

Big Data architektúrák és elemző módszerek Gyakorlat

Gombos Gergő

Feladat1 - Grep

- Készítsünk egy grep funkcionalitást elvégző MapReduce Jobot!
- Grep feladata:
 - Visszaadni azokat a sorokat, amelyben a keresett karaktersorozat megtalálható
 - Keresett szó: "elephant"
 - Eredmény:
 - An **elephant** gentle and mellow.
 - Hadoop is the **Elephant** King!
 - The **elephant** plays well with Sqoop.

Feladat1 - Grep

- Egészítsük ki a megoldást, hogy paraméterben kapjuk a keresett szótöredéket!
 - Driver:
 - `conf.set("text", args[0]);`
 - Mapper:
 - `Configuration conf = context.getConfiguration();`
 - `String grepString = conf.get("text");`

Feladat2 - Legmelegebb nap

- Készítsünk saját Writeable-t, amely tárolja az időpontot és a hőmérsékletet!
- Keressük meg minden naphoz a legmelegebb órát!
- Mapper: nap, (óra, hőmérséklet)
- Reducer: nap, (óra, maxhőmérséklet)

Feladat3 - WordMean

- Számoljuk meg az átlagos szóhosszt!
- Counter segítségével számoljuk meg a szavakat!
- Driver:
 - enum words {num}
- Mapper:
 - `context.getCounter(WMDriver.words.num).increment(1);`

Köszönöm a Figyelmet!