ASIAN BANKING SCHOOL

Let's Talk Digital Series #8

Big Data Analytics: from Laboratory to Life

Big Data Analytics: From Laboratory to Life

By Peter Kua Seng Choy

Big data includes a wide variety of large datasets that can be processed and analyzed using big data analytics or BDA. Among big data's most significant features are the 3Vs - Volume, Velocity and Variety.

VOLUME covers the large amounts of data that is produced and amassed by organizations. Over 40 billion videos and photos have been posted on Instagram since its launch; every year, Google handles over two trillion searches, and Netflix users consume over one billion hours of video weekly. BDA must be capable of handling and analyzing the enormous quantities of data that are generated by machines and people.

VELOCITY refers to the speed of data generated today. Data velocity can be best highlighted through social media examples. In a single minute, Twitter sees over half a million tweets, 4.5 million YouTube videos are watched, and Tinder users collectively swipe 1.4 million times. BDA must be capable of analyzing and processing such high-speed data.

VARIETY represents both unstructured and structured data. Unstructured data is unorganized data such as emails, handwritten documents, voicemails, ECG readings and audio recordings. Structured data on the other hand, represents strings, dates and numbers that can be arranged in database-like repositories. Over 80% of all data exists in an unstructured form. BDA needs to be competent in analyzing and processing these datasets, specifically unstructured data.

TRANSFORMING THE 3VS INTO VALUE

Unless value can be obtained to benefit the economy, government and society, big data and BDA are useless.

Organizations employ BDA to gather, process and analyze big data in a bid to gain competitive advantage. They utilize the insights from BDA to make decisions that will increase profits, drive down costs and maximize efficiencies.

Big data analytics can also help a government enhance its competency and productivity. And through BDA's positive effects on the government and the economy, society at large stands to gain.

BDA MADE POSSIBLE FOR EVERYONE TODAY

Big data analytics used to be confined to academic institutions and major R&D houses since only they had the resources to buy the computing power and storage required to perform advanced simulations and predictive analytics.

Today, cheap storage and powerful microprocessors are easily accessible to people who have the data and skills to implement advanced analytics.

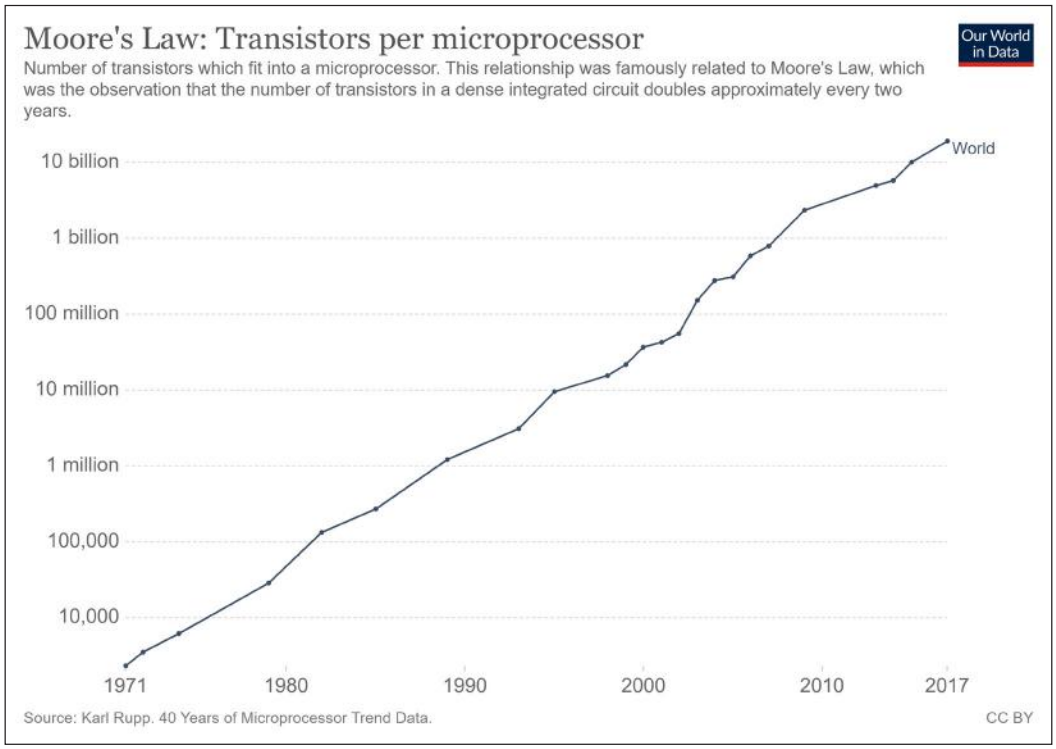

Figure 1: Computing power has doubled every 2 years

Ever since the first 2300-transistor Intel 4004 was introduced in 1971, integrated circuits have doubled their transistor count every two years. To put it simply, computing capabilities have doubled every 2 years. A 6-core Intel i7 microprocessor today contains over 2.6 billion transistors.

