

Nombre:	Matricula:
Donovan Gared Orozco García	3005606
Nombre del curso:	Nombre del profesor:
Infraestructura de Big Data	Miguel de Jesús Martinez Felipe
Módulo: 1	Actividad: Avance 2 Evidencia 1
Fecha: 14 Mar 24.	

¿Qué es Apache Ambari?

Apache Ambari es un proyecto de código abierto el cual facilita la gestión de clusteres de Hadoop, también ofrece una interfaz web demasiada intuitiva la cual esta hecha para.

- **Administrar y Aprovisionar el Cluster de Hadoop:** Esto permite instalar, iniciar, detener y actualizar los servicios de Hadoop.
- **Monitorear el Cluster:** Se ofrece una vista completa del estado del cluster en donde se incluye el uso de los recursos, la salud de los servicios y sus alertas.
- **Configurar los Servicios:** Permite personalizar la configuración de los servicios que ofrece Hadoop.
- **Aplicar Seguridad:** Permite configurar la autenticación, autorización y el cifrado de los datos.

¿Cuales son sus Beneficios?

- **Simplifica la Gestión de Clusteres en Hadoop:** Reduce la complejidad ademas de tambien el tiempo necesario para poder administrar clusteres de Hadoop.
- **Mejora la Disponibilidad y la Confiabilidad del Cluster:** Ayuda a prevenir y solucionar problemas de cluster.
- **Aumenta la Eficiencia:** Permite a los administradores de clusteres centrarse en tareas con mayor importancia.
- **Reduce Costos:** Puede ayudar a reducir los costos de las operaciones del cluster.

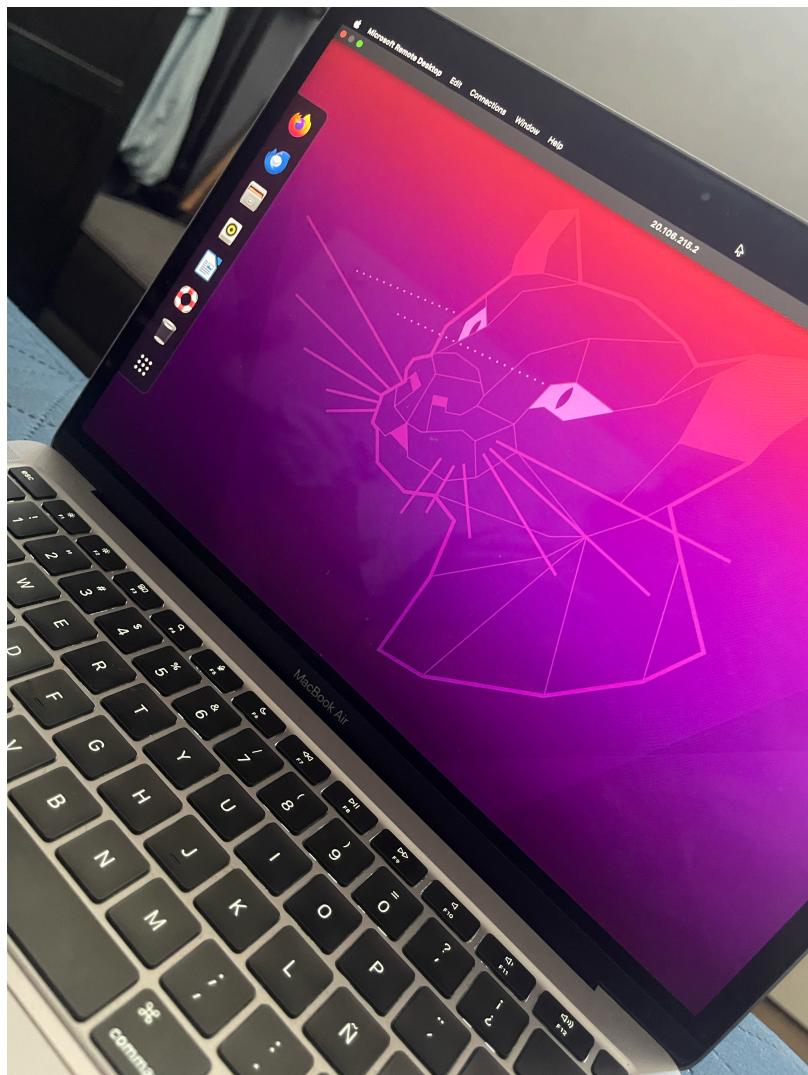
Sus Componentes.

- **Interfaz de Usuario Web:** Proporciona una interfaz gráfica para poder administrar el cluster.
- **API REST:** Permite a las herramientas de terceros para así poder integrarse con Ambari.
- **Ambari Server:** Se ejecuta en el nodo de Ambari y administra el cluster.
- **Agentes:** Dentro de cada nodo del cluster y recopilan datos de estado.

