Facilitating Role of Cloud Computing in Driving Big Data Emergence

Teah Yi Fan
School of Computing
Asia Pacific University of Technology & Innovation
Malaysia
tp061817@mail.apu.edu.my

Dr. Muhammad Ehsan Rana
School of Computing
Asia Pacific University of Technology & Innovation
Malaysia
muhd_ehsanrana@staffemail.apu.edu.my

***Abstract* — Big data emerges as an important technology that addresses the storage, processing and analytics aspects of massive data characterized by 5V's (volume, velocity, variety, veracity, value) which has grown exponentially beyond the handling capacity traditional data architectures. The most significant technologies include the parallel storage and processing framework which requires entirely new IT infrastructures to facilitate big data adoption. Cloud computing emerges as a successful paradigm in computing technology that shifted the business landscape of IT infrastructures towards service-oriented basis. Cloud service providers build IT infrastructures and technologies and offer them as services which can be accessed through internet to the consumers. This paper discusses on the facilitating role of cloud computing in the field of big data analytics. Cloud deployment models concerning the architectural aspect and the current trend of adoption are introduced. The fundamental cloud services models concerning the infrastructural and technological provisioning are introduced while the emerging cloud services models related to big data are discussed with examples of technology platforms offered by the big cloud service providers – Amazon, Google, Microsoft and Cloudera. The main advantages of cloud adoption in terms of availability and scalability for big data are reiterated. Lastly, the challenges concerning cloud security, data privacy and data governance of consuming and adopting big data in the cloud are highlighted.**

Keywords—Big data, cloud computing, Big Data as a Service (BDaaS), Database as a Service (DBaaS), Analytics as a Service (AaaS)

I. INTRODUCTION

Moving into the era of Big Data, the advancement in Internet of Things (IoT) and popularisation of Web 2.0 applications created nascent footprints in the organisations and corporates, expanding the data into multiple sources ranging from sensor-generated data, emails, social media and multimedia streaming output [1], of which the speed of data generation is beyond the processing power of traditional computing system and the structures no longer fit into the storage system of traditional database systems [2].

Big data emerges as the parallel storage and processing framework, a technological breakthrough in computing technology, achieves efficient storage and effective processing of massive unstructural data [3], provides feasibility to perform highly complex analysis to generate meaningful business insights [4]. The scientific research is among the communities that achieves significant progression where big data frameworks has made processing and analysing an estimated amount of 100 to 100000 Terabytes of genomics data generated from the Human Genome Project research [5], turning them towards the paradigm shift into personalised treatment and therapy in the healthcare industry[6]. On the other hand, the ability to store and process IoT sensor-generated data enable the manufacturing industry are gaining visibility and transparency on the machinery performance, which changes the landscape of machinary maintenance [7] from preventive measures to predictive measures [8], leading to significant reduce of manufacturing downtime, shifting the paradigm of manufacturing into smart manufacturing in the industry [9]. Furthermore, the ability to collect, store and process mobile technology-generated data enables the harnessing of insights into people's behaviour and activities to energy consumptions environmental conditions [10] and more brings constructive impact to the city planning communities to develop disruptive future concept on the development of smart cities [11].

The big data technology has brought pragmatic disruption and opens up opportunities in many areas, facilitating the large organisations and corporations to gain additional competitive advantages from the big data adoption [2]. Despite the promising impact of big data can bring to business, the adoption requires new infrastructure to house the parallel computing hardwares and technology frameworks involving huge monetary investment, which can be a burden to small and medium enterprises as well as the startups [12].

Cloud computing emerges as resource virtualization, an infrastructural breakthrough in computing technology, brings the sophisticated infrastructures required for big data adoption to the consumers as services through simple internet mechanism [3]. Moreover, the business model of cloud computing is built upon pay-as-you-go payment plan [15], not limited to renting the required resources but also the flexibility to scale-up and scale-down through simple on-demand self service mechanism[13], are among the main drivers of cloud computing adoption for big data [16]. It addresses the bottleneck of big data adoption by breaking down the infrastructural barrier in an economical feasible way, making cloud computing a favorable option for all scale of business [14].

The following sections aim to address various aspects of considerations in cloud adoption, which covers the cloud deployment models and cloud service models and their suitability for big data applications. Finally, the advantages as well as the risk and challenges are discussed to uncover the why and why-not for big data in cloud adoption.

