

C A S O P R A C T I C O
M O N I T O R I Z A C I Ó N F 1

BIG DATA

POR: ALVARO LUCIO-VILLEGRAS DE CEA



¿QUE TIPO ES Y PORQUE?



"En este deporte los datos ganan carreras"

PROCESAMIENTO A TIEMPO REAL DE DATOS

Desde los inicios del deporte hasta la actualidad.

Cada año se hace mas y mas sabido que los datos son fundamentales en este deporte, ya que el nivel de refinamiento que requieren estos vehiculos para diferenciarse del resto cada vez es mas complicado de conseguir.

LA SOLUCIÓN A LOS PROBLEMA !!

SENSORES

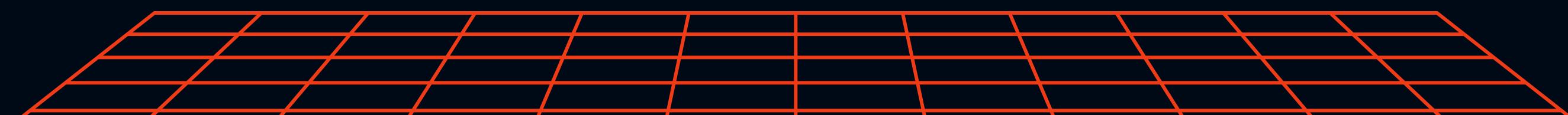
Carga aerodinamica
Desgaste de neumaticos
Nivel de Carga
Fuerzas G

MONITORIZACIÓN

Saber en todo momento el
estado del coche y el vehiculo

OPTIMIZACIÓN

La trazada mas corta de la pista
Entrada en Boxes



SENSORES

Como podemos ver en la imagen que existe una gran cantidad de Sensores, cada uno enviando la información al instante a la sala de control en el Paddock

Ej:

Un de los casos de uso de estos sensores en la amortiguación para dependiendo del estado de la pista se endurece o se hablanda para mayor aderencia del coche.

