

Project: MapReduce with Hadoop

Phase 2

Phase 2:

The input data is about various chess game results.

Example:

[Result "1-0"]
[Result "0-1"]
[Result "1-0"]
[Result "1/2-1/2"]

Each row is a different game. The first row states that White won that game (1-0). The second row states that Black has won that game (0-1). The last row states that game was draw ($\frac{1}{2}$ - $\frac{1}{2}$)

Your objective is to write a simple Hadoop MapReduce program that calculates how many games white won, how many games black won and how many games were draw. You also need to calculate the fraction of games won by each of them (between 0 and 1)

You need to solve this on AWS.

Example Output:

White	2	0.5
Black	1	0.25
Draw	1	0.25

Explanation:

White won 2 games. Fraction of winning is 0.5 i.e. 50% of the games (Total 4 games)
Black won 1 games. Fraction of winning is 0.25 i.e. 25% of the games (Total 4 games)
1 game was draw, Fraction of winning is 0.25 i.e. 25% of the games (Total 4 games)

For submission:

For submitting the phase 2 on SAKAI, you need to submit the following files:

- 1) Your MapReduce java file
- 2) Output files generated
- 3) Jar file you created
- 4) A small (less than 1 page) PDF report explaining how you solved this problem and any difficulties/challenges you faced. Also write an approximate amount of time you spent on phase 2.