II. Cloud Deployment Model

Cloud deployment model refers to construction, management and the operation of a cloud infrastructure. It defines the resource sharing and relationship of consumers within the cloud environment, commonly referred to as tenancy. Generally, there are 4 common deployment models – public cloud, private cloud, community cloud and hybrid cloud; and other newly developed concepts such as multi-cloud and rain cloud. A comparison table of requirements and specifications is summarized in Table I. The suitability of cloud deployment model depends on the current IT setup such as computing and networking environment of the organisation, the additional requirements required by the organisation, sensitivity of security level of applications in different business function, and most importantly, business goal.

A. Public cloud

The resources on *public cloud* are available to the public in a multitenant manner while the infrastructure resides with the cloud service provider, managed and administered by the cloud service providers. Having said that, the end users of public cloud are at ease of the operation and maintenance of which the responsibility falls on the cloud service provider while being able to access and utilise the complex high level resources from the current devices. On the other hand, this also imply that the end users have to rely heavily on cloud service provider in terms of the security and privacy through the logical isolation solution maintained by the cloud service providers; and business continuity in the case of the disruption or outages. The major incidence of service outage by Salesforce cloud service in 2016 [17] and 2019 [18] are the classic example of the limitation in private cloud. Nevertheless, with increasing stabilisation of technology, the global public cloud services market in 2020 registered USD 312 billion [19], which is equivalent to 53% surpluses from USD 203.4 billion, the value projected in 2017 [20]. Among the main drivers of adoption, public cloud is the most cost feasible model especially suited for individuals and startups whose business continuation outweigh data privacy.

B. Private cloud

The *private cloud* is owned by particular organization and intended for the consumption within the organisation. The private cloud can be hosted as on-premise or off-premise infrastructures where both types of hosting mechanisms differ significantly in terms of cost – both infrastructure and human resources. Hosting off-premise private cloud transfers the technological and infrastructural establishment costs to the cloud service provider, which means reducing the capital expenditure (CAPEX) of organisation shifting it to operational expenditures (OPEX). This is the gap in the cloud adoption market that cloud service providers penetrates by offering virtual private cloud as the solution to large and established organisations. The private cloud is a package of policy-based isolation, ensuring the cloud resources are prohibited from other tenants, and are provisioned to a specific business user. According to IDC, the global spending on private cloud infrastructure seizes 0.6% year over year increment in 2020, a significant drop compared to that in 2017 which is as high as 13.1%. On top of that, the spending on on-premise private cloud accounts for 63.2% of total spending on private cloud infrastructure, demonstrating a slight preference over on-premise model [21]. The gap is yet to be identified but the compromise on the limited control of the cloud is unarguably one of the profound factor of off-premised private cloud adoption among the large enterprises, whose business establishment are built upon proprietorship of data.

C. Community cloud

The *community cloud* deployment model is a multitenant cloud environment with restriction of users limited to the same domain business or expertise such as government, educational organisations, healthcare institutions, scientific research community etc. The typical use cases revolves around business joint ventures or the need for centralised system. For instance, the US Government employs the community cloud deployment model [22]. Similar to the private cloud, the infrastructures can be hosted on-premise or off-premise, where the cost of establishment are shared among the consumers within the community.

D. Hybrid cloud

The *hybrid cloud* comprises the combination of two or more types of cloud between the public, private and community cloud model. The key feature of hybrid cloud include increased interoperability for resource provisioning between public and private clouds, allowing companies to transition smoothly between traditional and cloud infrastructures. On the other hand, the portability of resources from private to public cloud makes hybrid cloud a suitable for big data-related use cases, where surge of data flows is highly expected in the data-intensive operations [23]. For instance, NASA employs the combination of private and public cloud as hybrid cloud deployment model during the development process of Nebula, where large datasets are shared among stakeholders through public cloud whereas the research and development processes resides in the private cloud [24].

E. Other cloud deployment models

Today, riding on the emerging big data trend and maturity of the cloud computing market, *multicloud* approaches becomes a favourable option for organisations to begin cloud adoption depending on the business demand. Multi-cloud cater the best possible cloud model to different business functions based on workloads and security requirements. Furthermore, no one cloud service providers is perfect in the applications for all business functions. Multi-cloud model allows businesses to select the vendors which provides the best services catered to business functions, thus helping organisations to reduce vendor lock-in. Moreover, multi-cloud adoption has emerged as a measure in helping organisations to meet compliance and regulatory requirements. For instance, Google's multi-cloud proposal has won the cybersecurity contract from the Pentagon of USA in 2020 [22].

