



**BIENVENIDOS**



**QUE TODOS LOS  
HUÉSPEDES QUE ENTREN,  
SE VAYAN SIENDO AMIGOS**



***App Airbnb***

**Juan Ruiz - Martín García - David Melo -  
Santiago Murgueitio - José Carrera**

# SELECCIÓN DEL DATASET OBJETIVO



## **New York Airbnb Open Data 2024**

Airbnb listings and metrics in NYC, NY, USA as of (05 January, 2024)

[k kaggle.com](https://www.kaggle.com)

## ALTERNATIVAS A ANALIZAR

Característica	Kubernetes	Docker Swarm	Apache Mesos
Lanzamiento	2014	2015	2009
Desarrollador	Google	Docker, Inc.	Apache Software Foundation
Modelo de gestión	Declarativo	Imperativo	Mixto (Declarativo e Imperativo)
Escalabilidad	Alta (miles de nodos)	Media (cientos de nodos)	Alta (miles de nodos)
Resiliencia	Alta con auto-reparación	Moderada	Alta
Balanceo de carga	Integrado y configurable	Integrado y más simple	Requiere configuración externa
Compatibilidad	Amplia con CNCF y herramientas de terceros	Buena con herramientas Docker	Requiere integración con Marathon u otras herramientas
Interfaz de usuario	Compleja (kubectl y dashboard UI)	Simple (CLI similar a Docker)	Compleja (requiere interfaces adicionales)
Comunidad y soporte	Muy amplia y activa	Amplia, pero menor que Kubernetes	Amplia en entornos de gran escala
Casos de uso típicos	Aplicaciones a gran escala, microservicios	Proyectos más pequeños o inicios rápidos	Grandes clústeres de datos, aplicaciones intensivas en recursos



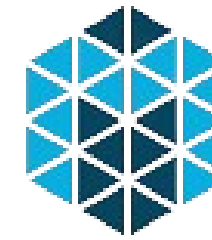
## KUBERNETES

- Completo y complejo.
- Ideal para grandes entornos de producción.
- Perfecto para microservicios.
- Alta disponibilidad y escalabilidad.
- Comunidad activa y soporte constante.



## DOCKER SWARM

- Simple de configurar y usar.
- Bueno para desarrolladores y proyectos pequeños.
- Rápido y menos complejo.
- Escalabilidad adecuada para muchos usos, pero inferior a Kubernetes.