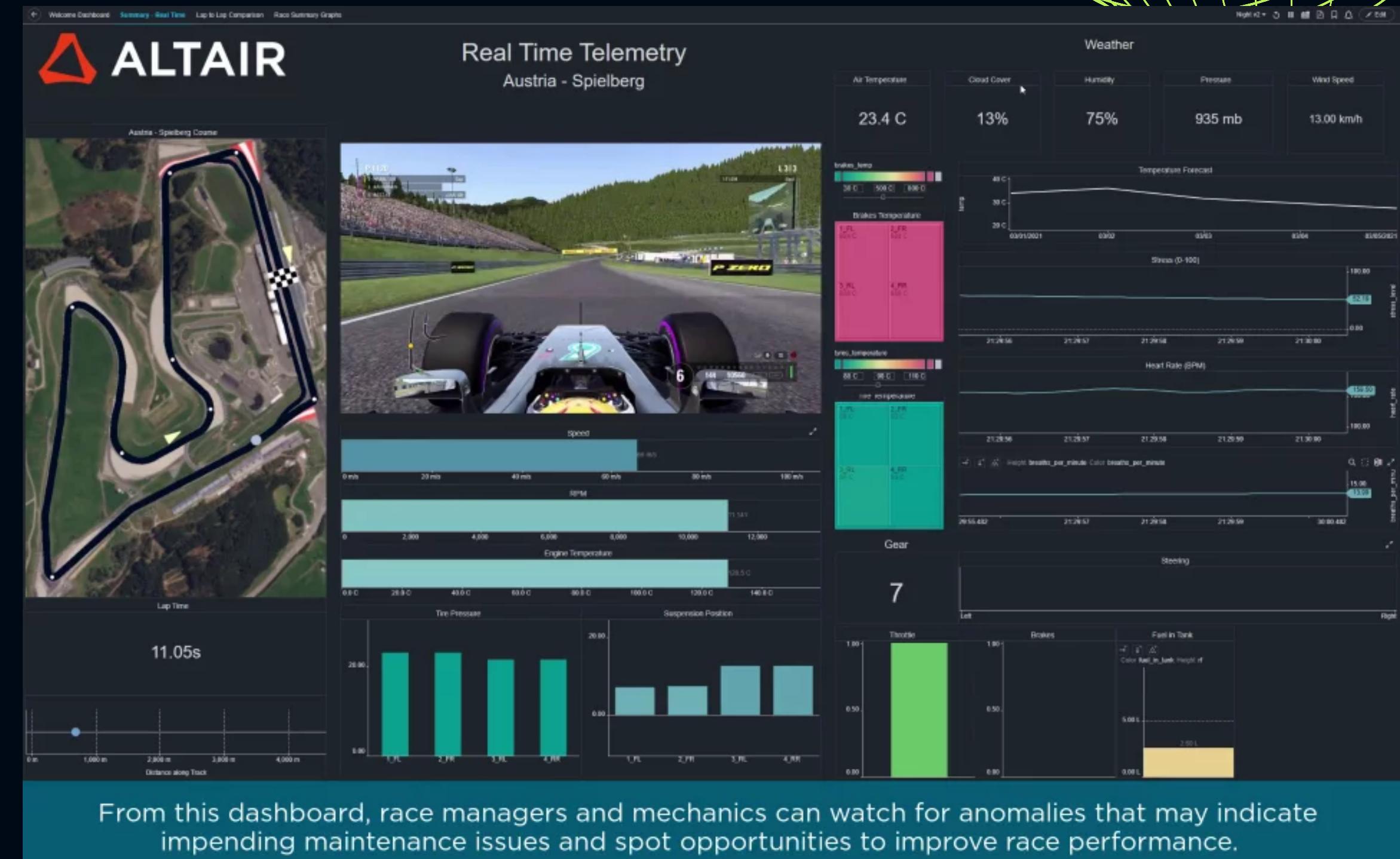


MONITORIZACIÓN

Aquí podemos ver un Dashboard de un coche de F1 en pista. Desde estos paneles se puede ver de un vistazo todo el estado del coche.

Con esta información los ingenieros de fábrica pueden optimizar y mejorar para próximas carreras.

Se recibe la información de los sensores y se les pasan distintos algoritmos de IA para convertir estos datos en información útil para el equipo









OPTIMIZACIÓN

01

Gracias a la aplicación de todas estas tecnologías en este deporte, se ha convertido en una carrera por ver quien construye el mejor coche.

02

Los pilotos tienen mas control sobre el coche y tienen mas ayudas para la conducción del vehículo.

03

Algunas escuderías hacen público los rendimientos en carrera para los mas fans del deporte puedan apreciar el rendimiento del coche.

PUESTA EN MARCHA

Vamos a realizar una simulación de una implantación de estos servicios de Big data en un caso real de F1.

La mayoría de escuderías no usan un servicio "Publico" tienen sus propias plataformas privadas para que la competencia no puedan copiarlos.



SELECCIÓN DE CAMINO

Pack Completos de Servicios.

Mismo Proveedor distintos servicios.



Servicios individuales

- Dashboards a tiempo real
- Almacenaje
- Procesamiento (Sensores del coche ...)
- Análisis (Optimización...)
- Etc

BIG DATA EN LA FORMULA 1

RITTER

Los coches de la Fórmula 1 generan terabytes de información durante una carrera. Docenas de ingenieros se enlazan desde la pista hasta la sede de sus escuderías analizando los datos durante la carrera en tiempo casi real, buscando cualquier ajuste que podría hacer la diferencia para ganar la carrera.

EN EL GRAN PREMIO DE AUSTIN TEXAS EN EL 2014 SE GENERARON 243 TERABYTES DE DATOS CON TODOS LOS VEHÍCULOS PARTICIPANTES.

243 TERABYTES DE DATOS COMPARADOS A...

Categoría	Tamaño (TB)
CAPACIDAD DE 10 CEREBROS HUMANOS	15 TB
DATOS DE TWITTER SEMANALES	42 TB
REGISTRO DE IMPUESTO DE LOS ÚLTIMOS 10 AÑOS	150 TB
TRÁFICO GLOBAL EN INTERNET 1994	200 TB
IBRERIA DEL CONGRESO DE LOS EU	235 TB

EQUIPADOS CON CIENTOS DE SENORES, LOS COCHES DE LA F1 PROPORCIONAN UN FLUJO DE DATOS QUE SE ANALIZA A MILES DE KILOMÉTROS DE DISTANCIA EN TIEMPO REAL.