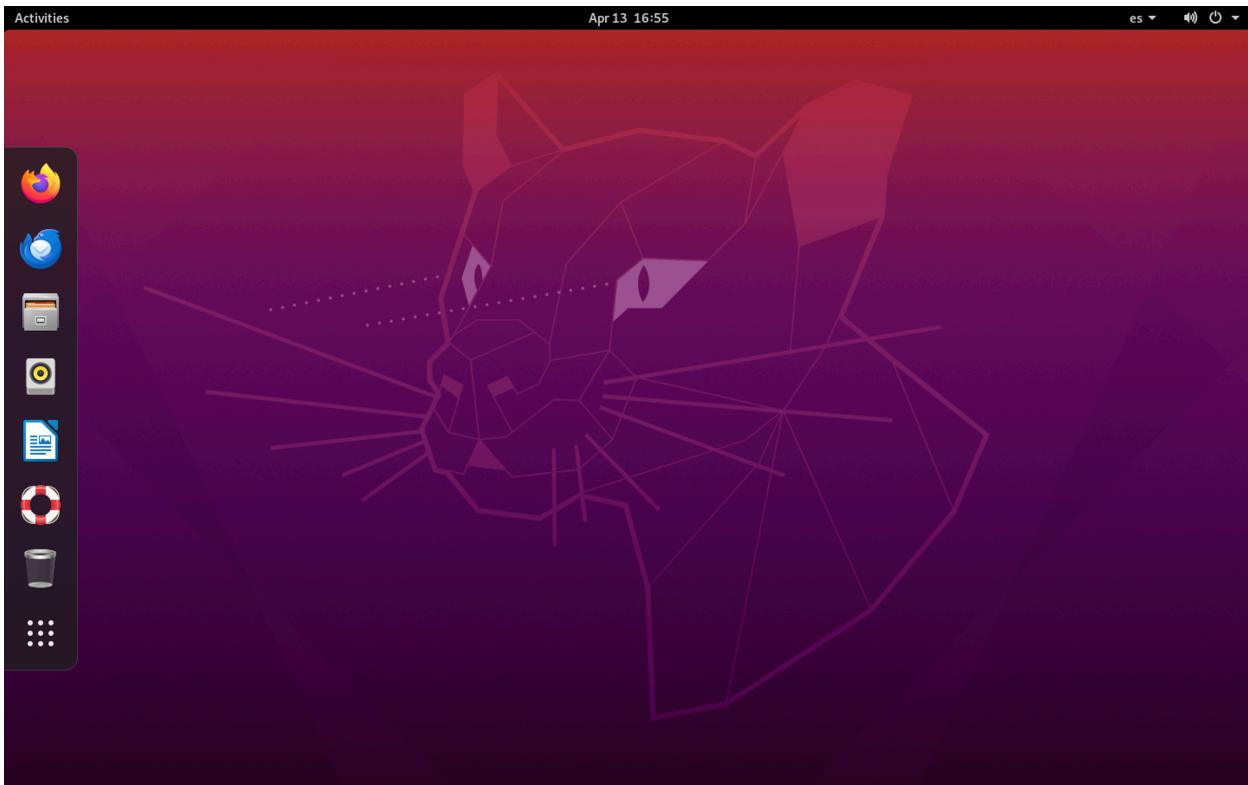
¿Para Que Se Usa?

- **Análisis de BigData:** Se usa para analizar un gran conjunto de datos en Hadoop.
- **Aprendizaje Automático:** Se usa para entrenar modelos de aprendizaje automático en Hadoop.
- **Ciencia de Datos:** Se usa para realizar análisis de ciencia de datos en Hadoop.
- **Almacenamiento de Datos:** Se usa para almacenar grandes cantidades de datos dentro de Hadoop.

Captura de Pantalla de Maquina Virtual en Azure.



