

Tagging origins

Grouping by keys

Local join

Mapper 1
R

Key	Value
1	R ₁
2	R ₂
3	R ₃
4	R ₄

map

Key	Value
1	'R',R ₁
2	'R',R ₂
3	'R',R ₃
4	'R',R ₄

Mapper 1
S

Key	Value
1	S ₁
2	S ₂
3	S ₃
4	S ₄

map

Key	Value
1	'S',S ₁
2	'S',S ₂
3	'S',S ₃
4	'S',S ₄

Partitioning by key (Round Robin)

Reducer 1

Key	Tuple
1	'R',R ₁
	'S',S ₁
3	'R',R ₃
	'S',S ₃

reduce

Result

R ₁	S ₁
R ₃	S ₃

Reducer 2

Key	Tuple
2	'R',R ₂
	'S',S ₂
4	'R',R ₄
	'S',S ₄

reduce

Result

R ₂	S ₂
R ₄	S ₄

Map Phase

Reduce Phase