





Tarbiat Modares
University

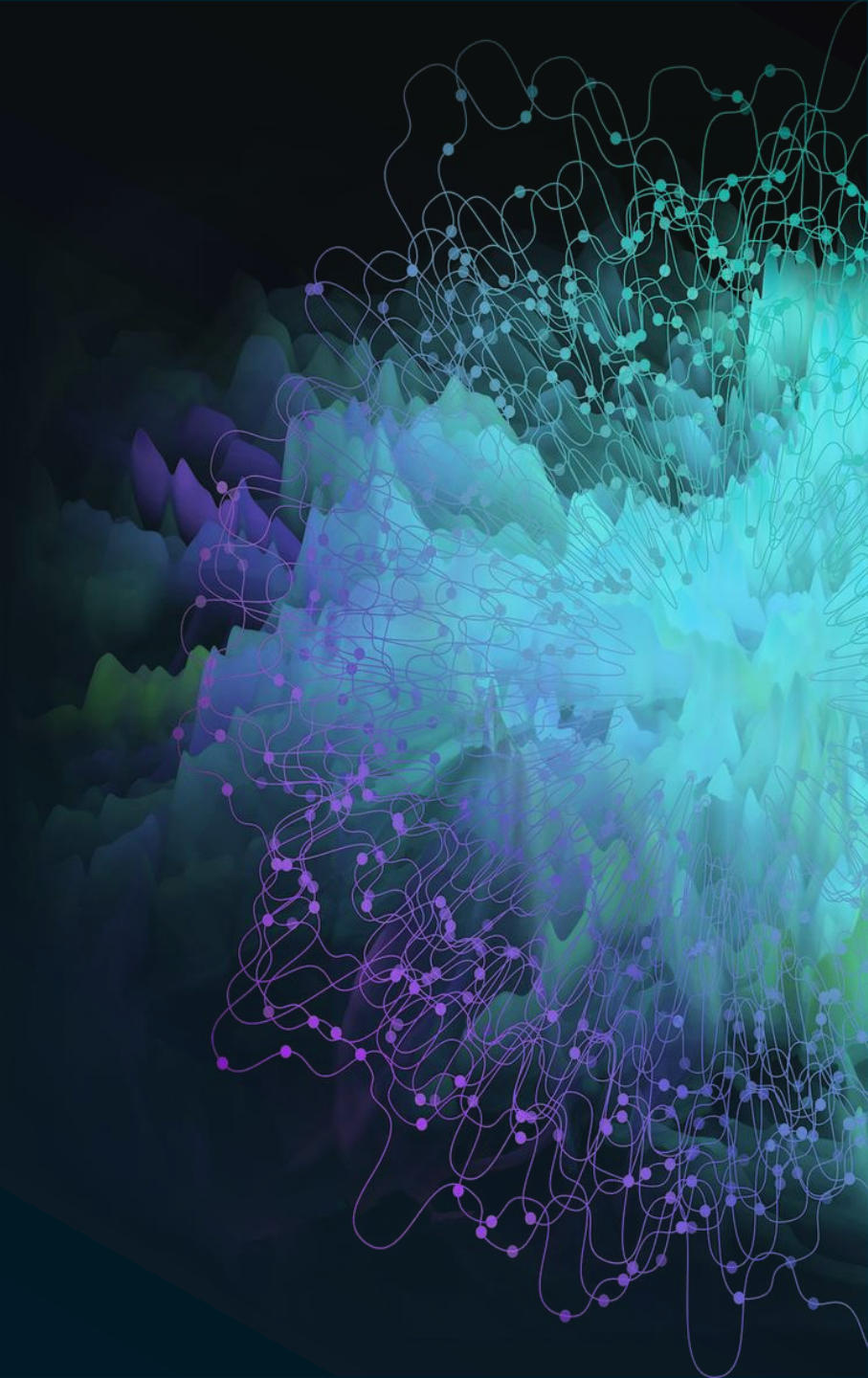
Big Data Course

Dr Akhondzade - Fall 2021

About me:

Amirreza Najafi

- Tech Lead at Carpo startup
- MSc of Information Technology at TMU
- Big Data science researcher



Lecture outlines

- A quick review of basic python
- MapReduce review
- mrjob Module
 - Homework 1
- Functional programming basics
 - Homework 2
 - Homework 3
- Introduction to Apache Spark
- Apache Spark concepts
- Apache Spark deployment modes
- Apache Spark setting up
 - Setting up task
- Apache Spark's structured APIs
- Spark SQL
- MLlib
 - Homework 4

