RDDs y Spark

1.- De lo que se puede ver en la presentación. ¿En qué sentido se puede decir que Spark es mejor que Hadoop (o la versión de MapReduce de Google)?

Spark es más fácil de usar que Hadoop, ya que viene equipado con APIs amigables para diversos lenguajes de programación (como Java, Python, Scala, etc.)

Spark permite que los procesos se aceleren, ya que trabaja en la memoria (in memory), si el sistema presenta una gran cantidad de memoria de almacenaje, los procesos que se ejecuten con Spark superarían en rápido rendimiento a los procesos que se ejecutan con Hadoop.

2.- ¿Cuál es el truco principal que usaron para esta mejora?

Aprovecharon la facilidad de muchos algoritmos paralelos para la recuperación de bajo nivel de los sistemas distribuidos.