

HADOOP, HERRAMIENTAS PARA APLICACIÓN DE BIG DATA, TABLEAU SOFTWARE.

AUTORES

David Sarmiento

Diego García

Cesar Ariza

José Salas

RESUMEN

Hadoop ha crecido hasta representar mucho más que una fuente pequeña de código para difundir el trabajo de un grupo de equipos. Este núcleo resultó ser tan útil que una gran colección de proyectos orbita ahora alrededor de él, algunos ofrecen limpieza de datos, otros monitorean el progreso y algunos entregan el almacenamiento sofisticado de datos, La comunidad Hadoop está evolucionando rápidamente para incluir a las empresas que ofrecen servicios de apoyo, alquilar tiempo en clústeres gestionados, construir mejoras sofisticadas para el núcleo de código abierto o añadir sus propias herramientas a la mezcla.

El presente artículo tiene como propósito mostrar las herramientas existentes al momento de utilizar Hadoop y describir algunas características de los componentes principales que abarca dentro de Big Data.

Palabras Clave. Big data, Hadoop, almacenamiento, base de datos, Tableau, BI y arquitectura.

REFERENCIAS

- [1] Gartner. (2012). "The Importance of Big Data" ; (2013) "Top Technology Predictions for 2013 and Beyond"
- [2] Artículo "In Perspective" de Fidelity Worldwide Investment (2012) about the Big Data.
- [3] MapReduce. Simplified Data Processing on Large Clusters. <http://research.google.com/archive/mapreduce-osdi04.pdf>. MapReduce. <http://developer.yahoo.com/hadoop/tutorial/module4.html>
- [4] Introduction to the Hadoop SoftwareEcosystem - <http://www.revelytix.com/?q=content/hadoop-ecosystem>. Consulta 07-112015
- [5] Apache PIG -<http://pig.apache.org/>
- [6] Apache Hive - <http://hive.apache.org/>
- [7] Apache Mahout - <http://mahout.apache.org/>
- [8] "VizQL: un lenguaje de consulta, análisis y visualización". SIGMOD '06 Actas de la conferencia internacional 2006ACM SIGMOD sobre la Gestión de los datos. . ACM doi: 10.1145 /1142473.1142560. ISBN 1-59593-434-0
- [9] Tableau Business Intelligence. Tableau Software. <http://www.tableau.com/businessintelligence> Consulta 2015-01-11.
- [10] A. Labrador, P. Wightman, A. Santander, D. Jabba and M. Jimeno, "Tis-Bad: A Time Series-Based Deobfuscation Algorithm", Investigación e Innovación en Ingenierías, vol. 3, no. 1, 2015. DOI: <https://doi.org/10.17081/invinno.3.1.2035>