Data Engineering

Introduction to Parallel Execution

Shivnath Babu

Introduction to Parallel Execution

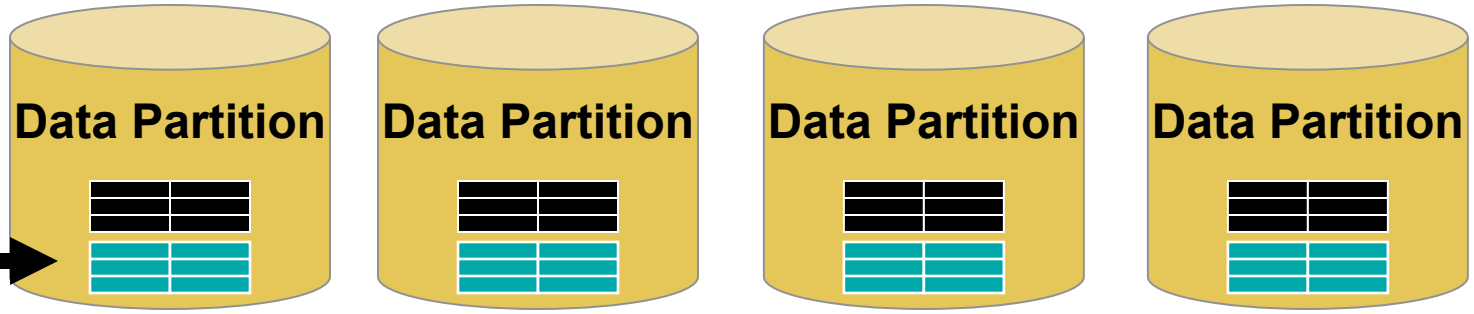
Query or
Spark Program Q

Answers

Spark
Driver

Translates Q into
an execution plan
and runs the plan

Data resides
on one or
more
machines



Example: At a Company

Query 1: Is there an employee named “Nemo”?
Query 2: What is “Nemo’s” salary?
Query 3: How many departments are there in the company?
Query 4: What is the name of “Nemo’s” department?
Query 5: How many employees are there in the “Accounts” department?

Employee

ID	Name	DeptID	Salary	…
10	Nemo	12	**120K**	…
20	Dory	156	**79K**	…
40	Gill	89	**76K**	…
52	Ray	34	**85K**	…
…	…	…	…	…

Department

ID	Name	…
12	IT	…
34	Accounts	…
89	HR	…
156	Marketing	…
…	…	…

Counting the number of records that will be read or transferred over the network in a parallel execution

- We have a dataset R with two attributes A and B
- There are 10000 records in R, with 2500 unique values of A and 5000 unique values of B
- See Figure 1 on next slide

