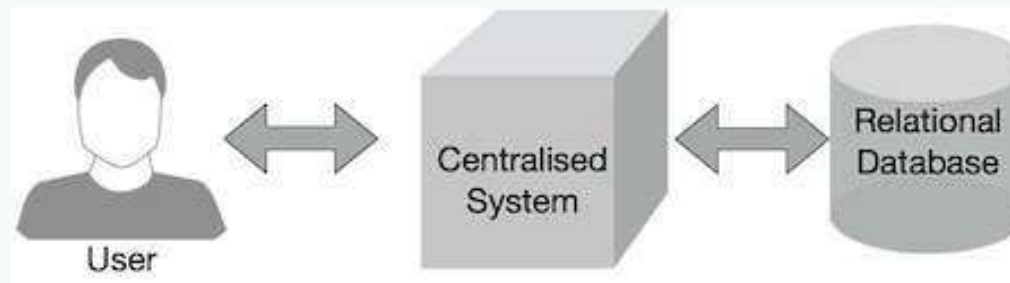


**Map Reduce**

# Traditional Enterprise Systems

---

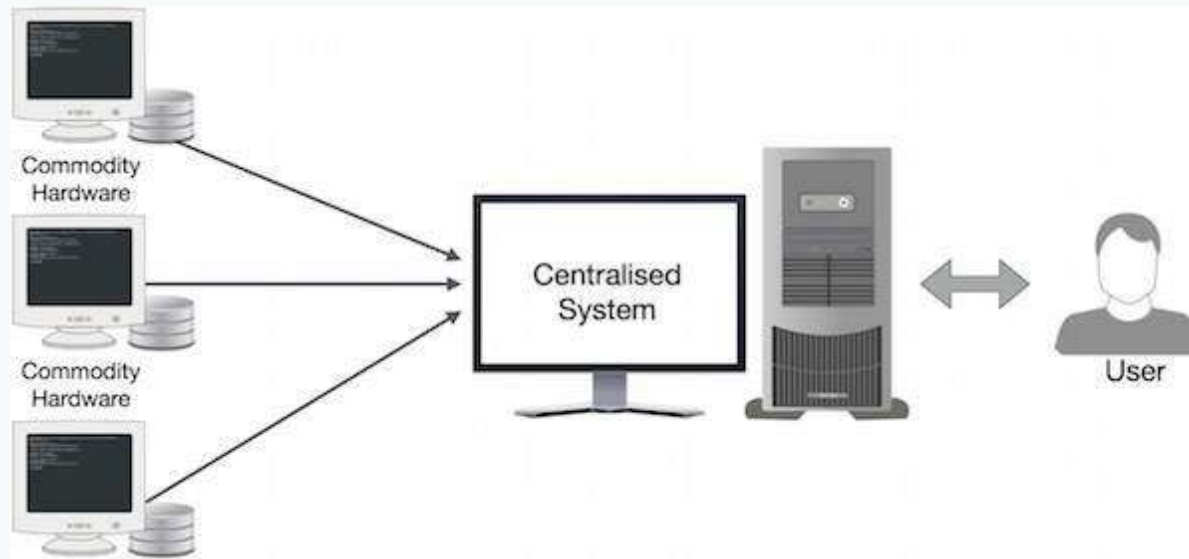
- Have a centralized server to store and process data
- Not suitable to process huge volumes of scalable data
- Creates too much of a bottleneck while processing multiple files simultaneously



# Map Reduce

---

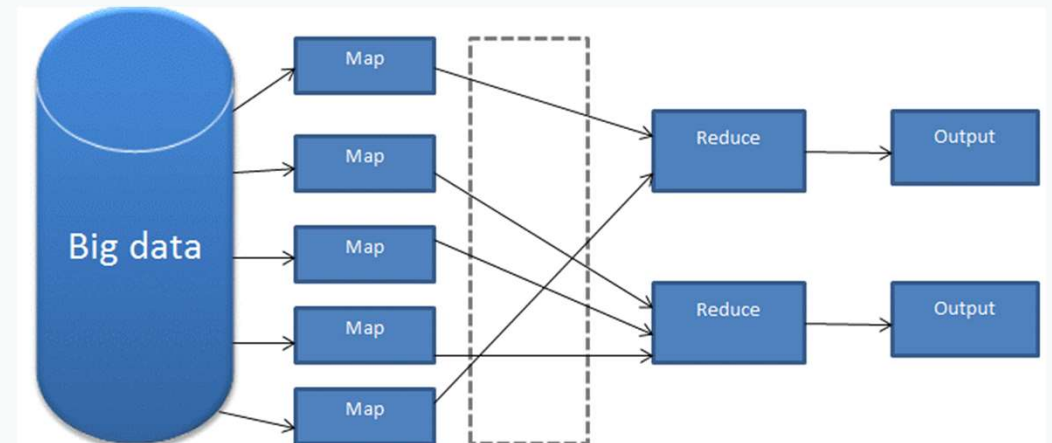
- Google solved this bottleneck issue using an algorithm called MapReduce.
- MapReduce divides a task into small parts and assigns them to many computers.
- Later, the results are collected at one place and integrated to form the result dataset.



