Spark pair RDD functions

Nhóm 6

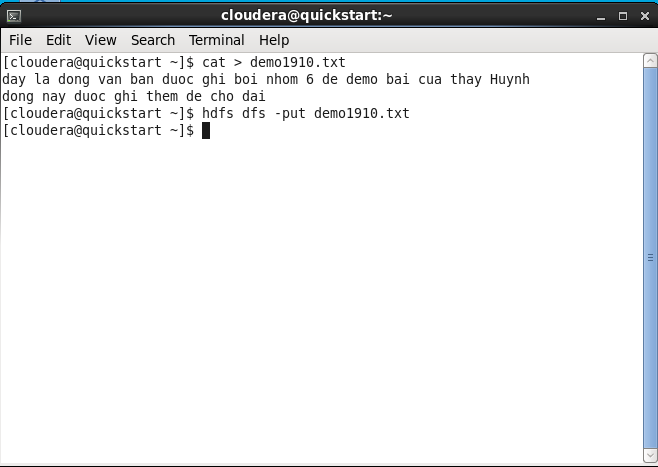
Ngô Phạm Anh Vũ

Phạm Thái Quang

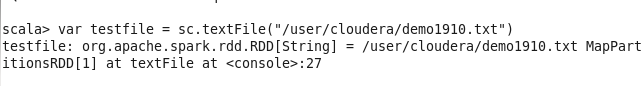
Nguyên Văn Tiến

Ví dụ về đếm từ

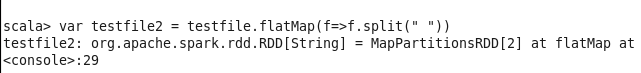
B1: Tạo tệp đặt tên bất kì với đuôi .txt và đặt vào trong hdfs



B2: Đọc tệp và tạo RDD



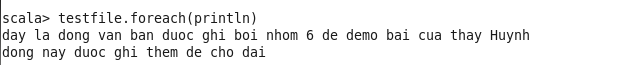
B3: Áp dụng phép biến đổi flatMap



B4: Ấp dụng phép biến đổi Map



In văn bản trong file .txt



In ra kết quả

