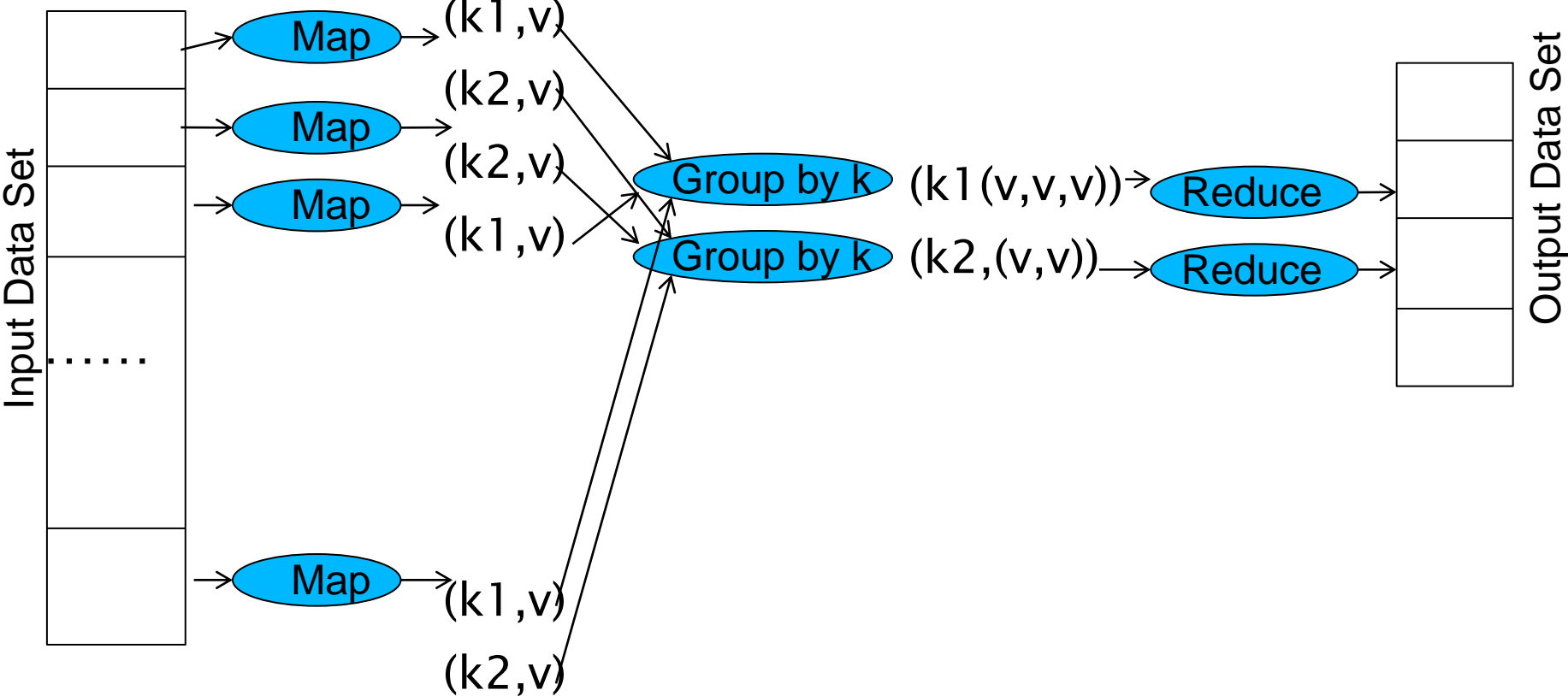


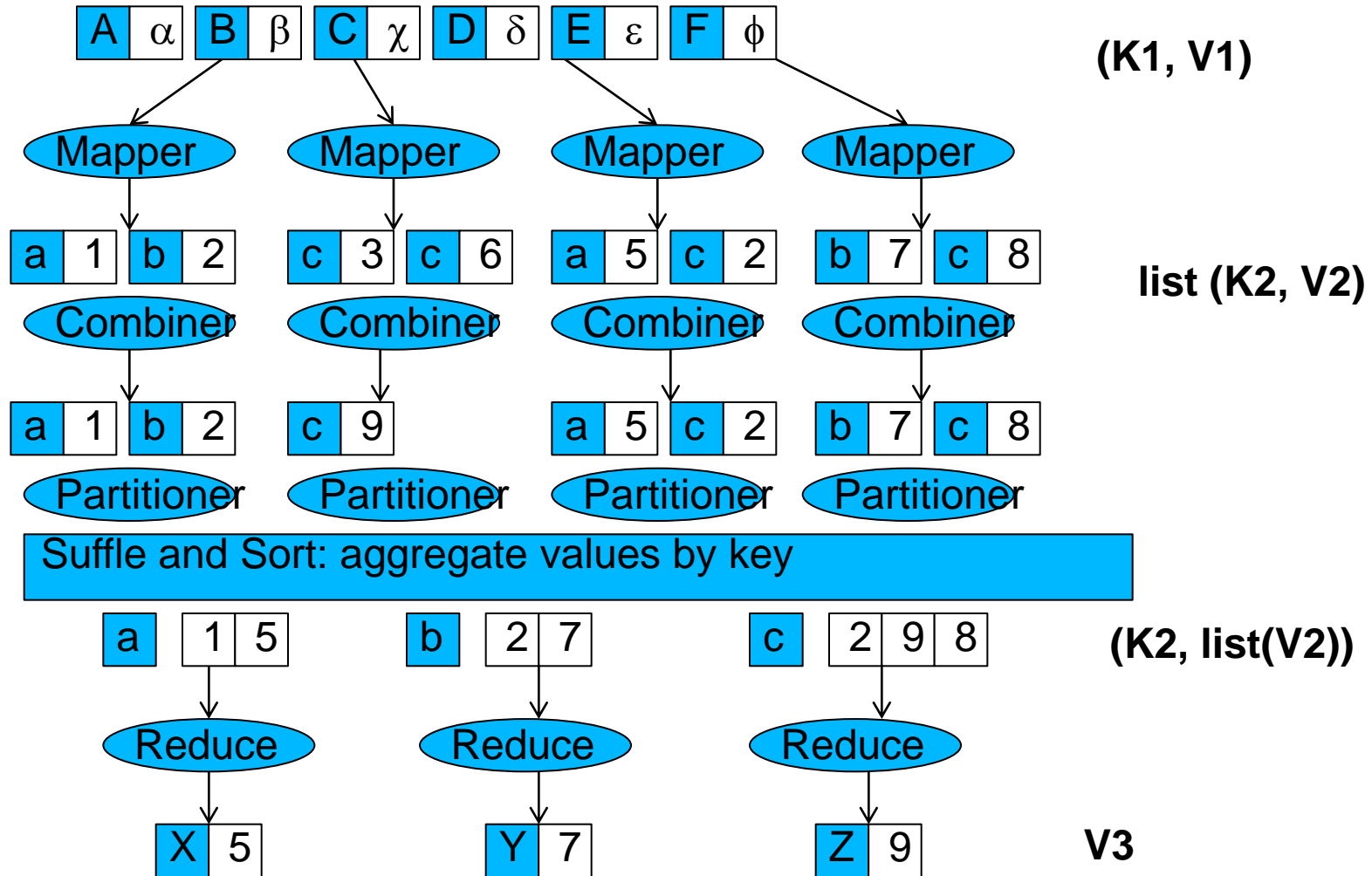
MapReduce Processing [Val 2010]

$(key1, val1) \xrightarrow{\text{map}} \text{list}(key2, val2)$

 $(key2, \text{list}(val2)) \xrightarrow{\text{reduce}} val3$



Combiner & Partitioner



A Simple Word Count Example [Lin 2010]

Count the number of occurrences of every word in a text collection

■ Method Map (docid id, doc d)

for all term t in d do

Emit(term t, count 1)

■ Method Reduce (term t, counts [c1, c2,, cn])

sum=0

for all count c in counts [c1, c2,, cn] do

Sum = sum+c

Emit(term t, count sum)

MapReduce Example [Val 2010]

- EMP (ENAME, TITLE, CITY); **Question:** For each city, return the number of employees whose name is "Smith"?
- SQL Query: SELECT CITY, COUNT(*) FROM EMP WHERE ENAME LIKE "\%Smith" GROUP BY CITY
- With MapReduce
 - Map (Input (TID,emp), Output: (CITY,1)) if emp.ENAME like "%Smith" return (CITY,1)
 - Reduce (Input (CITY,list(1)), Output: (CITY,SUM(list(1)))) return (CITY,SUM(1*))

References

- [Lin 2010] J. Lin & C. Dyer; “Data-Intensive Text Processing with MapReduce”; Publisher: Morgan & Claypool Publishers; ISBN: 9781608453429
- [Val 2010] P. Valduriez & E. Pacitti; « Data Management in the Cloud – Current Issue and Research Direction »; In: DNAC Congres, Paris, Nov. 2010.