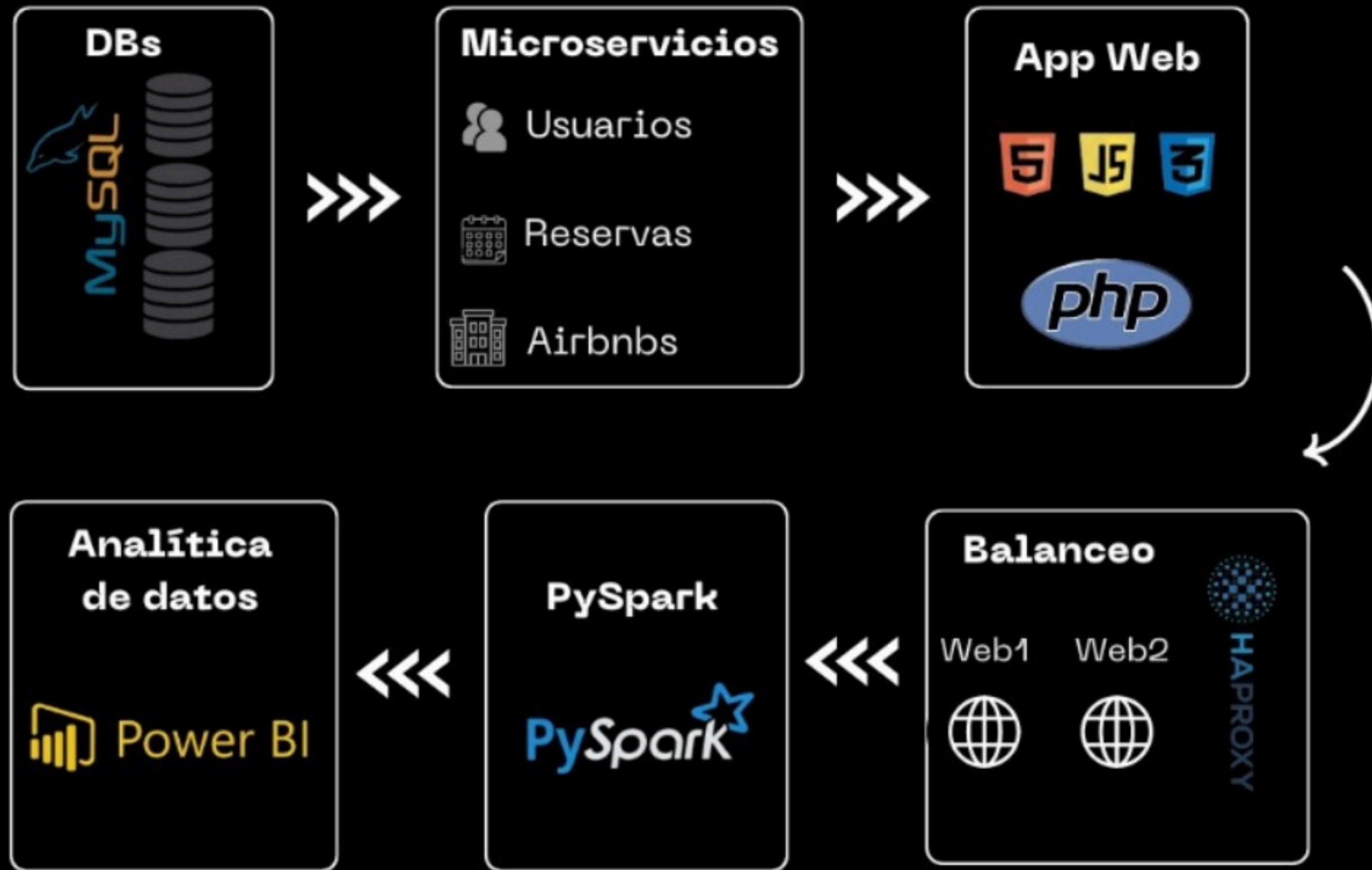


## APACHE MESOS

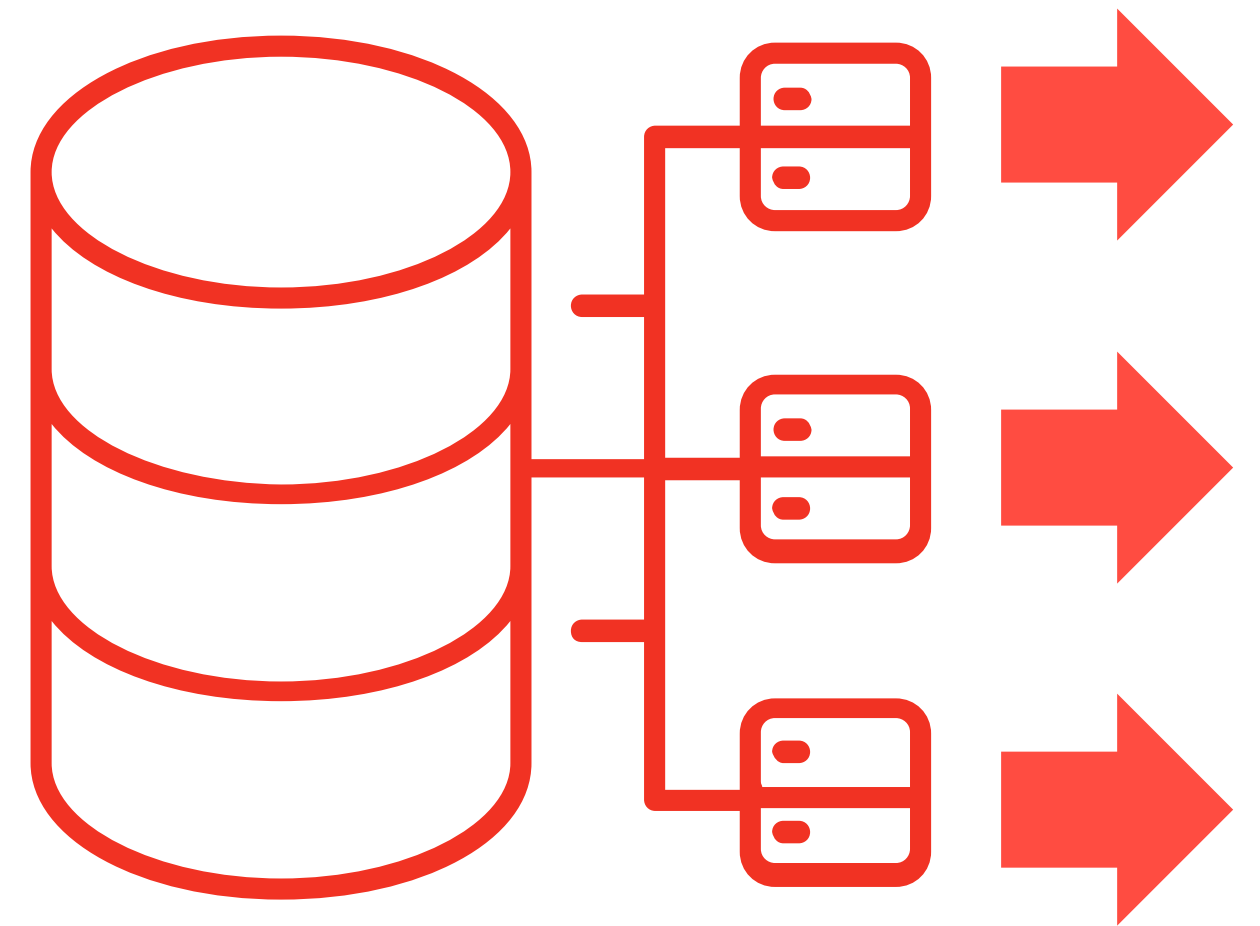
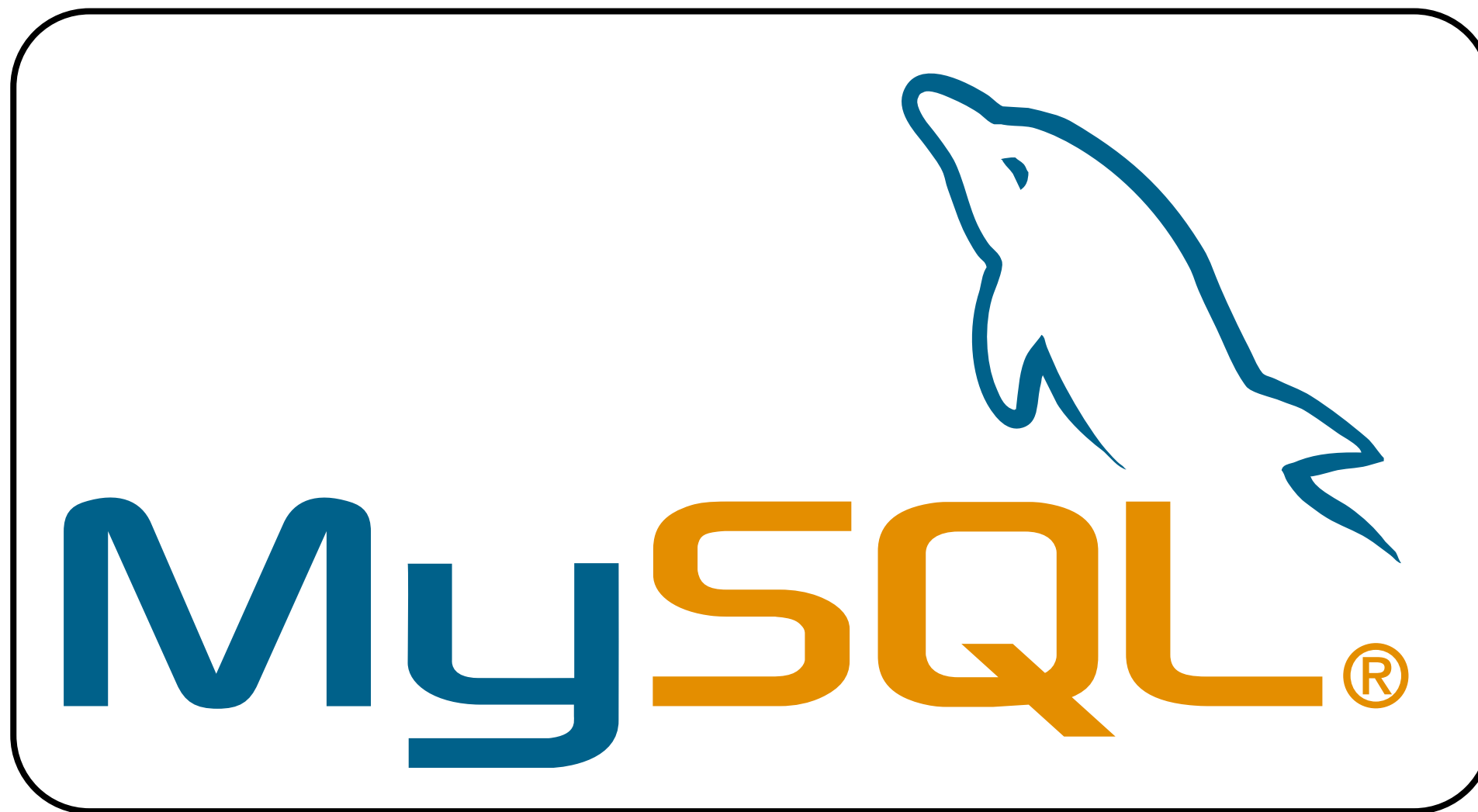
- Orientado a entornos de gran escala.
- Maneja aplicaciones que consumen muchos recursos.
- Efectivo en la gestión de grandes clústeres de datos.
- Comúnmente usado con Marathon para contenedores.



