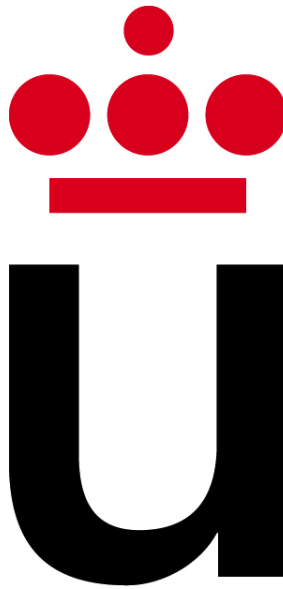


FORO DE PREGUNTAS

Análisis de Big Data

PREGUNTAS TEMA 2

José Ignacio Escribano



MÓSTOLES, 25 DE MARZO DE 2016

Índice

1. Preguntas	1
1.1. Las dos fases principales en cualquier tarea MapReduce son map y reduce. Sin embargo, también se puede añadir un nuevo tipo de fase llamada combine. Busque información sobre este tipo de fase y explique con un ejemplo qué ventajas proporciona cuando se añade a ciertos procesos MapReduce.	1

1. Preguntas

- 1.1. ¿Cuál es el resultado de la ejecución del segundo ejemplo proporcionado en el Tema 2, llamado top10_words.py? ¿Qué tareas realiza exactamente este archivo? ¿Por qué la lectura de la salida de un proceso con mrjob se debe poner en un archivo diferente (top10_words.py) de aquel en que declaramos la clase principal de la tarea (MRCountWords.py)?**
- 1.2. Las dos fases principales en cualquier tarea MapReduce son map y reduce. Sin embargo, también se puede añadir un nuevo tipo de fase llamada combine. Busque información sobre este tipo de fase y explique con un ejemplo qué ventajas proporciona cuando se añade a ciertos procesos MapReduce.**