



# Analyzing Service Models in Cloud Computing and Networking through Big Data Lens

Nagaraju Ankathi <sup>1 \*</sup>, Dr. Ramesh Kumar Thallati <sup>2</sup>

1. Programmer Analyst IV, Kinecta Federal Credit Union, CA, US

ankathinagarajuresearch@gmail.com ,

2. Professor, Department of Computer Science, CMJ University, India

**Abstract:** Cloud computing provides small and medium-sized businesses with the assurance of significant data application. The MapReduce programming standard is used for Big Data processing. However, the application of the MapReduce standard generally requires online connected storage space and similar processing capabilities. The computing requirements of MapReduce programming are typically beyond the capabilities of small and medium-sized businesses. In this paper, we briefly discussed cloud computing service models and network infrastructure driven by big data.

**Keywords:** Big Data, Data Mining, Cloud Computing

----- X -----

## INTRODUCTION

It is no secret that "big data" is one of the most frequently used words today. Everyone is talking about it, and it is believed that scientific research, business, industry, government, and society, among others, will undergo a significant transformation with the impact of big data. Technically, the process of handling big data involves collection, storage, transportation, and exploitation. The collection, storage, and transportation stages are required for the ultimate goal of exploitation through data analytics, which is the core of big data processing.

From an information analytics perspective, we recognize that "big data" has been defined by the four V's - Volume, Velocity, Veracity, and Variety. It is believed that all or any of these criteria must be met for a problem to be classified as a Big Data problem. Volume refers to the size of the data, which may be too large to be handled by current algorithms and/or systems. Velocity means data is flowing at speeds faster than that which can be managed by conventional algorithms and systems. Sensors are rapidly exploring and communicating streams of data. We are approaching the world of measured self, where data that was not readily available previously now exists. Veracity indicates that despite the data being provided, the quality of data is still a major concern. That is, we cannot assume that with big data comes higher quality. In fact, with size comes quality issues that need to be tackled either at the data pre-processing stage or by the learning algorithm. Variety is the most powerful of all V's as it involves data of various formats and methods for a given object under consideration.

Each of the V's is not new. Artificial intelligence and data mining scientists have been addressing these issues for years. However, the advent of Internet-based services has challenged most of the traditional process-oriented businesses - they now need to become knowledge-based businesses driven by data rather than by process.

The purpose of this article is to share the authors' opinions about big data from their information analytics perspectives. The four authors bring various viewpoints with different research experiences and expertise, covering computational intelligence, machine learning, data mining and science, and interdisciplinary research. The authors represent academia and industry across four different continents. This diversity combines an exciting perspective with coverage on examining data analytics in the context of today's big data.

It is worth highlighting that this article does not aim to provide a comprehensive evaluation of the current state of big data analysis, nor to provide a future big data analysis research direction. The intention is to present the authors' personal viewpoints and offer their perspectives on the future based on their views. Therefore, there will be always minimal indicative argument or literary support, given the rapidly changing landscape and significant lag of academic research coverage. Indeed, many critical issues and relevant approaches are not explicitly covered in this article and are best left to research papers.

While all authors have contributed to the overall study, each author has focused on their specific specialties in the following discussions. Zhou covers artificial intelligence, while Chawla brings a data mining and data science perspective. Jin provides a view from computational intelligence and meta-heuristic global optimization, and Williams draws on a machine learning and data mining background applied as a practicing data scientist and consultant to industry globally.

Every year, we have been observing a significant improvement in our ability to collect knowledge from various sensing devices, systems, in multiple formats, from independent or attached applications. This big data has exceeded our ability to process, analyze, store, and visualize these datasets. Consider the data on the internet. The websites indexed by Google were around 100 million in 1998, but quickly reached one billion in 2000 and have now exceeded one trillion in 2008. In 2016, it is around 1.3 trillion.

## **CLOUD COMPUTING SERVICE MODELS**

Cloud computing offers a variety of release versions, including Platform as a Service (PaaS), Software as a Service (SaaS), Infrastructure as a Service (IaaS), and Hardware as a Service (HaaS). These services can provide benefits to businesses that they may not be able to achieve otherwise. Companies can also use cloud deployment as a test run before implementing a new technology or application.