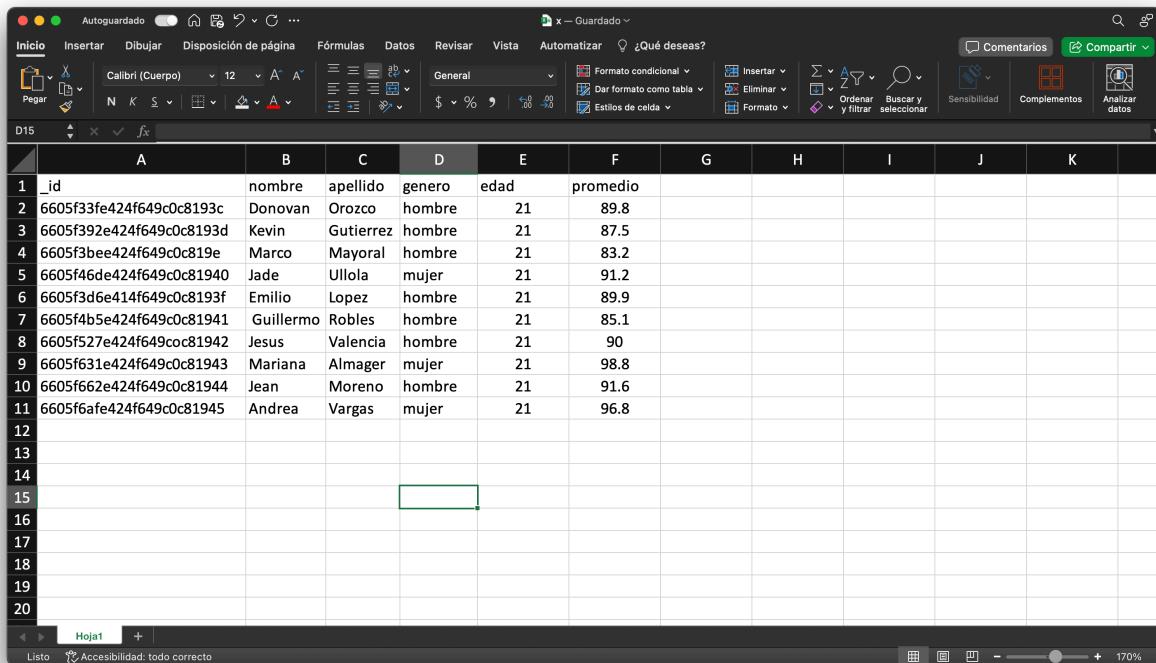
Captura de Pantalla con ejecución Ubuntu.



Captura de Pantalla Hadoop.

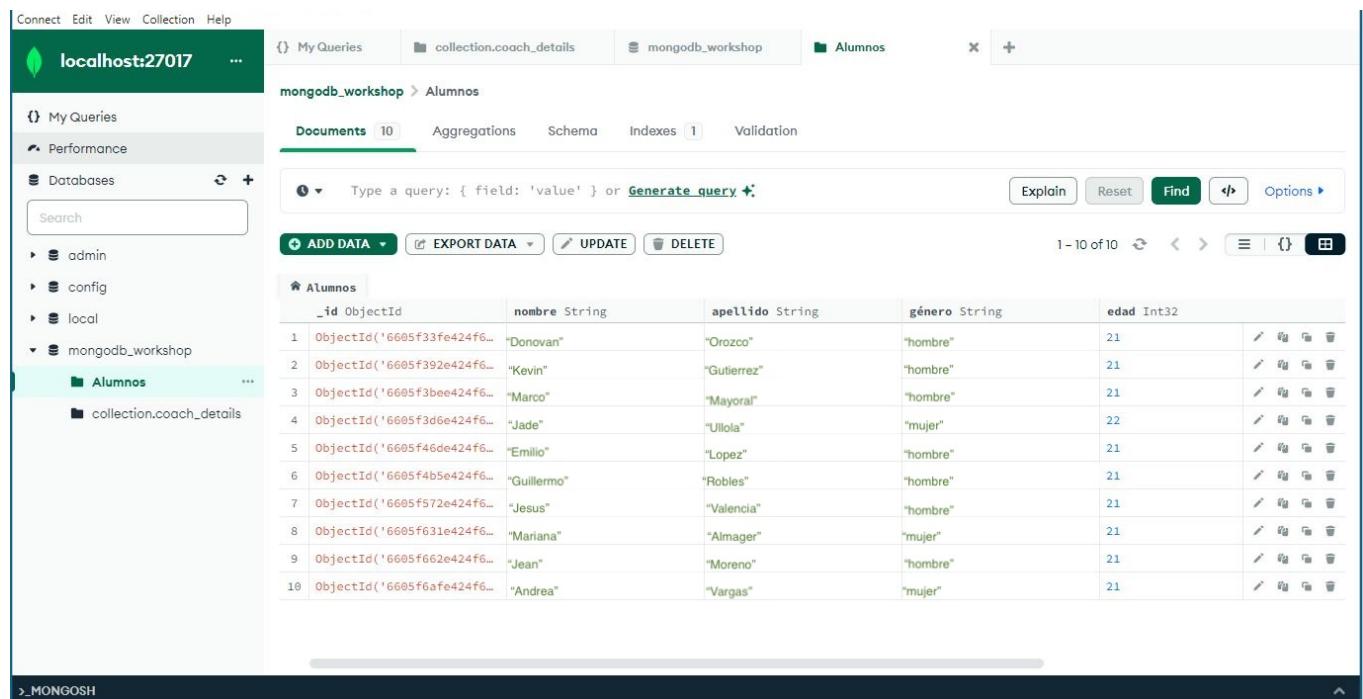
A screenshot of a Firefox browser window showing the Hadoop cluster management interface. The URL in the address bar is "localhost:8088/cluster". The page displays cluster metrics, including "Apps Submitted: 0", "Apps Pending: 0", "Apps Running: 0", "Apps Completed: 0", and "Containers Running: 0". It also shows "Active Nodes: 0" and "Decommissioning Nodes: 0". Under "Scheduler Metrics", it lists "Scheduler Type: Capacity Scheduler" and "Scheduling Resource Type: [memory-mb (unit=Mi), vcores] <memory:1024, vCore>". A table at the bottom shows application details with columns: ID, User, Name, Application Type, Application Tags, Queue, Application Priority, StartTime, LaunchTime, and FinishTime. A tooltip indicates "0 items selected (containing 0 items)". The browser's toolbar includes options like "Detach tab", "Fit to window size", "Fullscreen", "Set connection password", "Toggle scaling", "Hide bar", and "Reconnect". The status bar at the bottom shows the date and time as "Feb 6 23:21".

