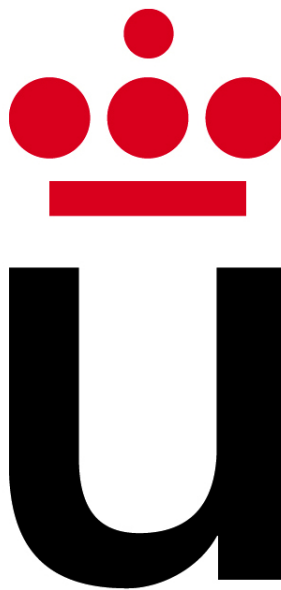


FORO DE PREGUNTAS

Análisis de Big Data

PREGUNTAS TEMA 6

José Ignacio Escribano



MÓSTOLES, 26 DE MARZO DE 2016

Índice

1. Preguntas	1
1.1. Cloudera Manager es el principal componente de la plataforma Cloudera Enterprise que no está disponible como software libre. Responda a las siguientes preguntas: Resuma brevemente las principales funciones que ofrece. ¿Cuál sería el componente equivalente en la distribución Hadoop de Hortonworks? ¿Considera que es una buena estrategia por parte de Cloudera mantener este componente como propietario? ¿Qué ventajas espera conseguir? ¿Qué riesgos asume ante la aparición de plataformas alternativas como ODPI?	1
1.2. Apache Zeppelin se está convirtiendo rápidamente en uno de los proyectos que más actividad y atención concentra en el ecosistema Apache Hadoop. Recabe información sobre el proyecto y conteste a las siguientes preguntas: ¿Cuáles son las principales funciones que ofrece Zeppelin? ¿Está ya integrado en alguna de las distribuciones Hadoop de referencia (Cloudera, MapR, Hortonworks, etc.)? ¿Por qué considera que el producto está orientado principalmente a su integración con Apache Spark? . . .	1

1. Preguntas

- 1.1. Cloudera Manager es el principal componente de la plataforma Cloudera Enterprise que no está disponible como software libre. Responda a las siguientes preguntas: Resuma brevemente las principales funciones que ofrece. ¿Cuál sería el componente equivalente en la distribución Hadoop de Hortonworks? ¿Considera que es una buena estrategia por parte de Cloudera mantener este componente como propietario? ¿Qué ventajas espera conseguir? ¿Qué riesgos asume ante la aparición de plataformas alternativas como ODPI?**
- 1.2. Apache Zeppelin se está convirtiendo rápidamente en uno de los proyectos que más actividad y atención concentra en el ecosistema Apache Hadoop. Recabe información sobre el proyecto y conteste a las siguientes preguntas: ¿Cuáles son las principales funciones que ofrece Zeppelin? ¿Está ya integrado en alguna de las distribuciones Hadoop de referencia (Cloudera, MapR, Hortonworks, etc.)? ¿Por qué considera que el producto está orientado principalmente a su integración con Apache Spark?**