PaaS provides businesses with a range of options for designing and developing applications. This includes application design and development tools, application testing, versioning, integration, deployment, hosting, state monitoring and other relevant development tools. PaaS can help businesses save costs through standardization and higher utilization of cloud-based computing across different applications. Other benefits of using PaaS include reducing risks by using pre-tested technologies, ensuring common services, improving software security, and reducing capacity requirements needed for new systems development. When it comes to big data, PaaS offers businesses a platform for creating and using customized applications required to analyze large amounts of unstructured data at a low cost and low risk in a secure environment.

SaaS provides businesses with applications that are stored and operated on virtual servers in the cloud.

Companies are not charged for hardware, only for the bandwidth and number of users required. The main advantage of SaaS is that businesses can shift the risks associated with software acquisition while moving from being capital-intensive to operational. Benefits of using SaaS include easier software management, automatic updates and patch management, software compatibility across the business, easier collaboration, and global accessibility. SaaS provides businesses analyzing big data with proven software solutions for data analysis. The difference between SaaS and PaaS in this case is that SaaS will not provide a customized solution, whereas PaaS will allow the business to develop a solution tailored to its needs.

In the IaaS model, a client company will pay for the use of hardware to support processing operations, including storage, servers, networking equipment, and more. IaaS is the cloud computing model that is receiving the most attention from the market, with an expectation of 25% of organizations planning to adopt a provider for IaaS. Services available to businesses through the IaaS model include disaster recovery, compute as a service, storage as a service, data center as a service, virtual desktop infrastructure, and cloud bursting, which provides peak load capacity for variable processes. Benefits of IaaS include increased financial flexibility, choice of services, business agility, cost-effective scalability, and improved security.

While not yet being used as widely as PaaS, SaaS, or IaaS, HaaS is a cloud service based on the time-sharing model used on minicomputers and mainframes in the 1960s and 1970s. Time-sharing evolved into the practice of managed services. In a managed services scenario, the managed provider would remotely monitor and administer equipment located at a client's site as contracted. A problem with managed services was the need for some MSPs to provide hardware on-site for clients, the cost of which needed to be incorporated into the MSP's fee. The HaaS model allows the customer to license the hardware directly from the provider, which reduces the associated costs. HaaS service providers include Google with its Chromebooks for Business, CharTec, and Equus.

## **TYPES OF CLOUDS**

There are three types of clouds: the public cloud, the private cloud, and the hybrid cloud. A public cloud is a pay-as-you-go service available to the general public. In this setup, a company does not own the technology resources and solutions but instead outsources them. A public cloud is considered an external cloud.

A private cloud is an internal data center of a company that is not accessible to the public but uses cloud infrastructure. In this setup, resources and solutions are owned by the company, with access available through the intranet. Since the technology is owned and managed by the company, this type of cloud is more expensive than a public cloud but is also more secure. A private cloud is an internal cloud, residing within the company's firewall and managed by the company.

When a company uses a hybrid cloud, it uses a public cloud for some tasks and a private cloud for others. In this model, the public cloud is used to accelerate tasks that cannot be easily run in the company's data center or on its private cloud. A hybrid cloud allows a company to keep critical, confidential data and information within its firewall while leveraging the public cloud for non-confidential data. The private cloud portion of the hybrid cloud is accessed by company personnel, both in the company and when

traveling, and is supported by the internal development team. The public cloud portion of the hybrid cloud is also accessed by the company's employees but is maintained by an outside provider. Each component of the hybrid cloud can connect to the other component. The type of cloud a company uses depends on its needs and resources. The public cloud is considered the least secure of the three types, with companies and resources accessible online through methods taken by the company. The communication protocols adopted by the provider are not necessarily secure, and the possibility of using secure or non-secure methods depends on the resources. The public cloud is also the least expensive of the cloud types, with cost savings in the areas of IT deployment, management, and maintenance.

