

Big data analytics

Big Data Analytics (University of Mumbai)



Scan to open on Studocu

Relational Algebra Operations using Mapreduce

> Some of the relational algebra operations are:

- 1) selection & Dibberence
- 2) Projection (6) Natural join
- 3) Intersection (Grouping & Aggregation.
- 4) Union

Selection using mapreduce

Selection Algo

map (Key, Value):

bor tuble in value:

if tuple satisfies condition:

emit (tuple, tuple)

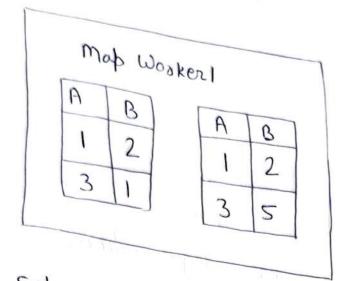
ruduce (key, values):

emit (key, key)

and on found building in

and as a disent a many property of the same of the

Example:



A	В	A	B
2	3	1	1
4	5	2	1

Selection Condition: B <= 2

Step1: Create (key, value) pair of each and every

Step 2: After (key, value) conversion a particular hash bunction will be applied onto map workers. So after applying this hash bunction the map worken task tusk is divided into two.

Downloaded by Adwait Purao (adwait.purao@spit.ac.in)

(2)

A hash function will be applied.

	mwi		mw2	
key —	Value	Key Value	key value	Key value
(1,2)	(1,2),(1,2)	(3,1)	(1,1) (1,1)	(2,1) (2,1)
		e		
			Swap bing	

Swapping will be done to discard redundancy that can be caused by duplicate tuples.

Step 3: Swap table 2 and table 3.

	RU	01				2 W 2		
(key 1	value (1,2)(1,2)	key (1,1)	Value (1,1)	(3,1)	رملىك (3,۱)	(2,1)	(2,1)
		1		2000	. h	au to	club.	

Step 4. Now at RW1 2 RW2 we have to check to there is any duplicate when combine we have to check to there values.

RW1 RW2

Key	1 value	Key	value
(1,2)	(1,2)(1,2)	(3,1)	(3,1)
	(1,1)	(2,1)	(2,1)
(1,1)	This document is available free of charge on	C etu	docu

Downloaded by Adwait Purao (adwait.purao@spit.ac.in)

	RW_1
A	В
١	2
1	1

R	ω_2	
A	B	
3	1	
2	1	

Projection using map reduce:

Projection Algorithm

let's Say S be subset containing the selected attributer.

map (key, value);
box tuple in value;
ts = tuple with only attributes in s
emit (ts,ts)

emit (Key, values):

Example

Project: (B,C)

1	B	C	A	B	C
	2	3	2	2	3
3	١	4	4	3	4

1	ß	C	A
	7	2	3
ı		1	4

Stepi

Key	Value
(2,3)	(2,3),(2,3
(1,4)	(1,4)
(3,4)	(3,4)

mW2.

Key	Value
(7,2)	(7,2)
(1,1)	(1,1)
(2,5)	(2,5)
(0,1)	(1,0)

Step 2

Key	value	Key	value
(2,3)	(2,3),(2,3)	(3,4)	(3,4)
(1,4)	(1,4)		

mw2

Ţ.	100 2		
Key	Value	Key	Value
(7,2)	(7,2)	(2,5)	(2,5)
(1,1)		(1,0)	(i,0)

This document is available free of charge on

5 studocu

Step 3

RWI

RW2

	Value	Key)	value
(2,3)	(2,3),(2,3)		(7,2)
(1,4)	(1,4)	(1,1)	(1,1)

Key	value	Key	Value
(3,4)	(3,4)	(2,5)	(2,5)
		(1,0)	(1,0)

Step 4

RW1

Key	Value
(2,3)	(2,3)(2,3)
(1,4)	(1,4)
(7,2)	(7,2)
1,1)	(1,1)

 $R\omega_2$

Key	value
(3,4)	(3,4)
(2,5)	(2,5)
(1,0)	(1,0)

Step 5

RWI

B	C
2	3
1	4
7	2
1	1

RW2

B	(
3	4
2	5
1	0

Union using met reduce:

Union algo:

map (key, value):

box tuple in value:

emit (tuple, tuple)

reduce (key, values).

emit (key, key)

Example

Table, Tablez ABAB	Mab	worker 1
	W* A	
		A B

Tabl	el	Table 2
A	B	AB
2	3	1 1
14	5	2 1

step1	m w1
Key	value
(1,2)	(1,2),(1,2)
(3,1)	(3,1)
(2,1)	(2,1)
	This document is available free of charge on

mw2	
Key	value
(2,3)	(2,3)
(4,5)	(4,5)
(1,1)	(1,1)
Stude	cu ^{2,1)}

Downloaded by Adwait Purao (adwait.purao@spit.ac.in)

	Key	1
(1	,2)	+
(:	3,1)	

mw1

m	(1)2
	~ ~

~ 11	
341	
7	
() ()	ĺ

	Value	Key	Value	Key	Value	Key	Value
(1,2)	(1,2),(1,2)	(2,1)	(2,1)	(2,3)	(2,3)	0	(1,1)
(3,1)	(3,1)			(4,5)	(4,5)		, ,
				7			

RWI

value	k
(1,2),(1,2)	(2
(3,1)	(4
	(1,2),(1,2)

Key 1	value	
(2 2)	(2.2.)	