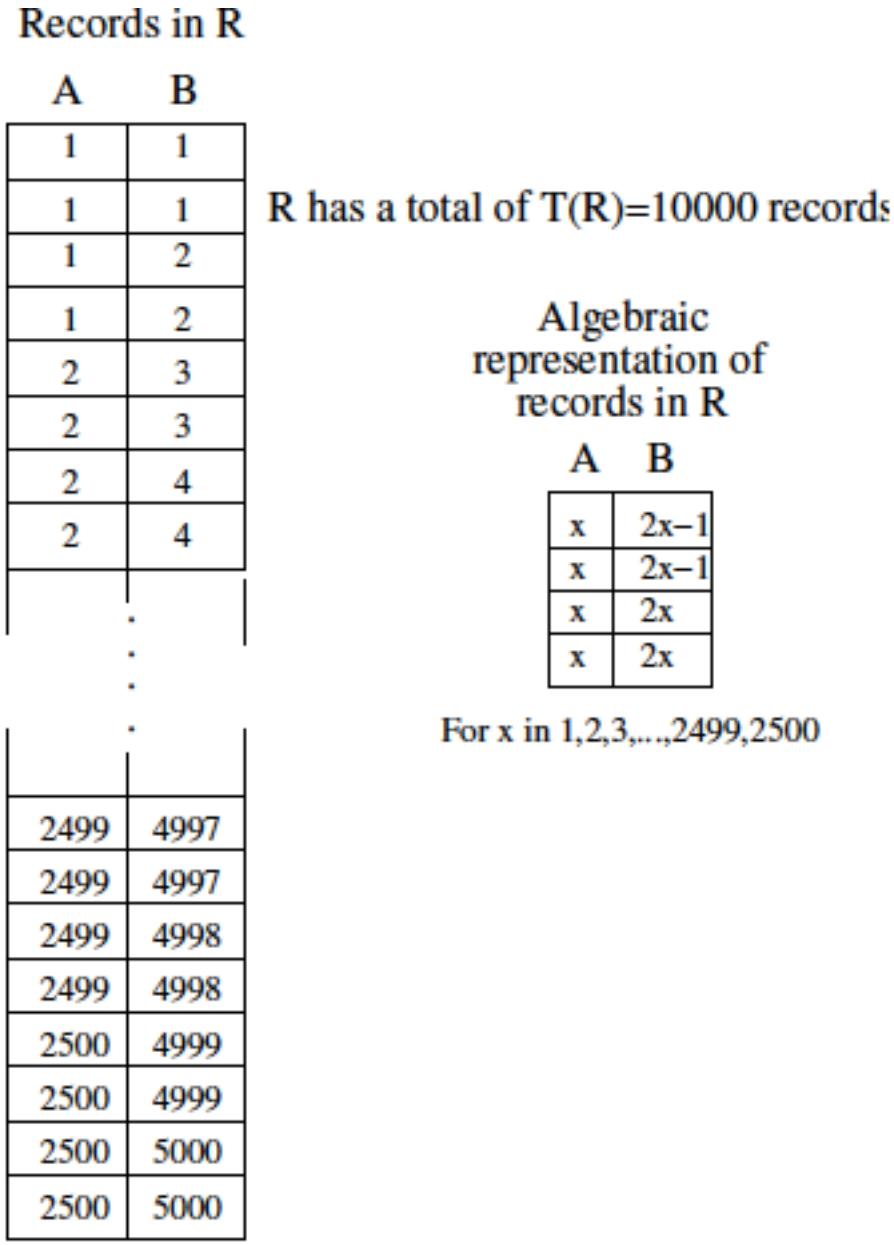


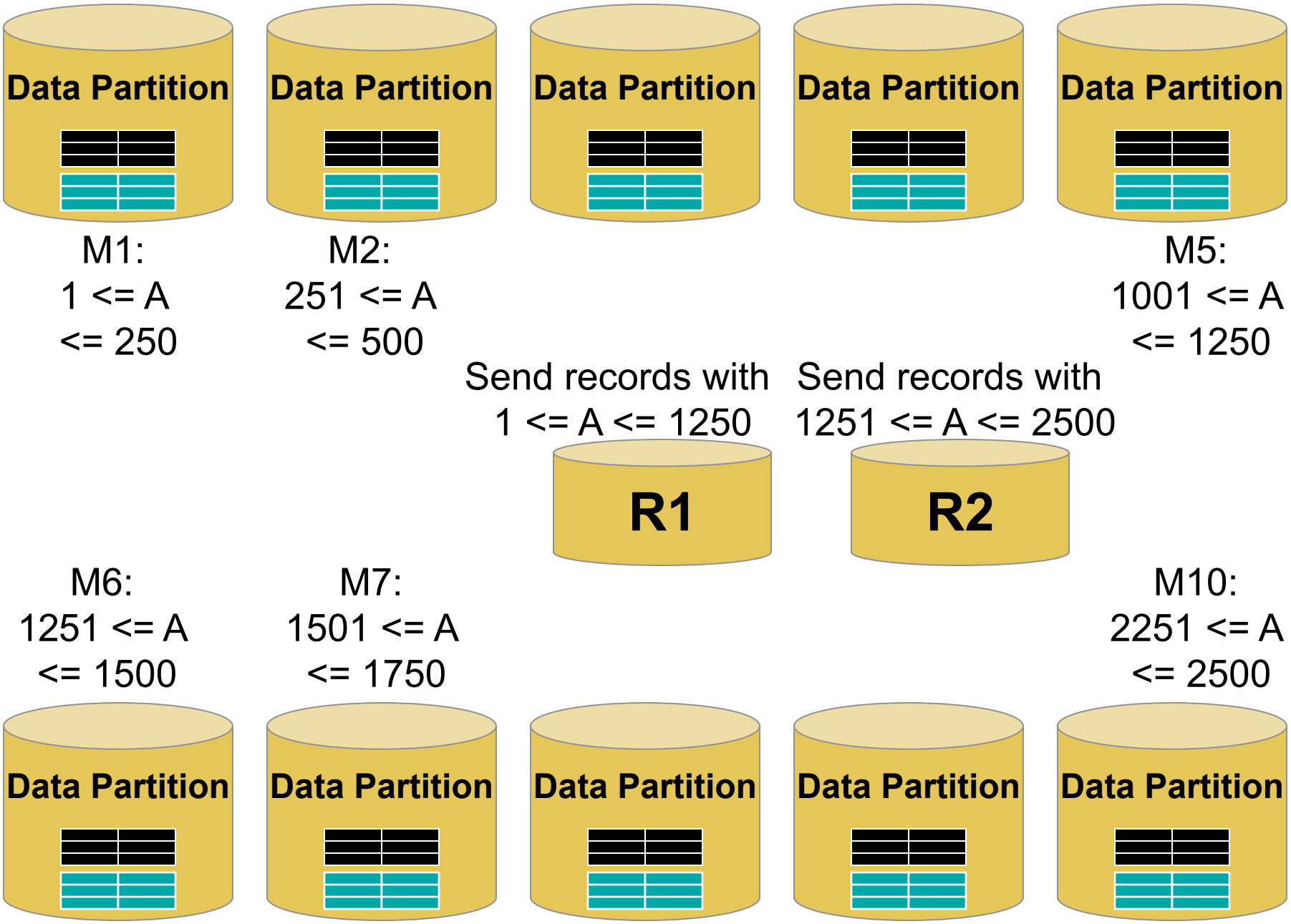
Figure 1: Figure showing the contents of records in R

Counting exercise (contd.)

- Dataset R has 10 partitions
- R is stored on 10 machines, with one partition per machine
- We want to find the result of:

```
Select    A, MAX(B)
From      R
Where     B >= 1000 and B < 2000
Group By  A
```

- See Figure on next slide


Data Partition
Data Partition
Data Partition
Data Partition
Data Partition
M1:
1 <= A
<= 250
M2:
251 <= A
<= 500
M5:
1001 <= A
<= 1250
Send records with
1 <= A <= 1250
Send records with
1251 <= A <= 2500
R1
R2
M6:
1251 <= A
<= 1500
M7:
1501 <= A
<= 1750
M10:
2251 <= A
<= 2500
Data Partition
Data Partition
Data Partition
Data Partition
Data Partition