The private cloud provides solutions to company employees through an intranet. If mobile employees can access the private cloud, access is usually through secure communication methods. All solutions and resources provided are tailored to the company's needs, and the company has complete control over the services and data. Due to the financial and personnel requirements to deploy, manage, and maintain the IT resources and solutions provided, the private cloud is the most expensive type of cloud. When a company uses a hybrid cloud, it has its own IT resources and services and will carry and deliver the resources and services internally. Non-critical services are outsourced and maintained on a public cloud. Typically, critical IT resources and services are mission-critical and are often confidential. Therefore, resources and services that need to be secure are hosted and protected on the private cloud, with the public cloud used for other services as a cost-saving measure.

As mentioned earlier, big data is often seen as critical for the success of the optimization of location systems. A lot of effort has been devoted to using data to improve the efficiency of meta-heuristic optimization formulations for solving complex problems in the presence of large amounts of uncertainties. It is believed that the boom in big data research may create new opportunities and impose new challenges to data-driven optimization. Addressing the following questions may be crucial to turning the challenges posed by big data into opportunities. Firstly, how can we effectively integrate modern learning and optimization techniques? Several advanced learning techniques, such as semi-supervised learning, incremental learning, active learning, and deep learning, have been developed over recent years. However, these techniques have rarely been used within optimization.

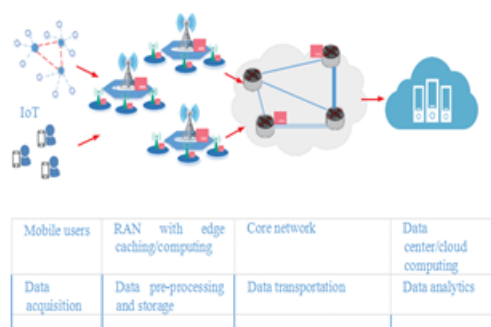
## **INDUSTRY, GOVERNMENT AND PEOPLE WITH BIG DATA**

Over the last two decades, there has been a significant increase in the collection of personal data by businesses and government agencies. Users have been incentivized to provide their personal data to these organizations in exchange for benefits. Companies like Google, Apple, and Facebook have access to a vast amount of personal data, including email, calendars, photos, and personal activities. This data can be used to better target services offered to users, using mathematical technologies to provide new insights and understanding.

However, the consumers who are driving this data collection are not the end-users of these services. Instead, the data is used by businesses and government agencies for their own purposes. This data is also shared with other organizations, either intentionally or unintentionally. In this context, the article discusses the impact of big data on society and examines how it is changing different industries and government practices. It provides a perspective on data analytics from the experiences of industry and government and

identifies areas where further research could make a significant impact.

## BIG DATA DRIVEN NETWORKING



**Figure 1 : Big data driven networking**

### B. Networking for Big Data

To effectively process and extract valuable insights from large amounts of data, it is crucial for 5G wireless networks to have the capabilities of handling high volume, high speed, and diverse data types. One way to achieve this is by increasing network capacity through techniques such as spectrum expansion, spectrum efficiency optimization, and network densification. This involves adding more spectrum resources, improving spectrum utilization, and enhancing spatial spectrum reuse. Additionally, data speed requires efficient data collection, preprocessing, and transmission. Furthermore, the wide range of data types must be supported through effective data handling and transportation.

Network slicing is an emerging solution for service-oriented networking that creates multiple virtual slices over the same physical infrastructure, extending from the access domain to the core network domain. Each slice operates independently and shares the same resource pool. A network slice is a collection of network resources for a given application or use case. It can be customized to meet the corresponding end-to-end service requirements, such as latency and reliability. Network slicing partitions heterogeneous network-wide resources for different slices to efficiently support diverse use cases. Effectively utilizing network resources while satisfying the needs of multiple applications or use cases is a challenging task.

Network function virtualization (NFV) can significantly facilitate network slicing. NFV enables network functions to be deployed in virtualized environments, improving network scalability and flexibility. In the core network, NFV can be used to compose a solution function chain (SFC) by virtualizing network functions such as traffic splitting, data collection, deep packet inspection, and firewall. An SFC can be created on-demand based on the service requirement of an application or use case. Multiple SFCs can be embedded over the same physical infrastructure for different applications/use cases. In addition, virtualization allows for dynamic scaling of network functions to adapt to network conditions and service requirements. For example, in the radio access network (RAN), virtual RAN instances can be created to support RAN slicing. Alternatively, resource allocation can be used to support slicing at a specific RAN component, such as a base station. By carefully allocating resources, the service needs of different slices can be met.