Medición	Tiempo (seg)	Detalles
.170 seg	.170 seg	Viaje redondo en el que los datos de la carrera se transfieren del Reino Unido a EU.
.300 seg	.300 seg	Viaje redondo en el que los datos de la carrera se transfieren del Reino Unido a Australia.
.600 seg	.600 seg	La diferencia entre el primer y el segundo lugar en el Gran Premio de España en el 2014.
1.923 seg	1.923 seg	Record mundial del pitstop más rápido, por Red Bull en Austin, 2013.

LOS FANS DE LA F1 GENERARON MÁS DE 2.3 TERABYTES DE DATOS MÓVILES DURANTE EL GRAN PREMIO DE ESTADOS UNIDOS GENERADO POR COMPARTIR FOTOS Y MANDAR TWEETS, MENOS DEL 1% COMPARADO CON LOS EQUIPOS DE LA F1.

PRÁCTICA, OCTUBRE 31 CLASIFICACIÓN, NOVIEMBRE 01 CARRERA, NOVIEMBRE 02

The image contains three separate bar charts stacked vertically. The first chart, labeled 'PRÁCTICA, OCTUBRE 31', shows data usage peaking at around 90GB between 2PM and 3PM. The second chart, labeled 'CLASIFICACIÓN, NOVIEMBRE 01', shows usage peaking at around 120GB between 2PM and 3PM. The third chart, labeled 'CARRERA, NOVIEMBRE 02', shows usage peaking at around 150GB between 2PM and 3PM. All charts have time markers from 8AM to 8PM on the x-axis and data volume in GB on the y-axis.

Pack Completos de Servicios.



- Dashboard on stream
- Proceso de IOT (Sensores)
- Bases de Datos
- Soporte tecnico
- APIs

Resumen de la estimación [Información](#)

Costo inicial 0,00 USD	Costo mensual 72.042,99 USD	Costo total de 12 months 864.515,88 USD Incluye el costo inicial
---------------------------	--------------------------------	---

Comenzar con AWS

[Comuníquese con el departamento de ventas](#)

[Iniciar sesión en la consola](#)

Mi estimación

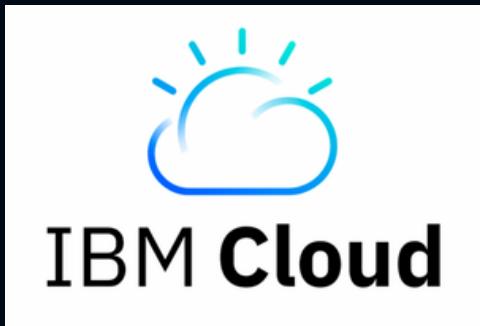
[Find resources](#)

<input type="checkbox"/> Nombre del servicio	Costo ini...	Costo mens...	Descrip...	Región	Resumen de la configuración
Amazon API Gateway	0,00 USD	36.030,02 USD	-	EE. UU. Este (Ohio)	Unidades de solicitudes de la API HTTP (millones), Tamaño promedio de cada solicitud (...)
Plan Enterprise Support	0,00 USD	15.000,00 USD	-	All regions	Supports 24/7 phone, chat, and email access to Cloud Support Engineers for unlimited ...
AWS Application Migration Service	0,00 USD	0,00 USD	-	EE. UU. Este (Ohio)	Número de servidores (12), Número de horas por servidor (6)
AWS IoT Events	0,00 USD	0,00 USD	-	EE. UU. Este (Ohio)	Número de dispositivos (40), Número de mensajes de un dispositivo incluidos los mensaj...
Amazon Kinesis Data Streams	0,00 USD	51,10 USD	-	EE. UU. Este (Ohio)	Duración de la retención de datos (48 horas)
Amazon Detective	0,00 USD	4072,00 USD	-	EE. UU. Este (Ohio)	Datos capturados por AWS CloudTrail, registros de flujo de Amazon VPC, Amazon Guard...
Amazon RDS for Oracle	0,00 USD	16.889,87 USD	-	EE. UU. Este (Ohio)	Almacenamiento para cada instancia RDS (SSD de uso general (gp2)), Cantidad de almac...