Captura Datos MongoDB.



A screenshot of Microsoft Excel showing a table of student data. The table has columns labeled A through K. Column A contains student IDs, column B contains names, column C contains surnames, column D contains gender, column E contains age, column F contains average, and column G contains other data. The data is identical to the one shown in the MongoDB interface below.

	A	B	C	D	E	F	G	H	I	J	K
1	_id	nombre	apellido	genero	edad	promedio					
2	6605f33fe424f649c0c8193c	Donovan	Orozco	hombre	21	89.8					
3	6605f392e424f649c0c8193d	Kevin	Gutierrez	hombre	21	87.5					
4	6605f3bee424f649c0c8193e	Marco	Mayoral	hombre	21	83.2					
5	6605f46de424f649c0c81940	Jade	Ullola	mujer	21	91.2					
6	6605f3d6e414f649c0c8193f	Emilio	Lopez	hombre	21	89.9					
7	6605f4b5e424f649c0c81941	Guillermo	Robles	hombre	21	85.1					
8	6605f527e424f649cc81942	Jesus	Valencia	hombre	21	90					
9	6605f631e424f649c0c81943	Mariana	Almager	mujer	21	98.8					
10	6605f662e424f649c0c81944	Jean	Moreno	hombre	21	91.6					
11	6605f6afe424f649c0c81945	Andrea	Vargas	mujer	21	96.8					
12											
13											
14											
15											
16											
17											
18											
19											
20											



A screenshot of the MongoDB Compass interface. On the left, there's a sidebar with databases like admin, config, local, and mongodb_workshop, and collections like Alumnos and collection.coach_details. The main area shows the 'Alumnos' collection with 10 documents. Each document has fields: _id, nombre, apellido, genero, and edad. The data is identical to the one shown in the Excel table above.

	_id	nombre	apellido	genero	edad
1	ObjectId('6605f33fe424f649c0c8193c')	Donovan	Orozco	hombre	21
2	ObjectId('6605f392e424f649c0c8193d')	Kevin	Gutierrez	hombre	21
3	ObjectId('6605f3bee424f649c0c8193e')	Marco	Mayoral	hombre	21
4	ObjectId('6605f3d6e414f649c0c8193f')	Jade	Ullola	mujer	22
5	ObjectId('6605f46de424f649c0c81940')	Emilio	Lopez	hombre	21
6	ObjectId('6605f4b5e424f649c0c81941')	Guillermo	Robles	hombre	21
7	ObjectId('6605f527e424f649cc81942')	Jesus	Valencia	hombre	21
8	ObjectId('6605f631e424f649c0c81943')	Mariana	Almager	mujer	21
9	ObjectId('6605f662e424f649c0c81944')	Jean	Moreno	hombre	21
10	ObjectId('6605f6afe424f649c0c81945')	Andrea	Vargas	mujer	21

Referencias.

Ambari User Guides - Apache Ambari - Apache Software Foundation.
(s. f.). <https://cwiki.apache.org/confluence/display/AMBARI/Ambari+User+Guides>

Ambari -. (s. f.-b). <https://ambari.apache.org/>

Apache Ambari | Cloudera. (2023, 18 julio). Cloudera. <https://es.cloudera.com/products/open-source/apache-hadoop/apache-ambari.html>

Cluster Hadoop con apache Ambari. (s. f.). <https://ccia.esei.uvigo.es/docencia/CDA/1516/practicas/ejercicio-hadoop/>