# DEFINICION DE ARQUITECTURA



# **BASES DE DATOS (DB)**



**ALMACENA Y GESTIONA LOS DATOS PERSISTENTES DEL SISTEMA**

# MICROSERVICIOS

MICROSERVICIOS

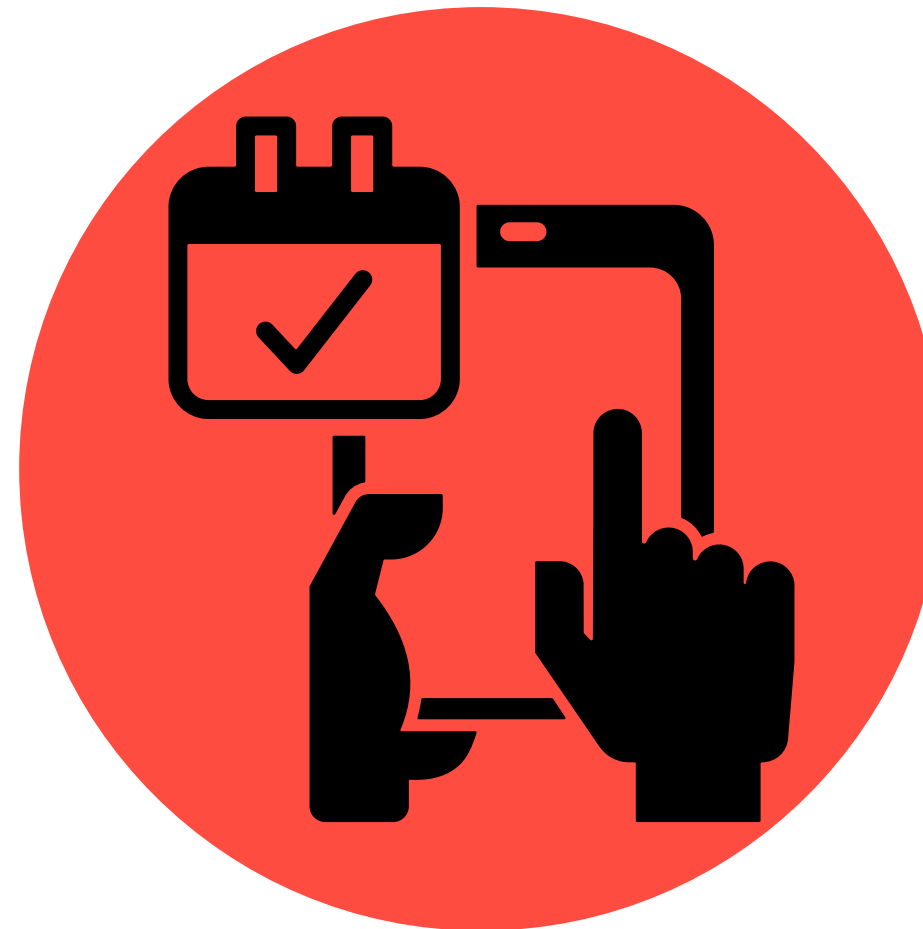


CREDENCIALES DE ACCESO

INFORMACION DEL PERFIL

HISTORIAL DE ACTIVIDAD

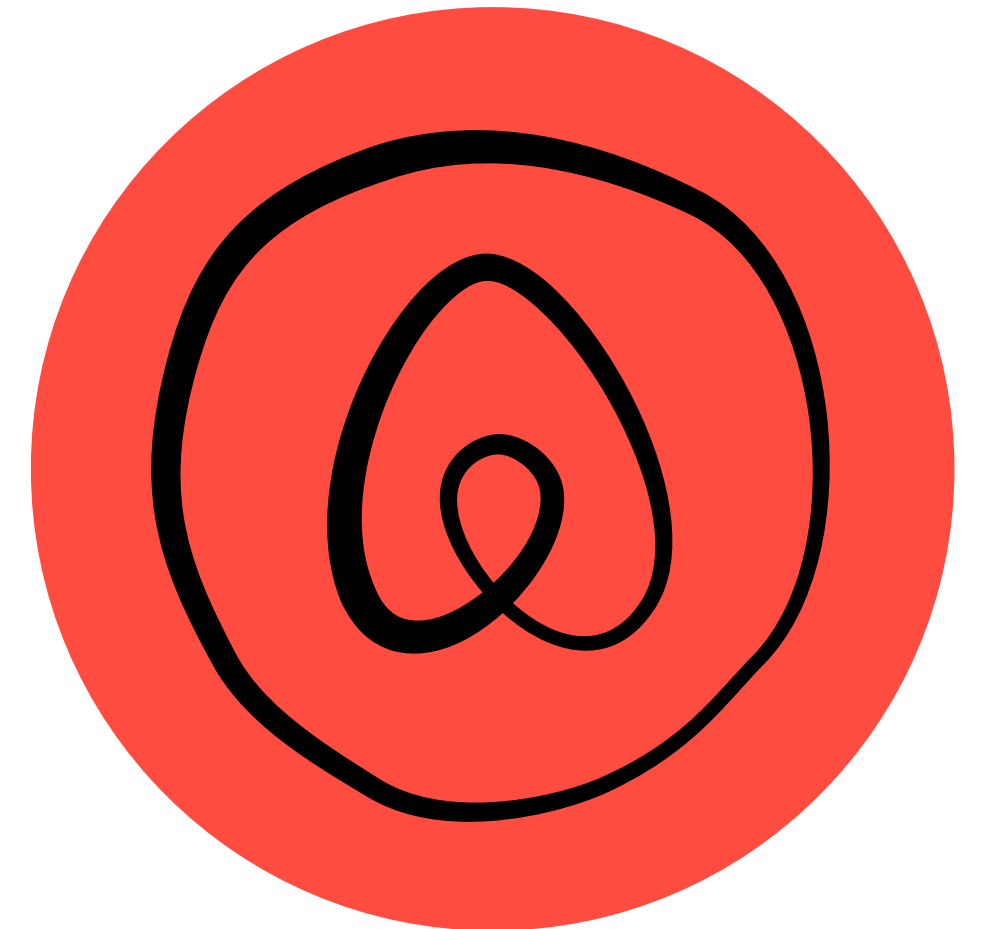
MICROSERVICIOS



DETALLES DE LA RESERVA

ESTADO DE LA RESERVA

MICROSERVICIOS



DETALLES DE LA PROPIEDAD



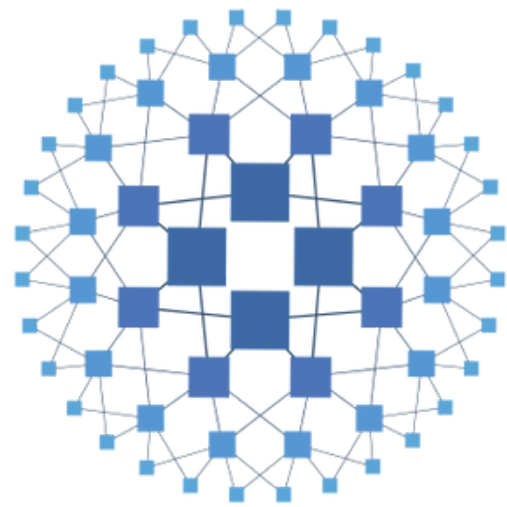
# **APP WEB**

**HTML**



**INTERFAZ INTERACTIVA Y AMIGABLE PARA QUE LOS USUARIOS PUEDAN  
INTERACTUAR CON EL SISTEMA**

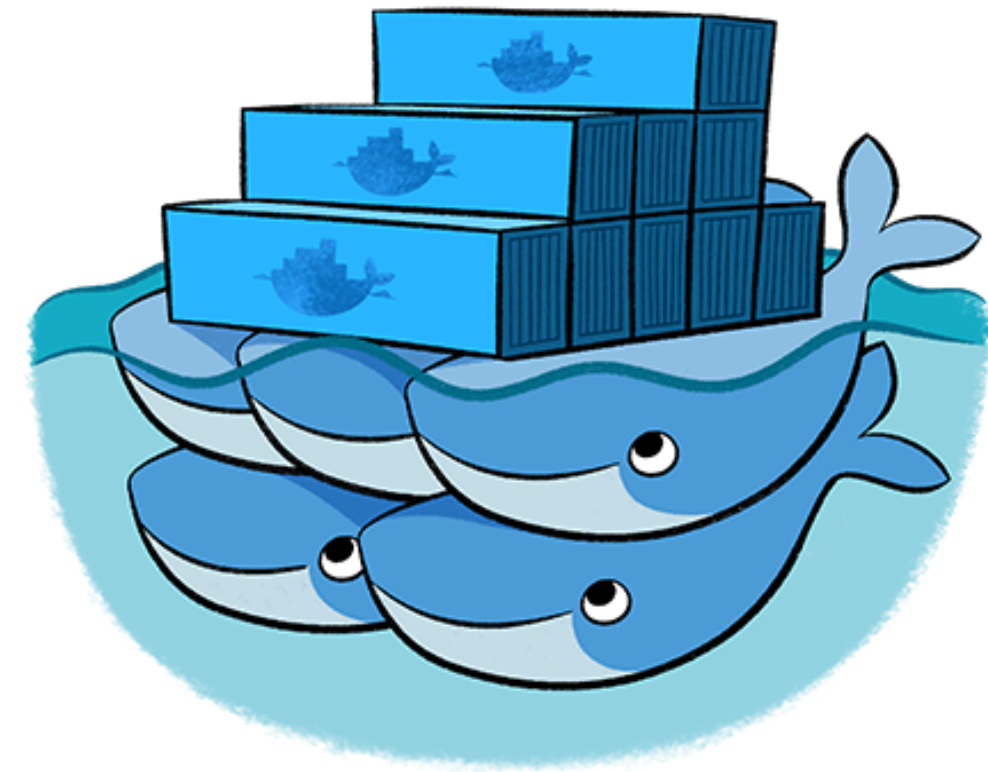
# BALANCEO DE CARGA



HAPROXY

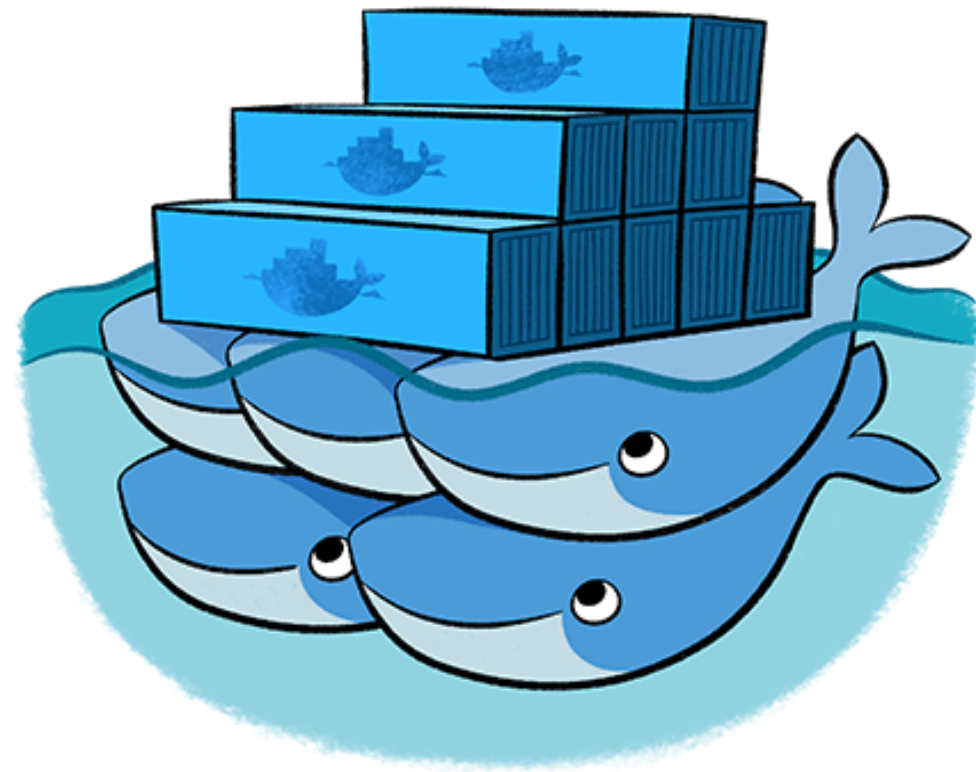
WEB 1

WEB 2



## EMPAQUETADO

# DOCKER SWARM



## COMPUESTO POR

**DB**

**USUARIOS**

**AIRBNBS**

**RESERVAS**

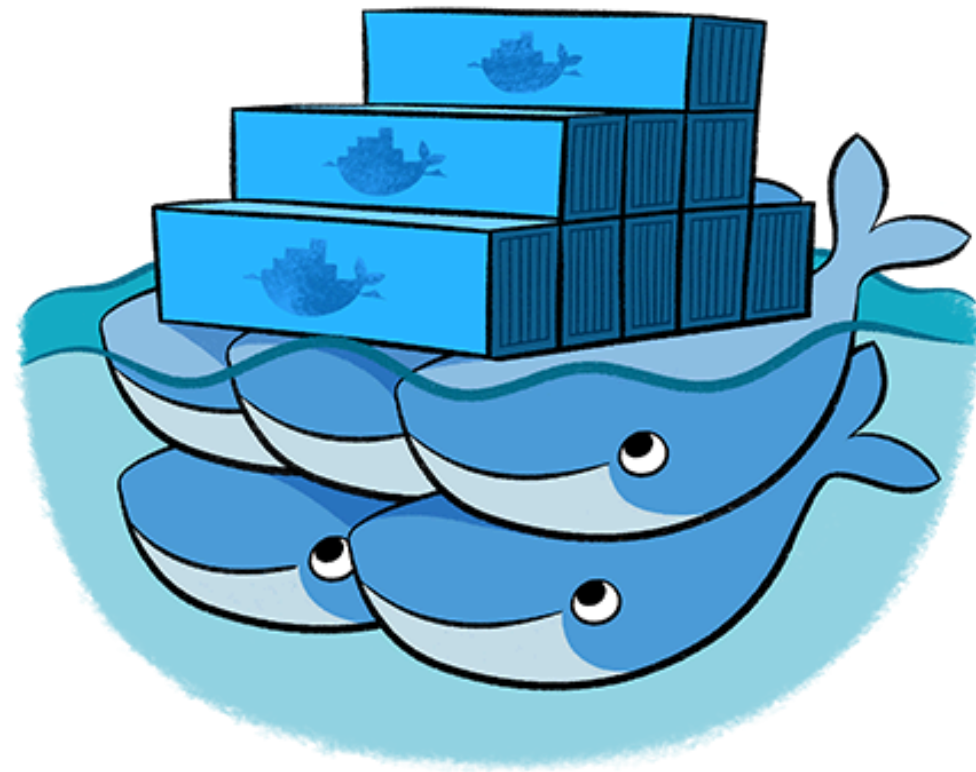
**WEB1**

**WEB2**

**HAPROXY**

**TRANSFERENCIA\_DF**

# DOCKER SWARM



## USO DE DEPLOY - PLACEMENT - CONSTRAINTS

```
web2:
  image: mitgar14/web-airbnbs
  depends_on:
    - usuarios
    - airbnbs
    - reservas
  deploy:
    placement:
      constraints:
        - node.hostname == workerAirbnb
```

```
haproxy:
  image: mitgar14/haproxy-airbnbs
  depends_on:
    - web1
    - web2
  ports:
    - "5080:80"
  deploy:
    placement:
      constraints:
        - node.hostname == serverAirbnb
```

<http://server-airbnb.eastus2.cloudapp.azure.com:5080>

The screenshot displays the Microsoft Azure portal interface. The top navigation bar shows the user is logged in as 'martin.garcia@uao.edu...' from the 'UNIVERSIDAD AUTONOMA DE O...'. The main content area is titled 'Máquinas virtuales' (Virtual Machines) and shows a list of VMs. The 'serverAirbnb' VM is selected, and its details are shown in the right-hand pane. The 'Información esencial' (Essential Information) section includes the resource group 'airbnb\_app', the state 'En ejecución' (Running), the location 'East US 2', the subscription 'Azure for Students', and the VM ID '0efd77ba-0410-4e3c-bb03-c89f42de4fdf'. The 'Sistema operativo' (Operating System) is 'Linux (ubuntu 20.04)', and the 'Tamaño' (Size) is 'Standard DS1 v2 (1 vcpu, 3.5 GiB de memo...)'