Partitioning and Load Balancing

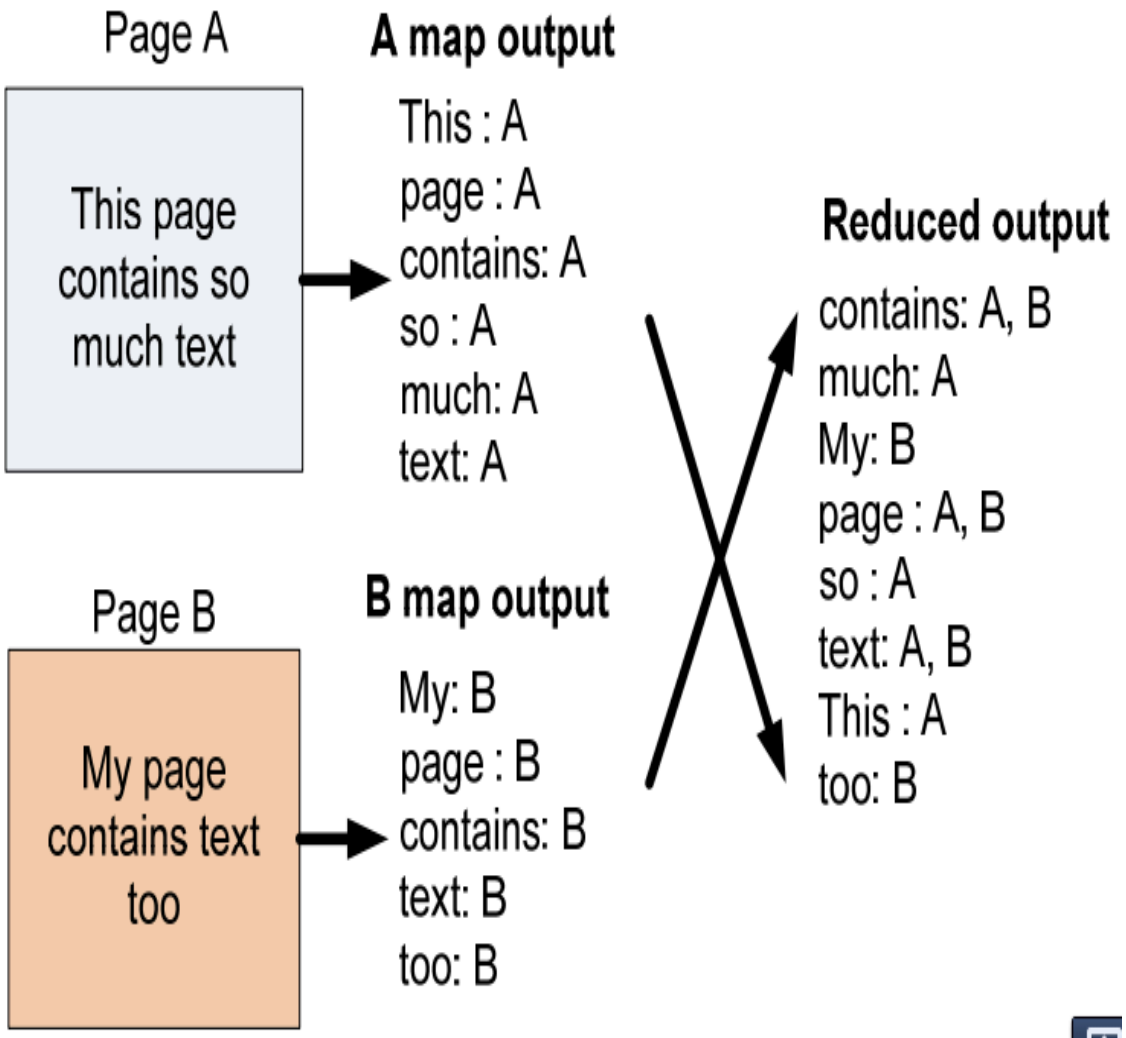
- Looking at data as Key-Value pairs
- The Map Vs. Reduce model of parallel execution
- Partitioning:
  - Range Partitioning
  - Hash partitioning
  - List partitioning
- The Shuffle step in parallel execution
- Load balancing and skew problems

Same example (contd.)

- Suppose records in R are partitioned randomly across M1-M10. What, if anything, changes?

Other Common Tasks that Need Parallel Execution

- Word counting
- Inverted indexes



Other Common Tasks that Need Parallel Execution

- TF-IDF
  - http://blog.cloudera.com/wp-content/uploads/2010/01/5-MapReduceAlgorithms.pdf

Other Common Tasks that Need Parallel Execution

- Sorting
- PageRank (will consider this when we study Graph Processing)
- …

Join Example: Store that Sells Cars

Owners of Honda Accords who are <= 23 years old

Make	Model	OwnerID	ID	Name	Age
Honda	Accord	12	12	Nemo	**22**
Honda	Accord	156	156	Dory	**21**

Join (Cars.OwnerID = Owners.ID)

Filter (Make = Honda and Model = Accord)

Filter (Age <= 23)

Cars

Make	Model	OwnerID
Honda	Accord	12
Toyota	Camry	34
Mini	Cooper	89
Honda	Accord	156
…	…	…