This incredible rise in computing power has accelerated BDA. It has enabled high-speed and complex predictive analytics. Advanced technologies like machine learning and deep neural networks require enormous amounts of computer processing power to learn and achieve a high degree of accuracy.

Computer memory prices have fallen drastically as well, which has helped the cause of big data analytics. Lots of storage and data are required for preparing advanced predictive analytic models like artificial neural networks. These models need access to plenty of historical data in order to make accurate forecasts.

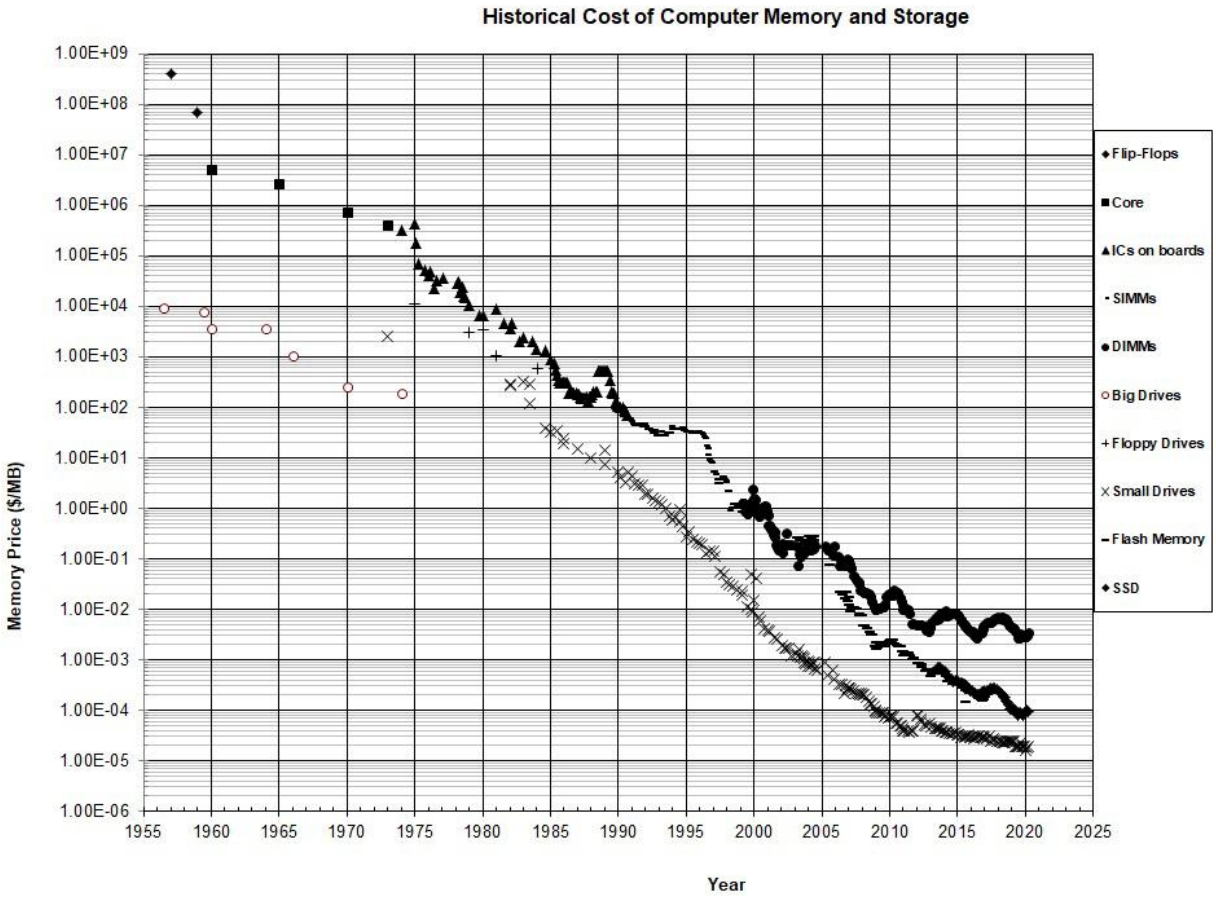

Figure 2: Price of computer storage has drastically fallen since its conception [(1)]

Back in the 80s, a 1 GB hard drive would have cost more than $500,000. Now, these drives can be bought for prices less than $0.03 per gigabyte.