. The 'Dirección IP pública' (Public IP Address) is '52.254.71.166', and the 'Red virtual/subred' (Virtual Network/Subnet) is 'airbnb\_network/default'. The 'Nombre DNS' (DNS Name) is 'server-airbnb.eastus2.cloudapp.azure.com', which is being copied to the clipboard. The 'Estado de mantenimiento' (Maintenance State) is '-', and the 'Hora de creación' (Creation Time) is '23/5/2024, 10:55 p.m. UTC'.

Microsoft Azure

Buscar recursos, servicios y documentos (G+)

Inicio > Máquinas virtuales >

Máquinas virtuales

Universidad Autónoma de Occidente (uao.edu.co)

+ Crear

Filtrar por cualquier campo...

Nombre ↑

serverAirbnb

workerAirbnb

serverAirbnb

Máquina virtual

Buscar

Conectar Iniciar Reiniciar Detener Hibernar (versión preliminar)

Información esencial

Grupo de recursos (mover) [airbnb\\_app](#)

Estado

En ejecución

Ubicación

East US 2

Suscripción (mover) [Azure for Students](#)

Id. de suscripción

0efd77ba-0410-4e3c-bb03-c89f42de4fdf

Sistema operativo

Linux (ubuntu 20.04)

Tamaño

Standard DS1 v2 (1 vcpu, 3.5 GiB de memo...)

Dirección IP pública

52.254.71.166

Red virtual/subred

[airbnb\\_network/default](#)

Nombre DNS

[server-airbnb.eastus2.cloudapp.azure.com](#)

Copiar al Portapapeles

Estado de mantenimiento

-

Hora de creación

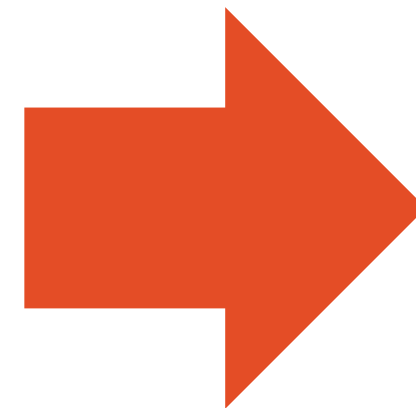
23/5/2024, 10:55 p.m. UTC

Etiquetas (editar) [Agregar etiquetas](#)

Página 1 de 1

# ANALITICA DE DATOS

PySpark




Power BI

**GENERAR REPORTES Y DASHBOARDS INTERACTIVOS.**




```
vagrant@servidorUbuntu: ~/l x + v
vagrant@servidorUbuntu:~/clusterAirbnb/reservasDistritoAirbnb$ rm -r
.part-00000-dd947131-f843-4249-bd4f-a7bfd05c3a18-c000.csv.crc
