Cloud Big Data Solutions   (https://cloud.google.com/products/big-data/)

What is big data?

Big data refers to data that would typically be too expensive to store, manage, and analyze using traditional (relational and/or monolithic) database systems. Usually, such systems are cost-inefficient because of their inflexibility for storing unstructured data (such as images, text, and video), accommodating "high-velocity" (real-time) data, or scaling to support very large (petabyte-scale) data volumes.

For this reason, the past few years has seen the mainstream adoption of new approaches to managing and processing big data, including Apache Hadoop and NoSQL database systems. However, those options often prove to be complex to deploy, manage, and use in an on-premises situation.

Where does big data come from?

Whereas in the past most customer data could be categorized as well-structured (such as bank) transactions, today, the massive "exhaust" that organizations produce daily in the form of unstructured online customer interaction data dwarfs that which was produced only a few years ago. The recent emergence of the "Internet of Things," the term describing the global network of billions of interconnected devices and sensors, has caused an explosion in the volume of data in the form of text, video, images, and even audio. Finally, in some regulated industries, access to data that would otherwise be archived is now often needed for compliance reasons.

Why is big data important?

The ability to consistently get business value from data is now a trait of successful organizations across every industry, and of every size. In some industries (such as Retail, Advertising, and Financial Services, with more constantly joining the list), it's even a matter of survival.

Data analytics only returns more value when you have access to more data, so organizations across multiple industries have found big data to be a rich resource for uncovering profound business insights. And, because machine-learning models get more efficient as they are "trained" with more data, machine learning and big data are highly complementary.

How will I know if my data is "big"?

Although many enterprises have yet to reach petabyte scale with respect to data volumes, it is possible that data has one of the other two defining characteristics of big data. And, if there is any single guarantee, it's that your data will grow over time–probably, exponentially. In that sense, all "big data" starts as "small data."

Why is the cloud the best platform for big data?

Cloud computing offers access to data storage, processing, and analytics on a more scalable, flexible, cost-effective, and even secure basis than can be achieved with an on-premises deployment. These characteristics are essential for customers when data volumes are growing exponentially–to make storage and processing resources available as needed, as well as to get value from that data. Furthermore, for those organizations that are just embarking on the journey toward doing big data analytics and machine learning, and that want to avoid the potential complexities of on-premises big data systems, the cloud offers a way to experiment with managed services (such as Google BigQuery and Google Cloud ML Engine) in a pay-as-you-go manner.

Learn more:

- Google BigQuery (https://cloud.google.com/bigquery/)

- Big Data Solutions (https://cloud.google.com/solutions/big-data/)

- [Why Google- Big Data and Analytics](https://cloud.google.com/why-google/big-data/) (https://cloud.google.com/why-google/big-data/)

Last updated September 27, 2019.