(2,3)	(2,3)
(4,5)	(4,5)
ata i	,

	2	(1)	R
1	2	00	1
	4		. `

				1120
1	Key	value	Key	Value
	(2,1)	(2,1)	(1,1)	(1,1)
	- 1		(2,1)	(1,1) (2,1)
1	,		1	

Key	value		
(1,2)	(1,2) (1,2)		
(3,1)	(3,1)		

$$(2,3)$$
 $(2,3)$ $(4,5)$

Key 1	value	
(2,1)	(2,1),	(2,1)
(1,1)	(1,1)	

1 \ - 1			
A	B		
١	2		
3	1		
2	3		

4 5

RW2

G	3
1	
1	
	1

Intersection using map reduce:

Intersection algo:

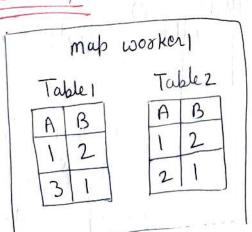
map (key, value):

box tuple in value: emit (tuple, tuple)

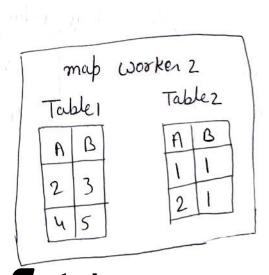
vieduce (key, values):

if values == [key, key] emit (key, key)

example



This document is available free of charge on



W	ω_1
Key 1	Value
(1,2)	(1,2),(1,2)
(3,1)	(3,1)
(2,1)	(2,1)

$m\omega 2$	1
Key	value
(2,3)	(2,3)
(4,5)	(4,5)
(1,1)	(1,1)
(2,1)	(2,1)

Apply hash bunction

	m wi	,	mw2	
Key	Value	key, value	Key value	Key val
(1,2)	(1,2),(1,2)	(2,1) (2,1)	(2,3) (2,3)	(1,1)
(3,1)	(3,1)	(1)	(4,5) (4,5)	(2,1) (2,
	1		7	
		K		

Swapping

	KMI		
Key 1	value	Key	value
(1,2)	(1,2),(1,2)	(2,3)	(2,3)
(3,1)	(3,1)	(4,5)	(4,5)

R	ω_{2}		4"
Key 1	value	key	value
(2,1)	(2,1)	(1,1)	(1,1) (2,1)
	1	1	

key | V (1,2) (1,2

R	u)	2	

key	value	
(2,11)	(2,1), (1,1)	(2,1)
(1,1)	(1,1)	
1.		

Key Value
(1,2)
(3,1)
(2,3)
(4,5)

Step 5 Ib length of "values" is greater than I, then emit.

Rey	valbus
(1,2)	(1,2)(12)

1, 191, 6	Rey	Balus.
(in the	(2,1)	(2,1)
	1111	10/ 4

Step 6

R	WI
M	B
1	2

Difference using mappiedice:

Dibberence algo

emit (tuple, R)

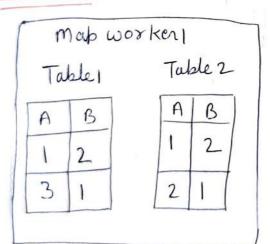
else: for tuple in value: emit (tuple, S)

vicduce (key, values):

if values == [R];

emit (key, key).

example



7	map w	orker 2.	
Tab	le I	Table	2,
A	B	A	B
2	3	1	1
4	5	2	1

R→ Tuble 1 S→ Tuble 2

m wi

mw2

Key	Value	Key	Value
(1,2)	[[,72]	(2,3)	[Ti]
		(4,5)	LT1]
(311)	[12]	(1,1) (2,1)	[T2]
		(2,1)	[T2]

cipply hash function.

(1,2) (3,1)	m WI Value [TI,T2] (TI]	(2,1)	value [T2]		=	Natue [Ti]	(1,1) (2,1)	Value [T2] [T2]
	\		K	l);	1	1	

RWI

Step3

Stepy

value

RW2 Key | value

[T2]

(2,1)

Key [T2]

[[2]

(1,1)

Key Value [Ti] (2,3)[T1, T2] (1,2)[Ti] (4,5) [Ti] (3,1)

RWI

Key

(2,1) RW2

value

[T1, T2] (1,2) [TI] (3,1) This document is available free of charge on (2,3)(4,5) Tpownloaded by Adwait Purao (adwait.purao@spit.ac.in)

Value

 $[T_2]$ (2,1)

Key

* eliminate keys that are present in both table he TIRTZ as well as only in Tz.



RW1

Key	Value	- 1
(311)	[7,]	-)
(2,3)	[7]	
(4,5)	TTI	

RW2 Key Value

Step 5

A	B
5	
3	1 1/1/
2	3
141	5

range of white good

Endland Calina

 $\Gamma(1 \mid -1, .)$

(,)

RW2 AB

al V Park

108 ...

11 17

111

4-1-11