.part-00001-dd947131-f843-4249-bd4f-a7bfd05c3a18-c000.csv.crc
.part-00002-dd947131-f843-4249-bd4f-a7bfd05c3a18-c000.csv.crc
_SUCCESS
._SUCCESS.crc
vagrant@servidorUbuntu:~/clusterAirbnb/reservasDistritoAirbnb$ rm -r *crc
rm: cannot remove '*crc': No such file or directory
vagrant@servidorUbuntu:~/clusterAirbnb/reservasDistritoAirbnb$ rm -r *csv
rm: cannot remove '*csv': No such file or directory
vagrant@servidorUbuntu:~/clusterAirbnb/reservasDistritoAirbnb$ ls
_SUCCESS
vagrant@servidorUbuntu:~/clusterAirbnb/reservasDistritoAirbnb$ cd ..
vagrant@servidorUbuntu:~/clusterAirbnb$ ls
clientesDestacados reservasDistritoAirbnb reservasTipoHabitacion totalReservasGeneradas
hostsDestacados reservasPorMes totalDinero
vagrant@servidorUbuntu:~/clusterAirbnb$ rm -r clientesDestacados
vagrant@servidorUbuntu:~/clusterAirbnb$ rm -r reservasDistritoAirbnb/
vagrant@servidorUbuntu:~/clusterAirbnb$ rm -r reservasTipoHabitacion/
vagrant@servidorUbuntu:~/clusterAirbnb$ rm -r totalReservasGeneradas/
vagrant@servidorUbuntu:~/clusterAirbnb$ rm -r hostsDestacados/
vagrant@servidorUbuntu:~/clusterAirbnb$ rm -r reservasPorMes/
vagrant@servidorUbuntu:~/clusterAirbnb$ rm -r totalDinero/
vagrant@servidorUbuntu:~/clusterAirbnb$ ls
vagrant@servidorUbuntu:~/clusterAirbnb$ cd
vagrant@servidorUbuntu:~$ cd labSpark/spark-3.5.1-bin-hadoop3/sbin/
vagrant@servidorUbuntu:~/labSpark/spark-3.5.1-bin-hadoop3/sbin$ ./start-master.sh
starting org.apache.spark.deploy.master.Master, logging to /home/vagrant/labSpark/spark-3.5.1-bin-hadoop3/logs/spark-vagrant-org.apache.spark.deploy.master.
Master-1-servidorUbuntu.out
vagrant@servidorUbuntu:~/labSpark/spark-3.5.1-bin-hadoop3/sbin$ ./start-master.sh
org.apache.spark.deploy.master.Master running as process 1616. Stop it first.
vagrant@servidorUbuntu:~/labSpark/spark-3.5.1-bin-hadoop3/sbin$ ./start-worker.sh
```

OBS 30.1.2 - Perfil: Sin Tf...



