Visualization architecture

One of the main goals of the project is to visualize geographical data about the stops of vehicles over Valle de Aburrá map to understand the current demand for the public transport system in this area.

Since this transport system does not have a fixed set of stops, it is necessary to associate any given observation to the closest nodes in the map, and that way it can be calculated a load of passengers in the nodes. Once this data is processed, the results are going to be visualized in a dashboard that allows the user to observe the demand at any moment over a period during the months processed.

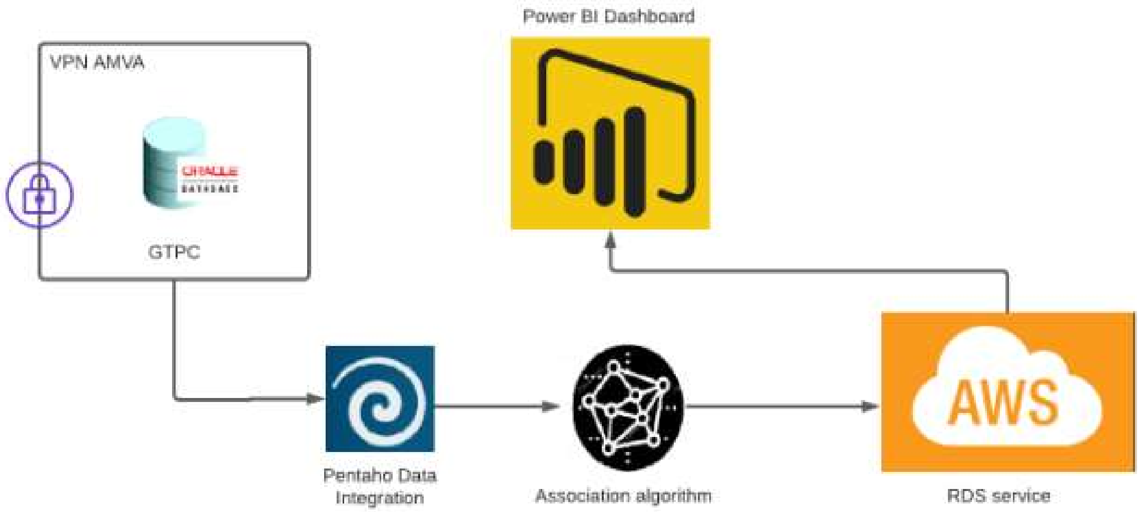

Fig.1 Visualization architecture

Database Infrastructure

In this section, it is described the databases available for the analysis that has been carried out throughout the project. After sourcing the data from the entity (Oracle database), these were consolidated in a PostgreSQL instance created in AWS (RDS service).

The data delivered is in the main schema called **source** and additionally, in another schema called **dw** there is a calendar-type table.

Relational Model

In the model, it is found information corresponding to the description of the event of interest and characteristics related to the vehicle in which they occur, as well as the spatial information of the routes carried out by each vehicle.

This is, in the first instance, both the boarding and the alighting at each stop the vehicle makes on its corresponding route (which is the table that we have been working on so far) and the characteristics associated with the vehicle that performs the routes: its associated model, the company to which it belongs and/or operates it, its capacity to transport standing, seated and total passengers.

On the other hand, there is also the data transformed into a row-column format (where each row represents an observation associated with a variable or column) corresponding to the stops of the routes in the metropolitan area (nodes) and the sections that connect them (links) which are the geometries to which we are interested in assigning the transportation demand at a specific time interval.