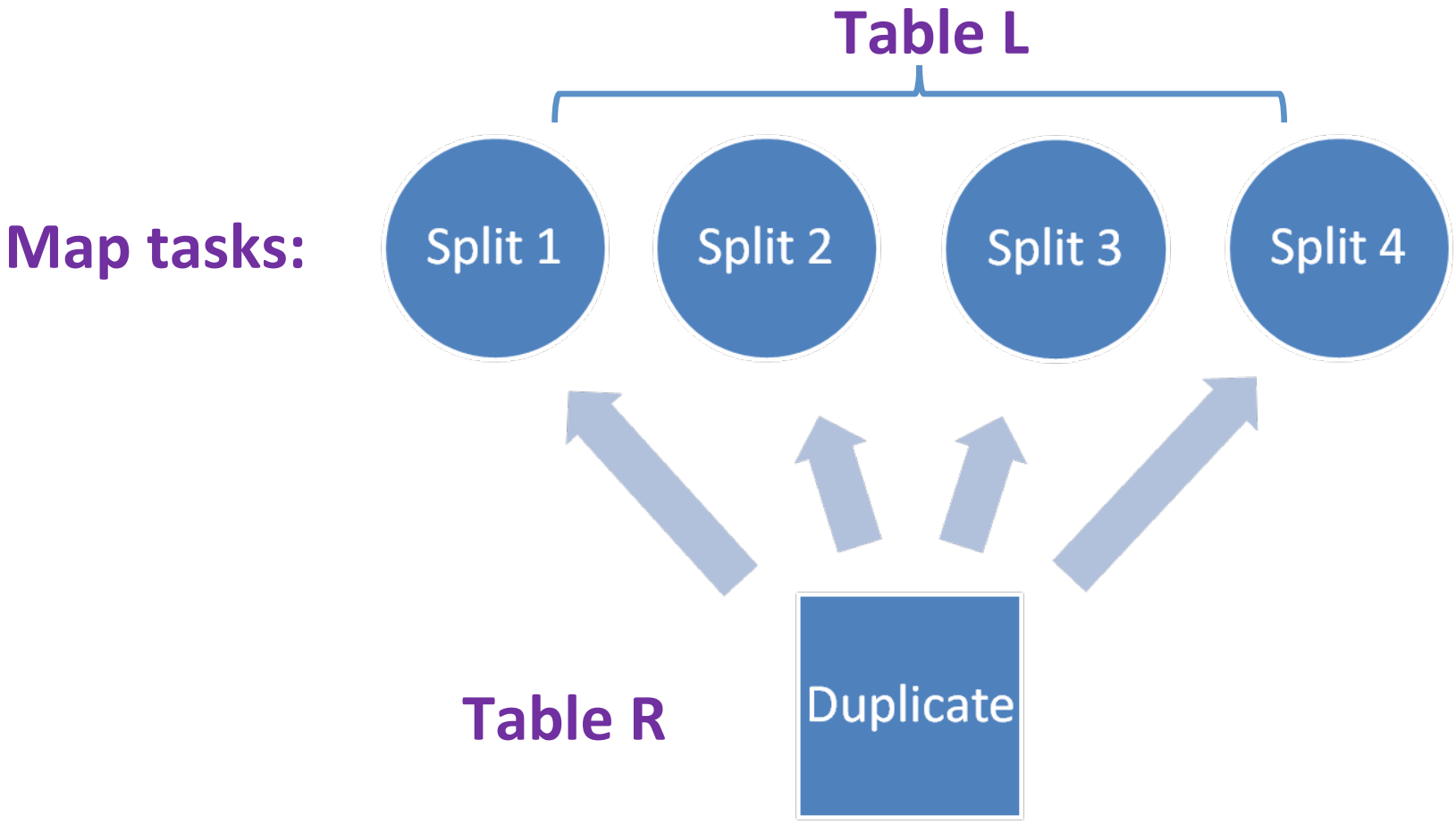
Owners

ID	Name	Age
12	Nemo	**22**
34	Ray	**42**
89	Gill	**36**
156	Dory	**21**
…	…	…

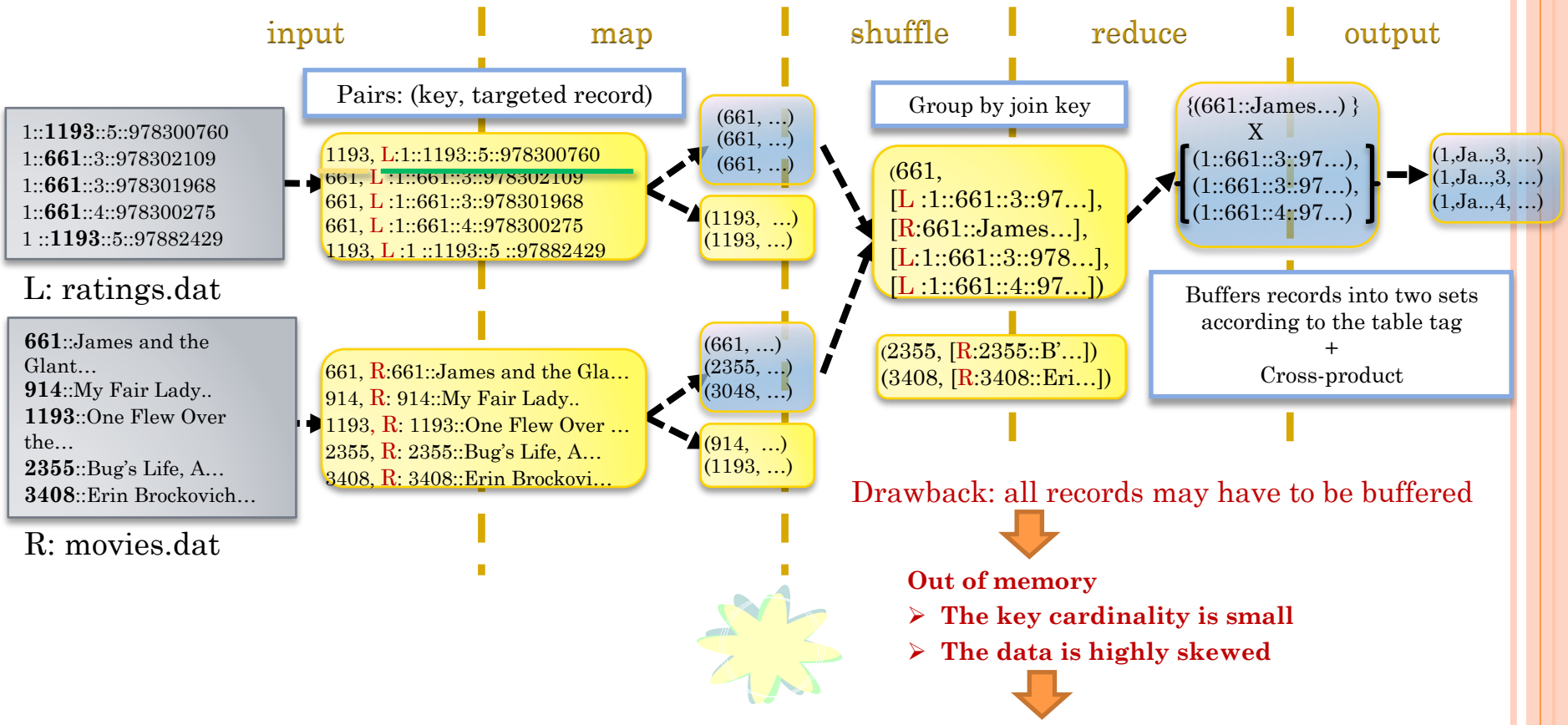
Parallel Execution of Joins in MapReduce Style

- Two typical types of join
  - Map-side join
  - Reduce-side join

Map-side Join



REDUCE-SIDE JOIN



Phase /Function	Improvement
Map Function	Output key is changed to a composite of the join key and the table tag.
Partitioning function	Hashcode is computed from just the join key part of the composite key
Grouping function	Records are grouped on just the join key