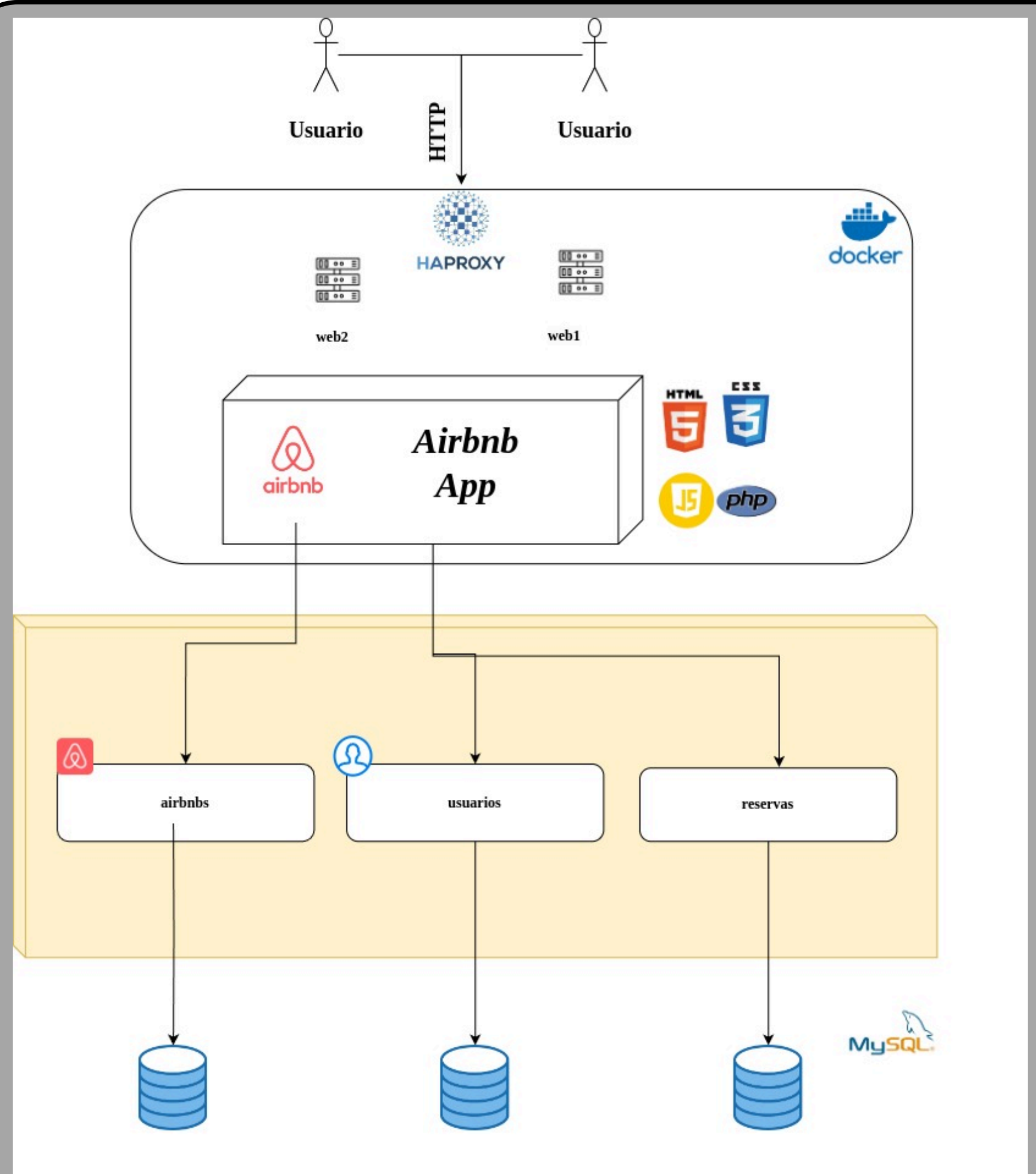
21°C  
Parc. soleado

Búsqueda



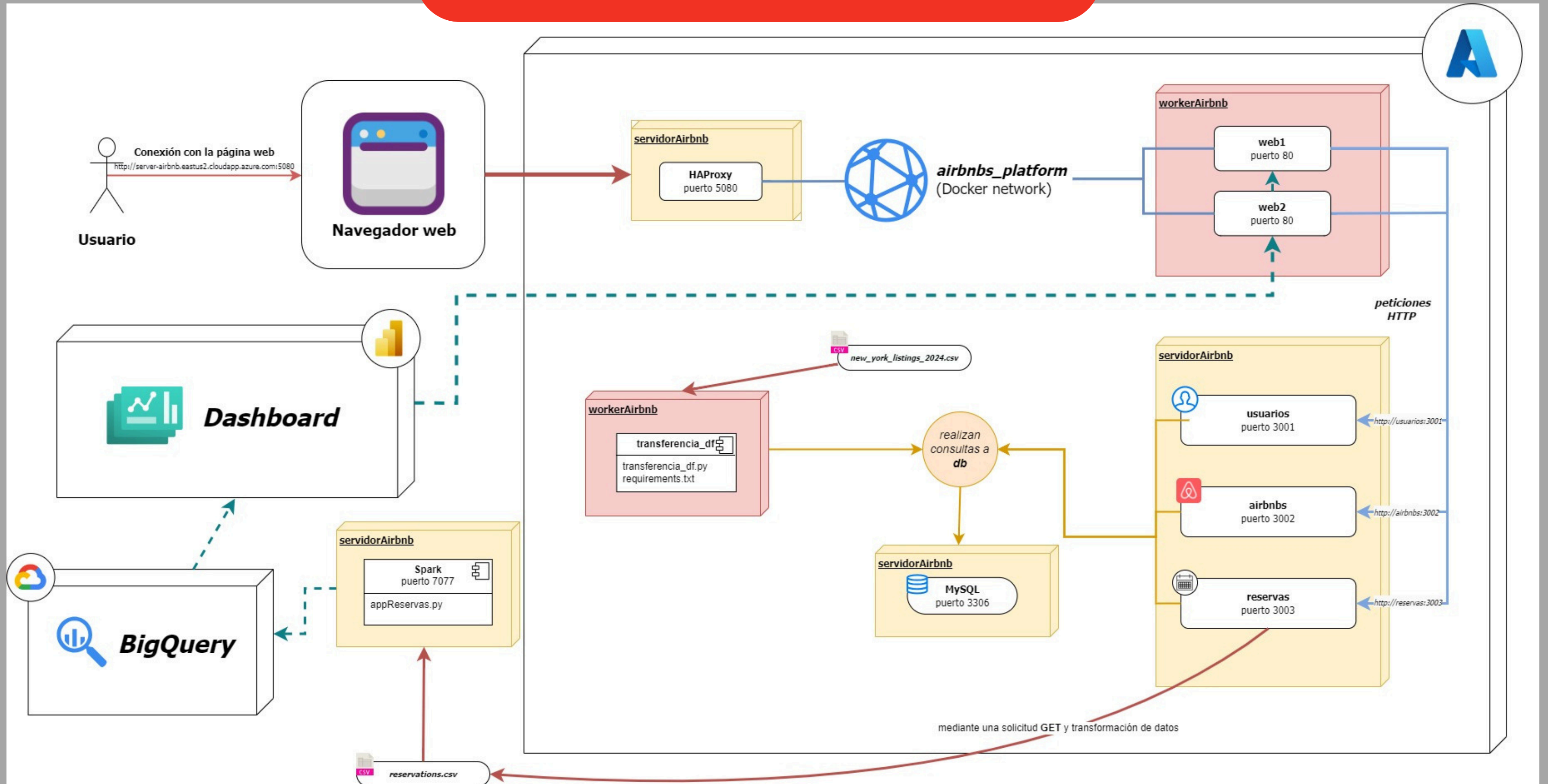
ESP LAA 7:30 a. m. 24/05/2024

# DIAGRAMA DE COMPONENTES



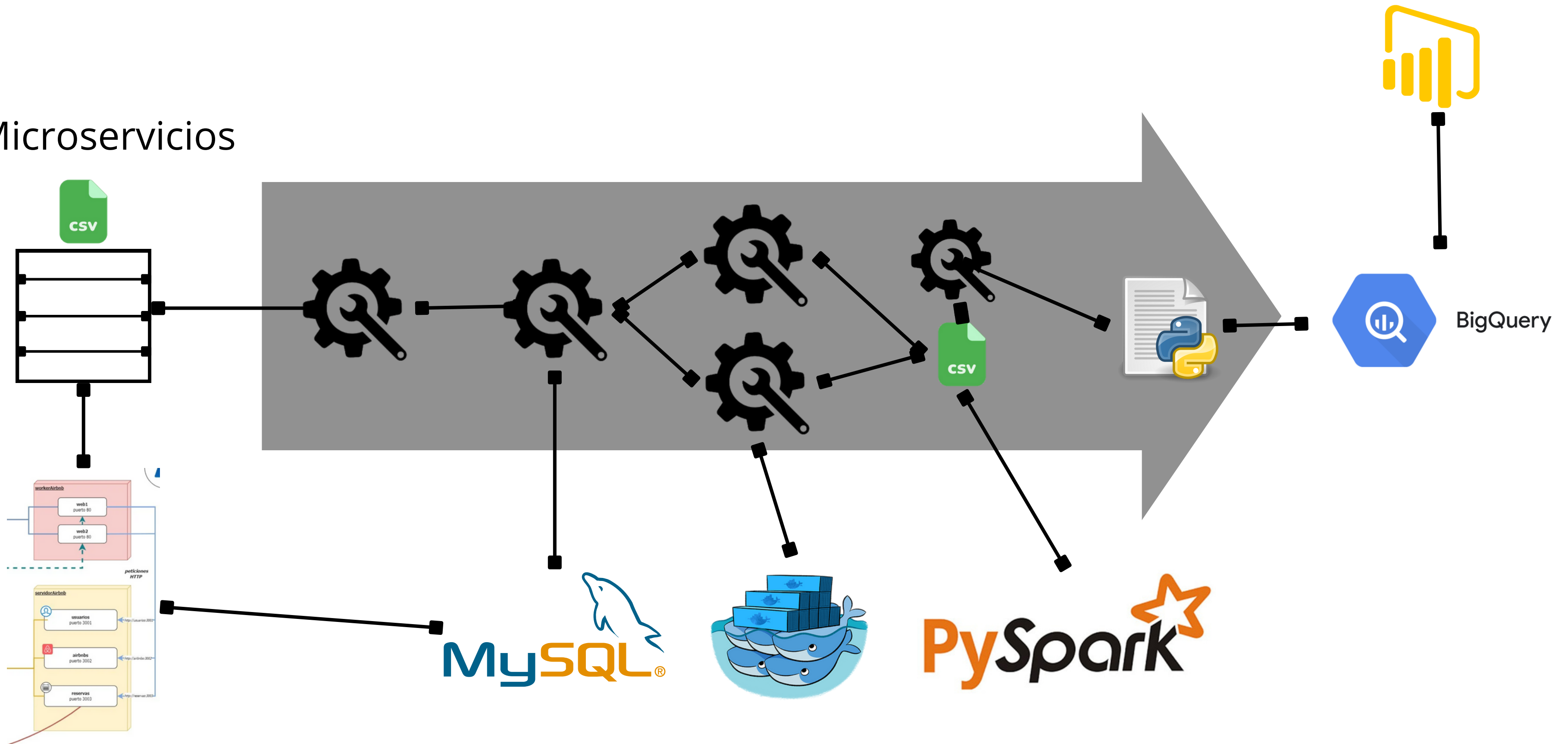


# DIAGRAMA DE DESPLIEGUE



# PIPELINE

Microservicios



# PRUEBAS DE ESCALABILIDAD

HAProxy version 2.3.21-3ce4ee0, released 2022/07/27

## Statistics Report for pid 7

### > General process information

pid = 7 (process #1, nbproc = 1, nbthread = 3)  
uptime = 0d 0h06m03s  
system limits: memmax = unlimited; ulimit-n = 1048576  
maxsock = 1048576; maxconn = 524271; maxpipes = 0  
current conns = 2; current pipes = 0/0; conn rate = 2/sec; bit rate = 0.000 kbps  
Running tasks: 0/2053; idle = 100 %

active UP  
active UP, going down  
active DOWN, going up  
active or backup DOWN  
active or backup DOWN for maintenance (MAINT)  
active or backup SOFT STOPPED for maintenance  
Note: "NOLB"/"DRAIN" = UP with load-balancing disabled.