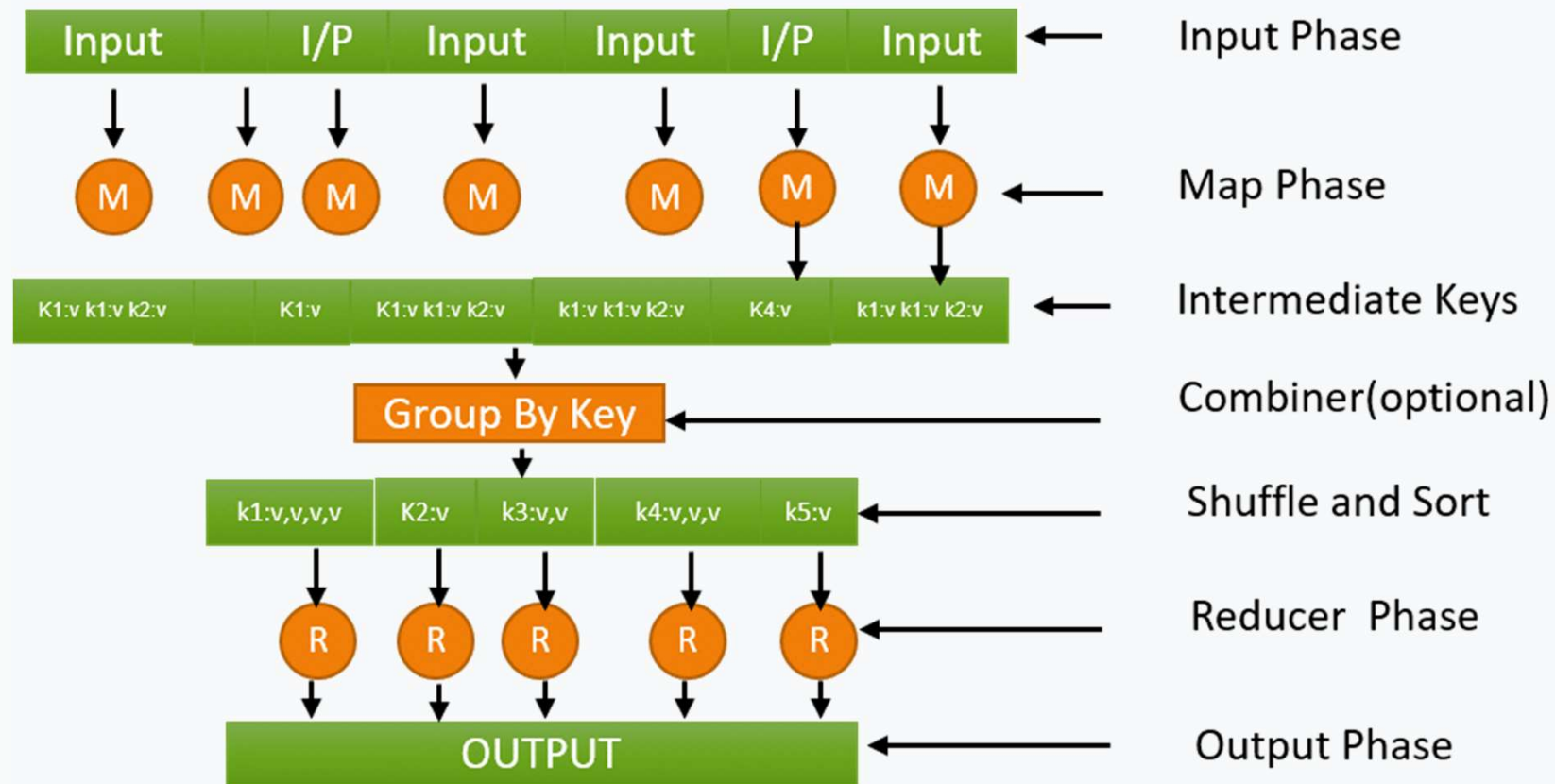
# How MapReduce Works?

---

- The MapReduce algorithm contains two important tasks, namely
  - Map task: takes a set of data and converts it into another set of data, where individual elements are broken down into tuples (key-value pairs).
  - Reduce task: takes the output from the Map as an input and combines those data tuples (key-value pairs) into a smaller set of tuples.
- The reduce task is always performed after the map job.



# Example



Thank You