72042,99 por mes

*Aviso que estos valores no son los reales pero son los mas cercanos a la realidad.

Pack Completos de Servicios.



- Dashboard on stream
- Proceso de IOT (Sensores)
- Bases de Datos

Detalles:

8587,89 por mes

Estimador de costes	
Productos	Acciones...
5	
Vaya al catálogo de IBM Cloud para explorar los productos disponibles que puede añadir a su estimación.	
Analytics Engine Standard Serverless Spark	4037,50 US\$/me
Mostrar detalles del producto	
Block Storage Espacio de almacenamiento: 4000 GB a 48.000 IOPS	4023,20 US\$/me
Mostrar detalles del producto	
Internet of Things Platform Lite	0,00 US\$/me
Mostrar detalles del producto	
IBM Cloud Monitoring with Sysdig Nivel graduado - Sysdig Secure + Monitor	279,69 US\$/me
Mostrar detalles del producto	
IBM Cognos Dashboard Embedded Pago por uso	247,50 US\$/me
Mostrar detalles del producto	
Total estimado El coste estimado no incluye impuestos	8587,89 US\$/me

Analytics Engine Standard Serverless Spark	4037,50 US\$/me
Plan	
Standard Serverless Spark	
Create a serverless Spark instance with customizations of your choice, and when you submit Spark applications, Spark clusters are created in seconds. Develop and deploy Spark SQL, data transformation, data science, or machine learning jobs using the Spark job API.	
Entradas aproximadas	
50 x Average virtual processor cores required per hour	3687,50 US\$
50 x Average memory (GiB) required per hour	350,00 US\$
500 x Hours of expected usage per month	
Características	
Serverless Spark instances – Spark clusters are created on-demand when workloads are submitted	
Ocultar detalles del producto	

*Aviso que estos valores no son los reales pero son los mas cercanos a la realidad.

Pack Completos de Servicios.



- Packs Prehechos
- Dashboard on stream
- Proceso de IOT (Sensores)
- Bases de Datos
- Soporte tecnico
- APIs

Detalles:

The diagram illustrates the Microsoft Azure Data Pipeline architecture. It starts with data inputs: Business/custom apps (structured) and Sensors and IoT (unstructured). Sensors and IoT data is processed by Event Hubs (1). Structured data goes to Azure Synapse Analytics (Pipelines) (2). Both paths lead to Azure Data Lake Storage (Store) (3). From there, data flows through PolyBase to Azure Synapse Analytics (Apache Spark) (4). This leads to Prep and train (5), then Model and serve (6), and finally to Azure Analysis Services and Power BI (7). Power BI also feeds into Real-time apps via Azure Cosmos DB.

Productos
Azure Synapse Analytics
Azure Analysis Services
Power BI Embedded
Cuentas de almacenamiento
Azure Cosmos DB
Event Hubs

Costo estimado:

Azure Synapse Analytics	Nivel: Optimizado para Compute Gen2, grupos de S... Por adelantado: 0,00 € Mensualmente: 6416,...
Azure Analysis Services	Desarrollador (horas), 1 instancia(s), 720 Horas Por adelantado: 0,00 € Mensualmente: 89,50 €
Power BI Embedded	1 nodos x 720 Horas, tipo de nodo: A2, 2 nodos virt... Por adelantado: 0,00 € Mensualmente: 1361,...
Cuentas de almacenamiento	Redundancia Data Lake Storage Gen2, Estándar y L... Por adelantado: 0,00 € Mensualmente: 14.89...
Azure Cosmos DB	Procesamiento aprovisionado estándar (manual), Ca... Por adelantado: 0,00 € Mensualmente: 257,4...
Event Hubs	Nivel Basic: 1 unidades de procesamiento x 730 Hor... Por adelantado: 0,00 € Mensualmente: 10,31 €

Soporte:

SOPORTE: Professional Direct
PROGRAMA DE LICENCIAS: Contrato de cliente de Microsoft (MCA)
MOSTRAR PRECIOS DE DESARROLLO Y PRUEBAS