The last table is a calendar that contains a series of dates for indexing and/ or filtering the events according to different time intervals required by the stakeholders, that is, by hour, day, and so on.

ds4a-database.cnjtnqqpofwy.us-east-2.rds.amazonaws.com ds4a_project dw Tables d_date

d_date

Columns - 29

Name	Type	Length	Precision	Scale	Default value	Allow null	Unique
date_dim_id	integer		32	0		false	false
date_actual	date					false	false
epoch	bigint		64	0		false	false
day_suffix	character varying	4				false	false
day_name	character varying	9				false	false
day_of_week	integer		32	0		false	false
day_of_month	integer		32	0		false	false
day_of_quarter	integer		32	0		false	false
day_of_year	integer		32	0		false	false
week_of_month	integer		32	0		false	false
week_of_year	integer		32	0		false	false
week_of_year_iso	character	10				false	false
month_actual	integer		32	0		false	false
month_name	character varying	9				false	false
month_name_abbreviated	character	3				false	false
quarter_actual	integer		32	0		false	false
quarter_name	character varying	9				false	false
year_actual	integer		32	0		false	false
first_day_of_week	date					false	false
last_day_of_week	date					false	false

first_day_of_month	date					false	false
last_day_of_month	date					false	false
first_day_of_quarter	date					false	false
last_day_of_quarter	date					false	false
first_day_of_year	date					false	false
last_day_of_year	date					false	false
mmyyyy	character	6				false	false
mmddyyyy	character	10				false	false
weekend_indr	boolean					false	false

ds4a-database.cnjtnqqpofwy.us-east-2.rds.amazonaws.com ds4a_project source Tables empresa

empresa

Columns - 4

Name	Type	Length	Precision	Scale	Default value	Allow null	Unique
idempresa	bigint		64	0		false	false
nombres	character varying	32				true	false
nroidentificacion	character varying	15				true	false
identidadopera	bigint		64	0		false	false

ds4a-database.cnjtnqqpofwy.us-east-2.rds.amazonaws.com ds4a_project source Tables links

links

Columns - 15

Name	Type	Length	Precision	Scale	Default value	Allow null	Unique
id	character varying	50				true	false
inode	bigint		64	0		true	false
jnode	bigint		64	0		true	false
TYPE	bigint		64	0		true	false
modes	character varying	50				true	false
lanes	numeric		19	0		true	false
vdf	bigint		64	0		true	false
LENGTH	numeric		16	7		true	false
data1	numeric		11	3		true	false
data2	numeric		18	2		true	false
data3	numeric		16	7		true	false
volax	numeric		16	7		true	false
volau	numeric		16	7		true	false
volad	numeric		16	7		true	false
timau	numeric		16	7		true	false

ds4a-database.cnjtnqqpofwy.us-east-2.rds.amazonaws.com ds4a_project source Tables nodes

nodes

Columns - 8

Name	Type	Length	Precision	Scale	Default value	Allow null	Unique
id	bigint		64	0		false	false
x	numeric		25	9		true	false
y	numeric		27	10		true	false
iszone	bigint		64	0		true	false
isintersec	bigint		64	0		true	false
label	bigint		64	0		true	false
lat	numeric		33	15		true	false
long	numeric		30	13		true	false

ds4a-database.cnjtnqqpofwy.us-east-2.rds.amazonaws.com ds4a_project source Tables passenger_route_vehicle

passenger_route_vehicle

Columns - 12

Name	Type	Length	Precision	Scale	Default value	Allow null	Unique
secuenciarecorrido	bigint		64	0		false	false
recorridofinalizado	character varying	1				true	false
idvehiculo	bigint		64	0		false	false
codigoruta	bigint		64	0		false	false
fecharegistro	timestamp without time zone					false	false
latitud	numeric		14	6		true	false
longitud	numeric		16	6		true	false
subendelantera	bigint		64	0		true	false
subentrasera	bigint		64	0		true	false
bajandelantera	bigint		64	0		true	false
bajantrasera	bigint		64	0		true	false
id	bigint		64	0	nextval('source.passenger_route_vehicle_id_seq'::regclass)	false	false

ds4a-database.cnjtnqqpofwy.us-east-2.rds.amazonaws.com ds4a_project source Tables vehiculo

vehiculo

Columns - 9

Name	Type	Length	Precision	Scale	Default value	Allow null	Unique
idvehiculo	bigint		64	0		false	false
placavehiculo	character varying	50				true	false
modelo	bigint		64	0		true	false
cappasajeros	bigint		64	0		true	false
cappasajerospie	bigint		64	0		true	false
cappasajerossentados	bigint		64	0		true	false
idempresa	bigint		64	0		true	false
nombres	character varying	100				true	false
identidadopera	bigint		64	0		true	false

