear y administrar aplicaciones de datos en entornos híbridos y multinube. Permite a los desarrolladores con abstracciones de datos y aplicaciones acelerar el desarrollo de aplicaciones de datos, abordando una gama más amplia de casos de uso en tiempo real y por lotes.

## **Función:**

Crear un procesamiento de datos ETL, gestionando tanto la extracción como la transformación y la carga de datos mediante el uso de una interfaz gráfica totalmente gestionada. Las siglas ETL significan Extract, Transform and Load que es el proceso que permite a las organizaciones mover datos desde múltiples fuentes, reformatearlos y limpiarlos, y cargarlos en otra base de datos, data mart, o data warehouse para analizar, o en otro sistema operacional para apoyar un proceso de negocio.

## **Características:**

- Un núcleo abierto que entrega una integración híbrida y de múltiples nubes
- Integración en las herramientas de macrodatos líderes de la industria de Google
- Integración de datos mediante la colaboración y la estandarización
- Autoservicio sin código
- Ingeniería de datos colaborativa
- Integración de datos en tiempo real
- Habilitación híbrida

## Ventajas:

- Evita cuellos de botella técnicos y aumenta la productividad
- Disminuye el costo total de propiedad de la canalización
- Compila con una base de administración de datos

## Clientes que utilizan cloud data fusion:

- LiveRamp
- Star Media Group
- Royal Resorts
- Mercado Libre
- BotMarker
- Minka

## Casos de uso:

- Data lakes más seguros y modernos en Google Cloud
- Almacenes de datos ágiles con BigQuery
- Entorno de estadísticas unificado
- Comunicación entre servicios en una configuración de microservicios

## Precios:

Los [precios de Cloud Data Fusion](#) se dividen entre las horas de instancia de la interfaz y las VM de Dataproc. La edición básica ofrece las primeras 120 horas mensuales por cuenta sin costo.

EDICIÓN	PRECIO POR HORA DE INSTANCIA DE CLOUD DATA FUSION	CANTIDAD DE CANALIZACIONES SIMULTÁNEAS ADMITIDAS	CANTIDAD DE USUARIOS ADMITIDOS
Desarrollador	\$0.35	2 (recomendado)	2 (recomendado)
Básica	\$1.80	Ilimitada	Ilimitada
Empresarial	\$4.20	Ilimitada	Ilimitada