Ảnh có chứa văn bản, ảnh chụp màn hình, Phông chữ

Mô tả được tạo tự động  
  
  
Parallel processing   
  
Ảnh có chứa văn bản, ảnh chụp màn hình, Phông chữ

Mô tả được tạo tự động

**BIG DATA TOOLS**Ảnh có chứa văn bản, ảnh chụp màn hình, Phông chữ

Mô tả được tạo tự động

**OPEN SOURCE AND BIG DATA**

Main components of Hadoop ecosystem :

+Hadoop

+ HDFS

+YARN

Ảnh có chứa văn bản, ảnh chụp màn hình, Phông chữ

Mô tả được tạo tự động

**BEYOND THE HYPE**

**1 zetta = 1 bill tera**

**Ảnh có chứa văn bản, ảnh chụp màn hình, Phông chữ, số

Mô tả được tạo tự động**

**In this lesson, you learned that:**

* **Personal assistants like Siri, Alexa, and Google Now use Big Data and IoT to gather data and devise answers.**
* **Big Data Analytics helps companies gain insights from the data collected by IoT devices.**
* **Big Data requires parallel processing on account of massive volumes of data that are too large to fit on any one computer.**
* **“Embarrassingly parallel” calculations are the kinds of workloads that can easily be divided and run independently of one another. If any single process fails, that process has no impact on the other processes and can simply be re-run.**
* **Open-source projects, which are free and completely transparent, run the world of Big Data and include the Hadoop project and big data tools like Apache Hive and Apache Spark.**
* **The Big Data tool ecosystem includes the following six main tooling categories: data technologies, analytics and visualization, business intelligence, cloud providers, NoSQL databases, and programming tools.**

**Ảnh có chứa văn bản, ảnh chụp màn hình, biểu đồ, Phông chữ

Mô tả được tạo tự động**