Fig2. Database schema

ER Diagram

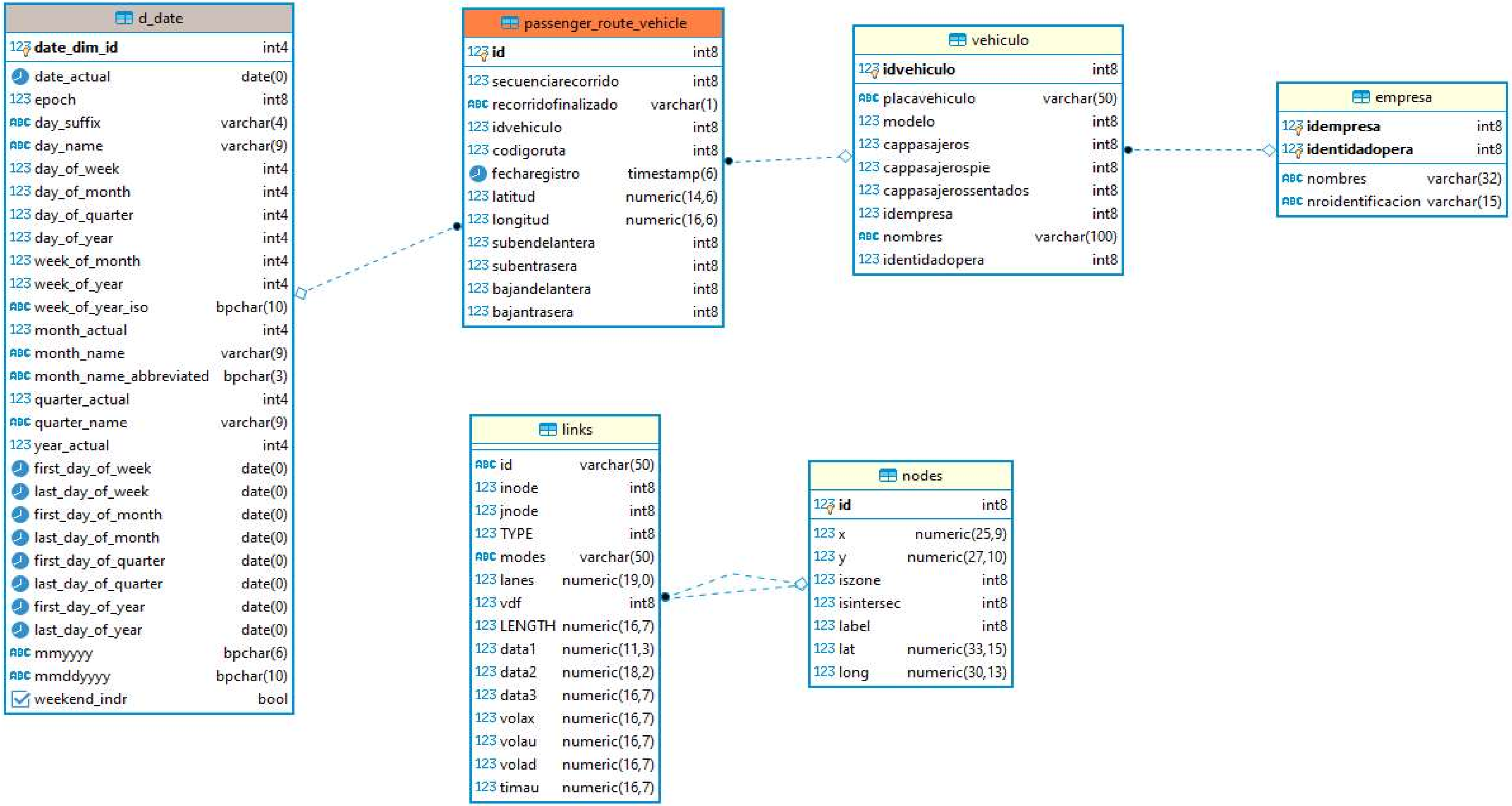

Fig3. ER Model