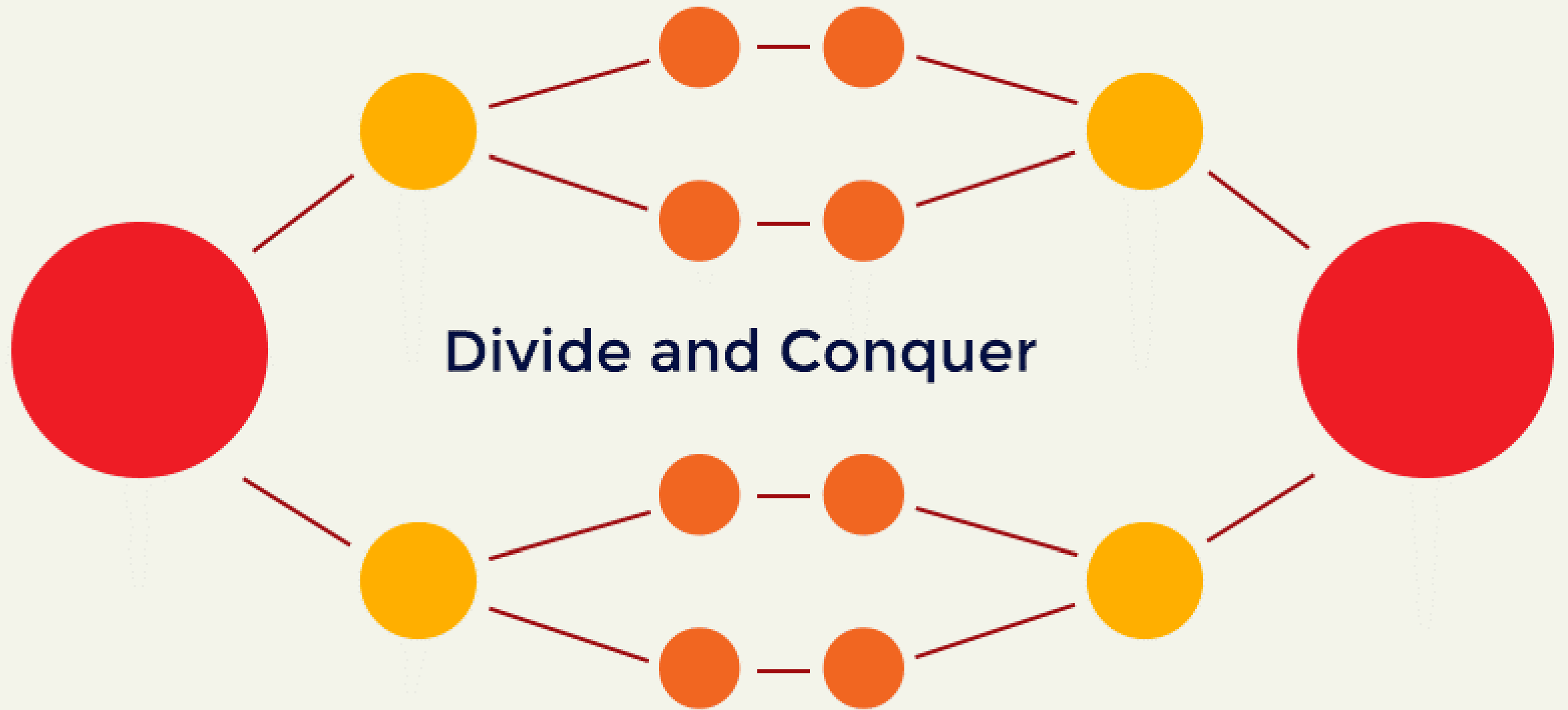


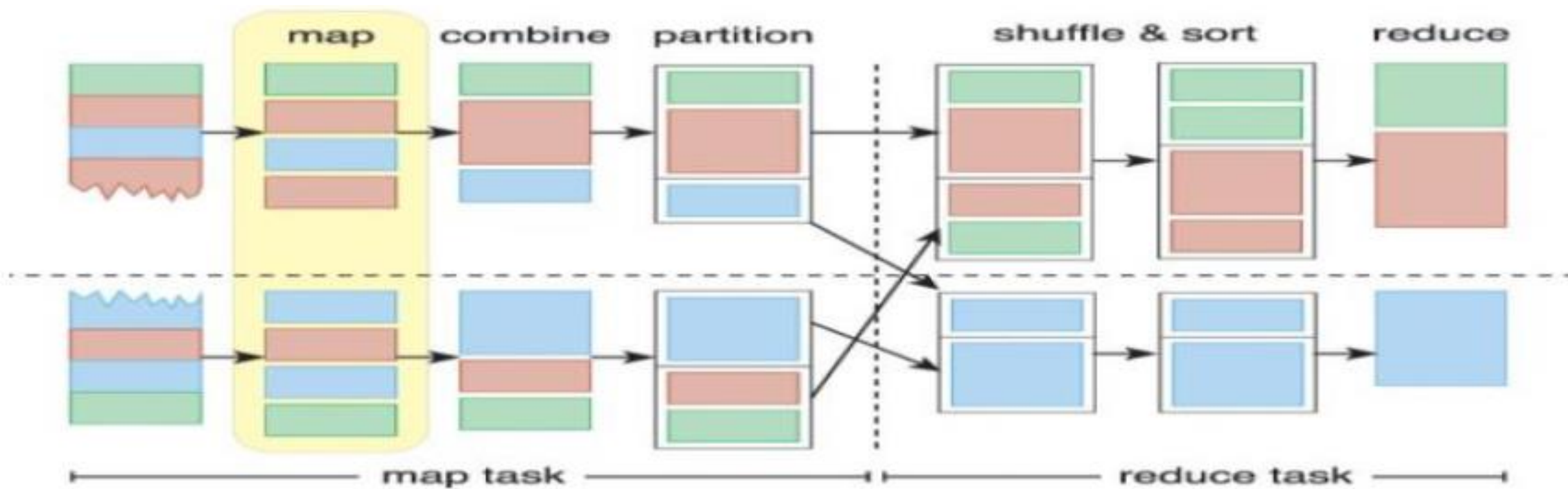
MapReduce

Big Data Programming Model

- Support Large Data
- Provide FT
- Scale Out







MRJob



Why use mrjob instead of X?

And

Why use X instead of mrjob?