Costo inicial estimado: 0,00 €
Coste mensual estimado: 23.974,36 €

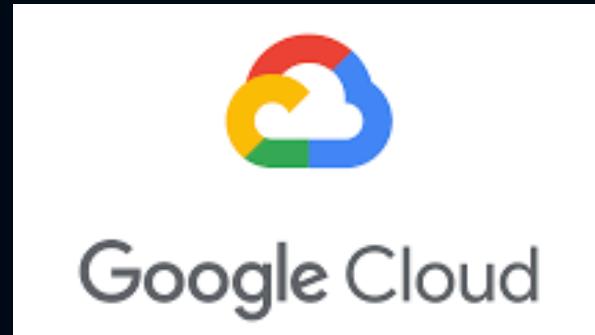
DIVISA: Zona euro: euro (€) EUR

23974,34 por mes

*Aviso que estos valores no son los reales pero son los mas cercanos a la realidad.

Pack Completos de Servicios.

Detalles:



- Dashboard on stream
- Proceso de IOT (Sensores)
- Bases de Datos
- Soporte tecnico Personal
- APIs

Estimate

Datastore

Iowa		
Stored data: 3,072 GiB	EUR 518.48	
Entity Reads: 2,500	EUR 0.00	
Entity Writes: 2,500	EUR 0.00	
Entity Deletes: 2,500	EUR 0.00	
EUR 518.48		
A part of your estimate fits within the Datastore free tier .		

Cloud Operations - Metrics (non-Prometheus)

5		
Volume of monitoring data: 5.569 MiB	EUR 0.00	

49.910.50 por mes

Cloud Operations - Metrics (non-Prometheus)		
5		
Volume of monitoring data: 5.569 MiB		
EUR 0.00		
AI Platform		
Region: United States		
ML Training Units: 33.783		
Job run time: 5,000 minutes		
Prediction Mode: online		
Total node hours: 5,000		
EUR 1,556.50		
Datastream		
Iowa		
Change data capture: 500 GiB		EUR 937.95
Backfill: 500 GiB		EUR 0.00
EUR 937.95		
Premium Support		
Premium Support		
Monthly GCP Charges: EUR 1.00		
Monthly Google Workspace Spend: EUR 1.00		
Premium Support Price: EUR 11,724.45		
TAM Service Level: Multi TAM		
Number of TAMs: 1		
Additional TAM Price: EUR 35,173.12		
EUR 46,897.58		
Total Estimated Cost: EUR 49,910.50 per 1 month		
Estimate Currency		
EUR - Euro		

*Aviso que estos valores no son los reales pero son los mas cercanos a la realidad.

Servicios individuales

01 DASHBOARDS A TIEMPO REAL

Desde los inicios hasta la actualidad.

02 ALMACENAJE

Nuestro campo de actuación internacional y productos disponibles.

03 PROCESAMIENTO (SENSORES DEL COCHE ...)

Informe de resultados.

04 ANALISIS (OPTIMIZACIÓN ...)

Los planes de crecimiento y nuevos mercados.



Servicios individuales

01 DASHBOARDS A TIEMPO REAL

Desde los inicios hasta la actualidad.



The screenshot shows the AWS Pricing Calculator interface. At the top, it displays "My Estimate" and "Editar". On the right, there are "Export" and "Compartir" buttons. Below this, a summary box shows the following details:

Costo inicial	Costo mensual	Costo total de 12 months
0,00 USD	45.141,81 USD	541.701,72 USD Incluye el costo inicial

To the right of the summary is a "Comenzar con AWS" section with links to "Comuníquese con el departamento de ventas" and "Iniciar sesión en la consola".

Below the summary is a table titled "Mi estimación" with the following columns: Nombre del servicio, Costo inicial, Costo mensual, Descripción, Región, and Resumen. The table contains one row for "Amazon Kinesis Data Streams" with values: 0,00 USD, 45.141,81 USD, -, EE. UU. Este (Ohio), and Número de días ...

- 50000 Registros por segundo
- 1KB Tamaño
- 7 Clientes
- Retencion de 1 dia

Servicios individuales

02

ALMACENAJE

Nuestro campo de actuación internacional y productos disponibles.

A screenshot of the Azure Storage Block Blob pricing calculator. The interface shows the following settings:

- Region: East US
- Storage Type: Almacenamiento de blobs en bloque
- Redundancy: Premium (LRS)
- Name Space: Espacio de nombres plano
- Capacity: 1000 TB
- Cost: 144.653,20 €

The calculator also includes sections for "Operaciones y transferencias de datos" and "Operaciones de escritura".