Additionally, cloud computing has enabled individuals and organizations to rent the storage and computing power they require for analyzing and processing big data.

Cloud computing has lowered the entry barriers for BDA through its cost-effective pay-per-use business model. Cloud computing frees organizations from having to purchase and manage physical machines, databases and licenses.

1. www.jcmit.net

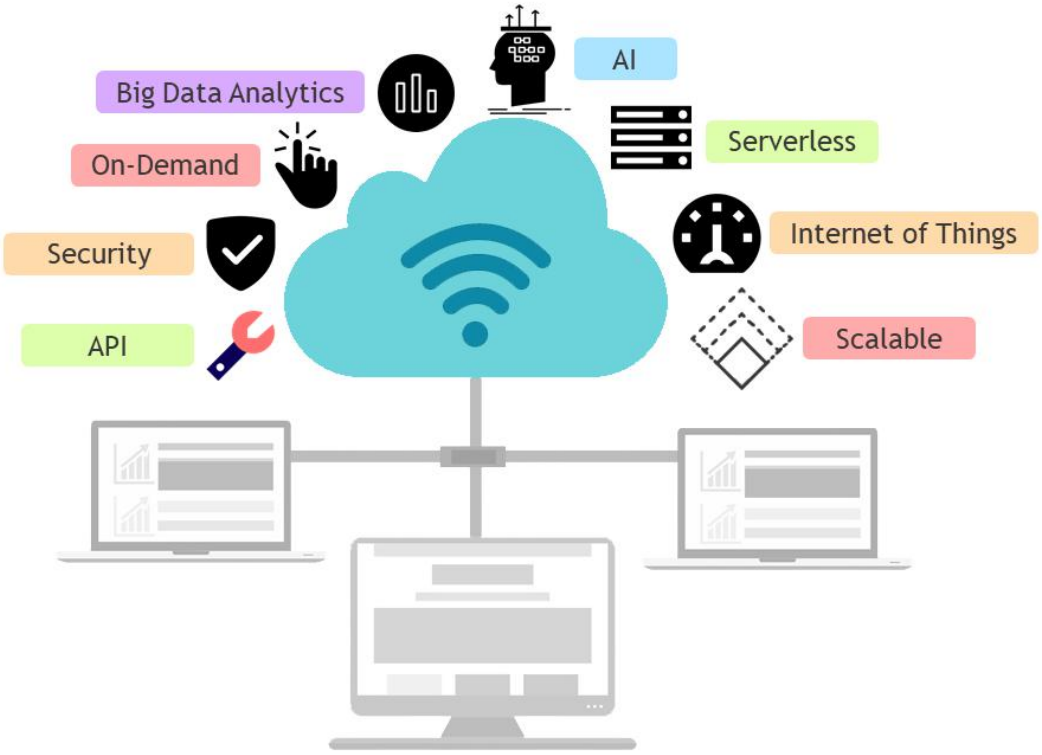

Figure 3: Cloud computing empowers on-demand big data analytics

It is now possible to lease an unlimited amount of computing power and storage. Huge amounts of data can be processed by cloud-based infrastructure. The training and validation of predictive models can also be done on the cloud. Once these intelligent models have been trained and corroborated, the cloud's virtual machines can be terminated.

ASIAN BANKING SCHOOL

This article is part of the Digital Banking Learning Series, 'Let's Talk Digital', an initiative by the ABS Center for Digital Banking. It is written by industry practitioners and are aimed at educating the general public on the intricacies of digital applications in banking and other related industries, including the latest insights and trends of Digital Banking.

As the industry's preferred partner in learning and development, ABS offers relevant training programmes that covers a comprehensive list of banking areas that are designed and developed in-house by our Specialist Training Consultancy Team or in collaboration with strategic learning partners that includes some of the top business schools in the world. It also provides specialised consulting services and tailored learning solutions to meet the specific needs of its clients.

For more information, visit our website at **www.asianbankingschool.com** or email us at **digitalbanking@asianbankingschool.com**