

MAP REDUCE ILLUSTRATED

Ex: New search:

"who cover more news about
Mr. Barack Obama since Nov. 4"

Inputs:

url	content
cnn.com/index.html	
cnn.com/money/	
washpost.com/...	
nytimes.com/...	
	↓
	hundreds of millions of pages

MAP
REDUCE
MAGIC

news	#pages covered B.O.
cnn.com	512 pages
nytimes	478 pages
⋮	⋮

①

BASIC MAP REDUCE

Inputs

key	value
a1	
a2	
a3	
a4	

Map (key, value) {



3 generate intermediate key-value

Intermediate table

key	value
b1	v
b2	v
b2	v
b3	v
b3	v
b3	v



(MapReduce merge matching keys)



Final result

key	value
c1	..
c2	..
c3	..
c4	..
c5	..

Reduce (key, values),



3 generate final key-value

b1	v
b2	v v
b3	v v v

Ex : Simple Math

1	x
2	x
3	x
4	x

→ Map()

1	1
1	4
1	9
1	16

1	1	4	9	16
---	---	---	---	----

↓
Reduce()

x	30
---	----

⇓
final
result .

Ex: count words in all web pages

url	content
cnn.com/...	
espn/...	
wop/...	
⋮	⋮

→ Map () →

Obama	1
Garnett	1
Kevin	1
Barack	1
Obama	1
Barack	1
⋮	⋮

↓

Obama	1	1
Barack	1	1
Garnett	1	
Kevin	1	
⋮	⋮	

Reduce () ←

Word	Total Count
Barack	2
Obama	2
Garnett	1
Kevin	1
⋮	⋮

Final table.