## CONCLUSION

To manage and process data efficiently, communication, computing, and storage resources are necessary at different points along the path from data acquisition to data centers. To address cost and diversity, personalized service-oriented end-to-end media is required, which includes relevant heterogeneous data and functions, tailored to the specific requirements of big data applications and use cases. In this paper, we provide a brief overview of cloud computing service models and big data-driven social networking.

---

## References

1. Smitha.T, Dr.V.Sundaram, "Comparative Study of Data Mining Algorithms for High Dimensional Data Evaluation"- published in International journal of Advancements in Engineering & Modern Technology (IJAET) ISSN2231-1963 173 in vol 4, issue2, sept 2012 PP 15-20.
2. P. Zikopoulos, C. Eaton, D. deRoos, T. Deutsch, and also G. Lapis. IBM Understanding Big Data: Analytics for Venture Course Hadoop as well as Streaming Data. McGraw-Hill Business, Integrated, 2011.
3. S. M. Weiss and N. Indurkha. Anticipating data mining: a functional overview. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1998.
4. Peddyreddy. Swathi, "Approaches And Objectives towards Financial Management" , International Journal of Advanced in Management, Technology and Engineering Sciences, Volume IV, Issue I, 2014
5. Peddyreddy. Swathi, "An Overview On The Types Of Capitalization" , International Journal of Advanced in Management, Technology and Engineering Sciences, Volume VI, Issue I, 2016
6. Peddyreddy. Swathi, "Architecture And Editions of Sql Server" , International Journal of Scientific Research in Computer Science, Engineering and Information Technology, Volume 2, Issue 4, May-June-2017
7. Peddyreddy. Swathi, "Scope of Financial Management and Functions of Finance", International Journal of Advanced in Management, Technology and Engineering Sciences, Volume III, Issue 1, 2013
8. Nagaraju Ankathi, Dr. Rajashekar Kummala,"Optimizing Big Data Workflows in Cloud Computing Environments", "International Journal of Scientific Research in Science, Engineering and Technology", Volume 3, Issue 3, 2017
9. Nagaraju Ankathi, Dr. Rajashekar Kummala, "Deployment Models and Web 2.0 Interfaces for Enhanced Business Solutions", "International Journal of Scientific Research in Science, Engineering and Technology", Volume 1, Issue 4, 2015
10. Nagaraju Ankathi, Dr. Kameshwar Rao, "Design Cycle and Deployment Considerations towards Efficient Implementation of Big Data Analytics in the Cloud", "International Journal of Scientific Research in Science and Technology", January-February-2016 [(2)1: 273-281 ]



11. Nagaraju Ankathi, "Optimizing Network Performance via Big Data Utilization", "International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering", Vol. 8, Issue 4, April 2019
12. Nagaraju Ankathi, "Critical Elements of Cloud Computing Infrastructure", "International Journal of Advanced Research in Arts, Science, Engineering & Management (IJARASEM)", Volume 1, Issue 2, November 2014
13. Nagaraju Ankathi, "The Journey of Cloud Computing Service Models: Evolution and Trends", "International Journal of Advanced Research in Arts, Science, Engineering & Management (IJARASEM)", Volume 5, Issue 5, September 2018
14. Keerthi Vuppula, Dr. Narsimha Reddy, "Facial emotion detection using machine learning algorithm K-nearest neighbor", "international journal for research & development in technology", Volume-13, Issue-2(Feb-20)
15. Keerthi Vuppula, "Internet of things based Smart Watch for Health Monitoring of Elderly People", "International Journal on Applications in Information and Communication Engineering", Volume 5, Issue 1, August 2019 , pp 82 –88
16. Keerthi Vuppula, "Design of Internet of things-based human-computer interface system", "International Journal on Applications in Basic and Applied Sciences", Volume 1, Issue 5, December 2013, pp 18-23.