Display option:  
Scope :  
Hide DOWN servers  
Refresh now  
CSV export  
JSON export (schema)

External resources:  
Primary site  
Updates (v2.3)  
Online manual

web-backend																															
	Queue			Session rate			Sessions					Bytes		Denied		Errors		Warnings		Server											
	Cur	Max	Limit	Cur	Max	Limit	Cur	Max	Limit	Total	LbTot	Last	In	Out	Req	Resp	Req	Conn	Resp	Retr	Redis	Status	LastChk	Wght	Act	Bck	Chk	Dwn	Dwntme	Thrtle	
web1	0	0	-	0	823	0	2 044	-	24 915	22 559	1m23s	2 504 349	35 461 451			0			72	674	2356	0	2m53s UP	L4OK in 0ms	1/1	Y	-	14	3	22s	-
web2	0	0	-	0	827	0	2 287	-	25 696	23 141	1m23s	2 568 651	36 147 938			0			68	885	2555	0	2m54s UP	L4OK in 0ms	1/1	Y	-	13	2	22s	-
Backend	0	0		1	1 647	1	4 247		49 600	45 700	0s	5 507 052	72 498 261	0	0				4 036	1 559	4911	0	2m54s UP		2/2	2	0		2	13s	

http																															
	Queue			Session rate			Sessions						Bytes		Denied		Errors		Warnings		Server										
	Cur	Max	Limit	Cur	Max	Limit	Cur	Max	Limit	Total	LbTot	Last	In	Out	Req	Resp	Req	Conn	Resp	Retr	Redis	Status	LastChk	Wght	Act	Bck	Chk	Dwn	Dwntme	Thrtle	
Frontend				2	1 060	-	2	4 962	524 271	10 707			5 507 052	72 510 195	0	0	54					OPEN									

FUNCIONA EL  
BALANCEO DE  
CARGA

Reporte resumen.jmx (C:\Users\juann\Downloads\apache-jmeter-5.6.3\apache-jmeter-5.6.3\bin\Reporte resumen.jmx) - Apache JMeter (5.6.3)

Archivo Editar Search Lanzar Opciones Tools Ayuda

Plan de Pruebas

Grupo de Hilos

Petición HTTP

Reporte resumen

Reporte resumen

Nombre: Reporte resumen

Comentarios

Escribir todos los datos a Archivo

Nombre de archivo

Navegar...

Log/Mostrar sólo: ☐ Escribir en Log ☐ Sólo Errores ☐ Éxitos

Configurar

Etiqueta	# Muestras	Media	Mín	Máx	Desv. Estándar	% Error	Rendimiento	Kb/sec	Sent KB/sec	Media de Bytes
Petición HTTP	50000	6178	1	82816	12581,06	12,15%	286,4/sec	427,99	34,07	1530,2
Total	50000	6178	1	82816	12581,06	12,15%	286,4/sec	427,99	34,07	1530,2

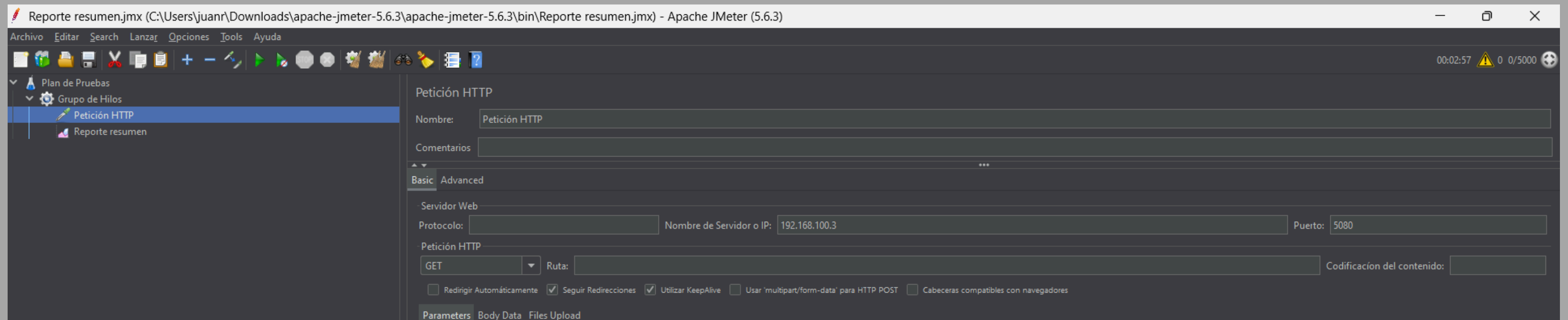
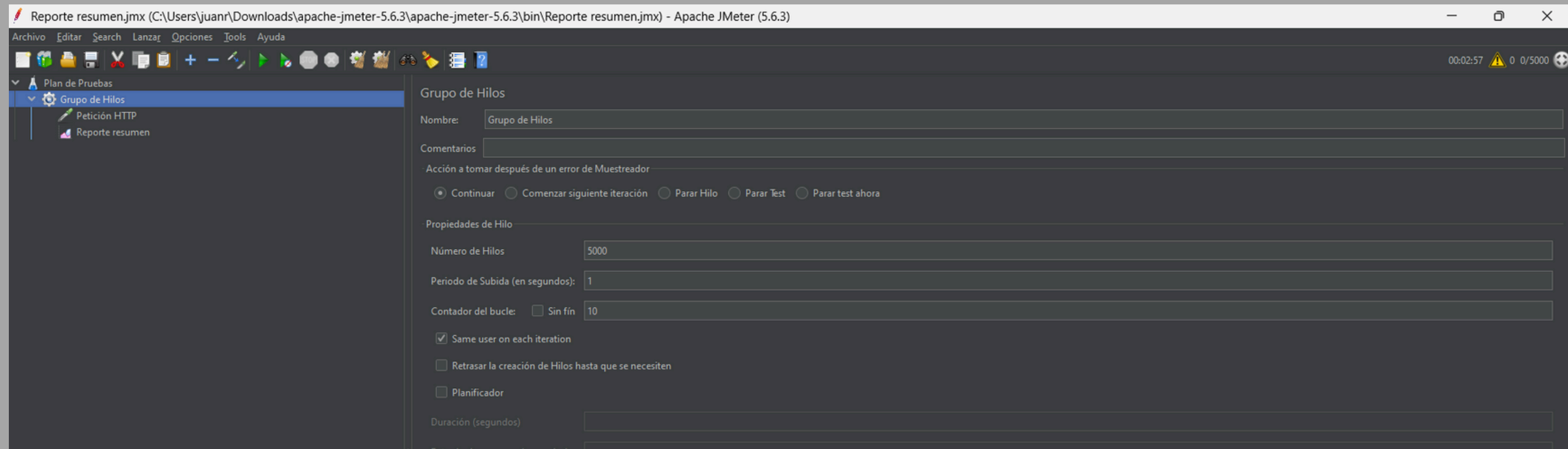
## PRUEBAS DE ESCALABILIDAD

**EN COMPOSE CON HAPROXY Y SE REPITE LA PRUEBA EN SWARM**

[illegible]



# PRUEBAS DE ESCALABILIDAD



# PRUEBAS DE ESCALABILIDAD

*Tablas de resultados – Pruebas de carga normal*

Compon ente	Répli cas	Path (GET - HTTP )	Err or %	Tiempo de respuesta promedio (THROUG HPUT)
HAProx y	1	N/A	0.0 0	10.1
Micro Usuarios	1	/usuar ios	0.0 0	3.6

*Tablas de resultados – Pruebas de carga alta*

Compon ente	Répli cas	Path (GET - HTTP )	Err or %	Tiempo de respuesta promedio (THROUG HPUT)
HAProx y	1	N/A	0.0 0	4.8
Micro Usuarios	1	/usuar ios	0.0 0	18.0

*Tablas de resultados – Pruebas de carga estrés*

Compon ente	Répli cas	Path (GET - HTTP )	Err or %	Tiempo de respuesta promedio (THROUG HPUT)
HAProx y	1	N/A	0.0 0	332.8
Micro Usuarios	1	/usuar ios	0.0 0	85.7

# CONCLUSIONES

THANKYOU