- 1000TB
- Servicio Técnico Personalizado

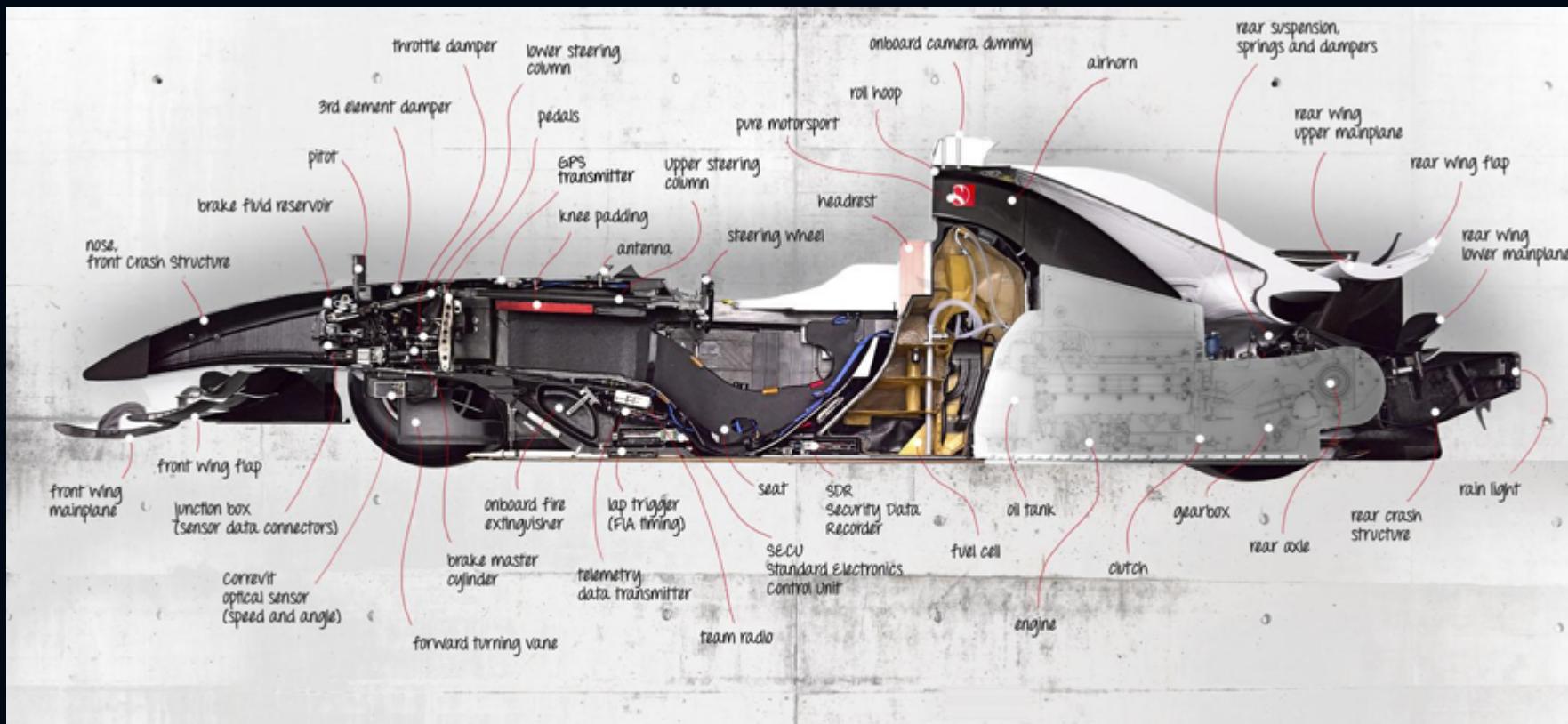
Servicios individuales

03

PROCESAMIENTO (SENSORES DEL COCHE ...)

Informe de resultados.

- 100 Dispositivos
- 25.000 mensajes por dia
- 5MiB por mensaje



*PARTE DE LOS SENSORES

Google Cloud

Estimate

IoT Core



IoT Core

Total Data Exchanged: 380,208,333.333 MiB

EUR 168,698.83

Total Estimated Cost: EUR 168,698.83 per 1 month

Estimate Currency

EUR - Euro

168.698,83€/Mes

Servicios individuales

04 ANALISIS (OPTIMIZACIÓN ...)

Los planes de crecimiento y nuevos merados.

