Big Data y Visualización MapReduce

Universidad Autónoma de Manizales

Docente: Alejandro Mora Rubio





Contenido

MapReduce

MapReduce sobre datos estructurados

MapReduce sobre texto plano



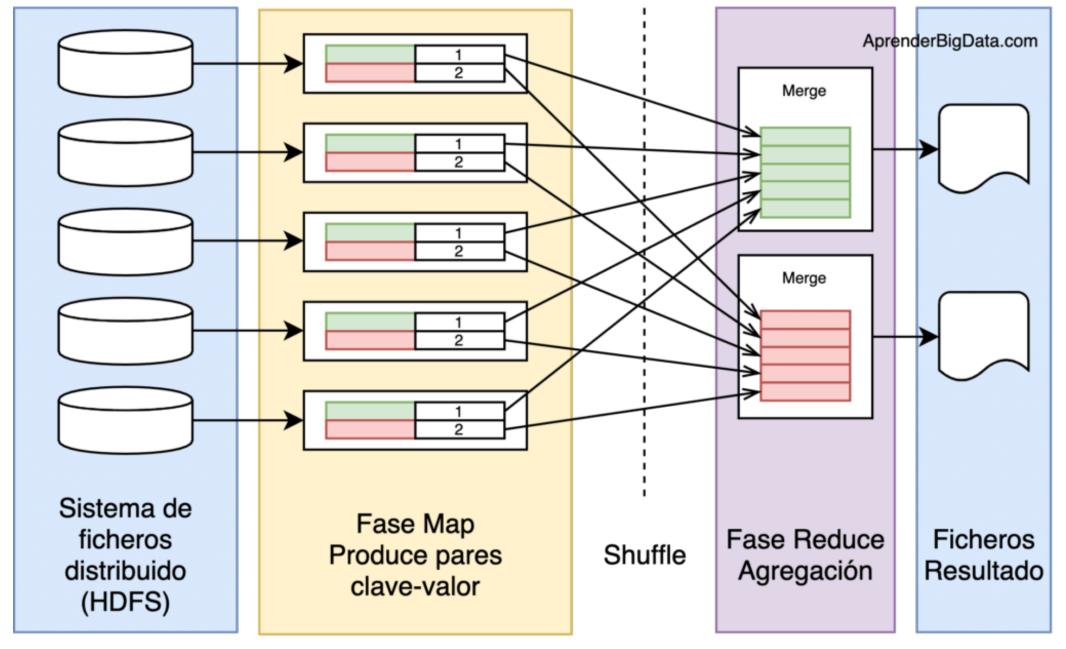
MapReduce

Modelo de programación distribuida basada en dos tareas:

Map operaciones para extraer información en el formato llave ightarrow valor

Reduce agrega los resultados para cada una de las llaves generando un

resultado final





Sobre datos estructurados

Ejemplo: conocer el importe total por ciudad

				ap	(BOG, 88.56)
Fecha	Hora	Ciudad	Produc	orte	(BGA, 62,41) Medio de pago
2012-01-01	9:08	BOG	Libros	88.56	(MED, 93.3/) Discover (MED 205.02)
2012-01-01	9:52	BGA	Libros	62.41	Discover
2012-01-01	10:01	MED	Musica -	93.37	Sort
2012-01-01	B06,	88.56) Redu	C6 _{95.93}	(BYST-C88.56)
	(BGA	, 62.41)		(BGA, 62.41)	
	MED,	488.67	7)	(MED, [93.37, 395.93])	



Sobre datos estructurados

Ejemplo: contar los contactos/amigos de cada usuario

```
["juan", "pepe"]
                                          (juan, 1)
                             Map
["juan", "sebastian"]
                                           (juan, 1)
                                           (raul, 1)
   ["raul", "pepe"]
   ["ana", "pepe"]
                                           (ana, 1)
                         Reduce
           (juan, 2)
(raul, 1)
                                        (juan, [1, 1])
                                           (ana, 1)
            (ana, 1)
```



Sobre texto plano

The Overall MapReduce Word Count Process

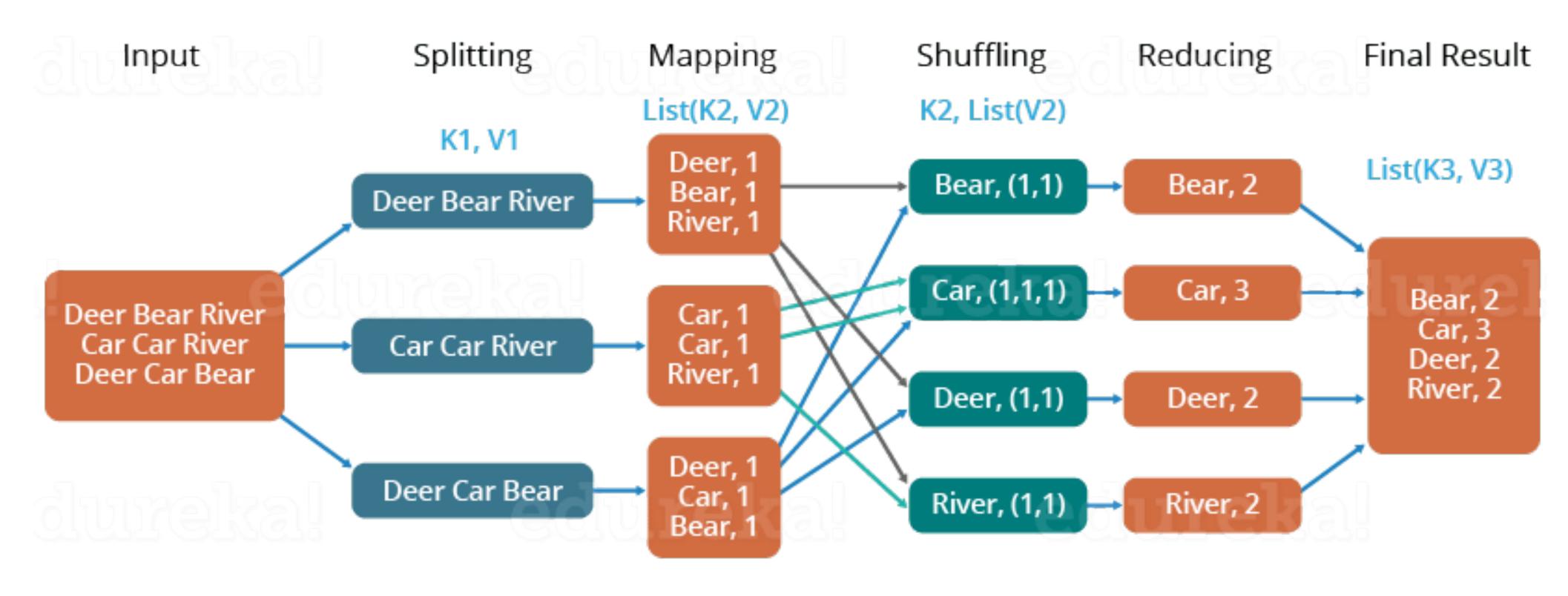


Imagen de <u>Edureka</u>



edureka!