IBM Cloud

Tecnologías :
Hadoop y Spark

The screenshot shows the IBM Cloud Catalog interface. At the top, there's a navigation bar with icons for Catalogo, Estimador de costes, Ayuda, and Preferencias. Below the header, the title "Analytics Engine" is displayed with a subdescription: "Infraestructura flexible para desplegar aplicaciones analíticas de Hadoop y Spark." There are two tabs: "Crear" (Create) and "Acerca de" (About). The "Acerca de" tab is selected. On the left, a sidebar provides service details: Tipo Servicio (Analytics Engine), Proveedor (IBM), Última actualización (11/30/2022), Categoría (Analítica), and Conformidad (Habilitado para HIPAA, Habilitado para IAM, Punto final de servicio soportado). The main content area shows a table for "Planes de precios".

Plan	Características y funciones	Tarifas
Spark estándar sin servidor	Los clústeres de Spark se crean bajo demanda cuando se envían las cargas de trabajo	0,147633 € EUR/Virtual processor core hours 0,014013 € EUR/Gigabyte hours

A descriptive text at the bottom explains: "Cree una instancia de Spark sin servidor con las personalizaciones que elija, y cuando envíe aplicaciones Spark los clústeres de Spark se crearán en segundos. Desarrolle y despliegue trabajos de SQL de Spark, de transformación de datos, de ciencia de datos o de aprendizaje automático utilizando la API de trabajo de Spark. Este plan proporciona un modelo de facturación por segundo adecuado para el escalado bajo demanda y el uso intermitente."

IBM Cloud Search resources and offerings... Catalog Docs Support Manage

Catalog / Analytics Engine

Configure service instance

Hardware configuration

Default (selected) 4 vCPU, 16 GB RAM, 2 x 300 GB HDFS disk on each compute node

Memory Intensive 32 vCPU, 128 GB RAM, 3 x 300 GB HDFS disk on each compute node

Number of compute nodes 1 20 1

Software package AE 1.2 Spark and Hive

Components

• Apache Spark 2.3.2	• Hadoop 3.1.1	• Apache Livy 0.4.1	• Knox 1.0.0
• Ambari 2.7.3	• Anaconda-Py 3.7.1	• Hive 3.1.0	• Jupyter Enterprise Gateway 0.8.0

Ambari configuration (optional)
You can customize Ambari configuration options for cluster components here. You must enter the configuration options in JSON format as described in this [document](#).

Summary United States

Analytics Engine Region: Dallas Plan: Standard-Hourly Instance name: Analytics Engine-hv Cloud service endpoints: public Software package: AE 1.2 Spark and Hive Hardware configuration: Default

Price Management nodes x 1 USD 0.70/hr Compute nodes x 1 USD 0.70/hr

Total price USD 1.40/hr estimated

Create Cancel Add to estimate

Resource list /

Analytics Engine-sv

Active

Add tags

Details

Actions...

Getting started

Manage

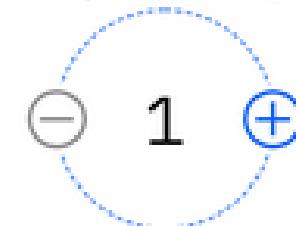
Service credentials

Connections

Plan

Details

Compute nodes



Software package

AE 1.2 Spark and Hive

Components

- Apache Spark 2.3.2
- Apache Livy 0.4.1
- Ambari 2.7.3
- Hive 3.1.0
- Hadoop 3.1.1
- Knox 1.0.0
- Anaconda-Py 3.7.1
- Jupyter Enterprise Gateway 0.8.0

Status

Active

Price estimate

USD **1.40/hour**

United States

Credentials

User ID

clsadmin

All clusters must be password-protected. Set or reset the password for your cluster. See [Credentials in IBM Analytics Engine](#).

[Generate password](#)

Nodes

Node type

Node name

Commission Time

GRACIAS POR SU ATENCIÓN

POR: ALVARO LUCIO-VILLEGRAS DE CEA