TABLE I. COMPARISONS OF REQUIREMENT FOR ADOPTION OF CLOUD DEPLOYMENT MODELS

	Public	**Private**	**Community**	**Hybrid**
Infrastructural and Technological setup	Minimal	Requires specialised cloud proficiency	Requires specialised cloud proficiency	Requires specialised cloud proficiency
Access and Control to Data	Low	High	High	High
Scalability and flexibility	High	High	Fixed capacity	High
Cost of adoption	The most cost effective model	Cost-intensive; the most expensive model	Cost is shared among community members	Depends on model
Requirements on In-house On-premise	No	Depends on on-premise or off-premise	Depends on on-premise or off-premise	Depends on the model combinantion

III. CLOUD SERVICE MODELS

Cloud service model refers to the resources which are provisioned to the end user. Generally, there are 7 types of resources in the cloud – network, storage, compute system, operating system, programming framework, database and applications; and 3 basic levels of service models i.e. infrastructure (IaaS), platform (PaaS) and software (SaaS). Regardless of the service models, three of the resources, primary resources – network, storage and compute system are exquisitely owned by the cloud service providers while the consumers are given control on the secondary resources– operating system, programming framework, database and applications, depending on the level of service model.

A. Fundamental service models – IaaS, PaaS and SaaS

Infrastructure as a service (IaaS) provides operating system, programming framework, database and applications which runs on the virtual machines, processing power and storage owned by the cloud service provider to the consumers. For instances, Amazon EC2, Google Compute Engine, OpenStack and CloudStack are examples of leading IaaS that provision computing infrastructure as services on pay-as-you-go payment schemes.

Platform as a service (PaaS) provides only application resource to the consumers, which means the users have less control on other secondary resources. As such, the server maintenance and software updates work are transferred to the cloud service provider. To elaborate further, the consumers are given options of platform (operating system, programming framework and database) for development work such as WISA stands for Windows, IIS, SQL and ASP.net; LAMP stands for Linux, Apache, mySQL, php; are examples of 'stack' that specifies the operating system, programming framework and database environments for users to develop, build and deploy applications on the cloud. Microsoft Azure, Google App Engine and CloudFoundry are among the most renowned PaaS.

Software as a service (SaaS) provides the least control of resources to the consumer, where only configurations are allowed at the consumers' end. In other words, this benefits users at ease of accessing the software online even without installation, thus allowing the access to the service from any devices. For example, Salesforce.com, Google Docs and Dropbox provide software or application which can be accessed by users directly on the cloud.

B. Big data-related cloud service models

The increasing demand of big data analytics in the business world, drives the extension of the cloud services into more diverse forms of service brokerage, aiming to add values to the current cloud service models. Some examples of cloud service brokerage related big data analytics are Big Data as a Service (BDaaS) [25], Database as a Service (DBaaS) [26] and Analytics as a Service (AaaS) [16]. A comparison table of these services offered by the main players in the cloud computing industry are summarized in Table II. More recently, a newer concept of cloud service model termed as Cloud-based Big Data Analytics as a serviced (CLAaaS) services has been discussed in recent publications [27]. These cloud service brokerage models targets on facilitating the infrastructural and technological aspects of big data analytics requirements – processing power and scalable storage. Moreover, certain brokerage models also provide the analytical aspects such as data mining, predictive analytics and artificial intelligence which focuses on business intelligence. Taken together, the cloud service providers are gaining increasing popularity on these brokerage models for provisioning the appropriate operational resources for management of big data, allowing organisations to focus on analytic applications that translates big data into actionable insights.

Big Data as a Service (BDaaS) focuses in providing big data processing framework and processing engines as additional services built upon the three basic cloud service models. This service simplifies the process of setting up big data environment, provides flexibility and scalability to the organisations which are newly venturing into big data [25]. Undoubtedly, the key players in the cloud computing industry – Amazon, Google, Microsoft and Cloudera are not excluded from this business landscape, where the competition of product offerings revolves around the renowned big data processing frameworks such as Hadoop distributed processing framework and Map Reduce as well as Apache Spark data processing and analytics engine. Serving these big data framework on the cloud allows organisations especially those in small and medium enterprises which lacks in-house data centers to gain access to these big data framework from the cloud on pay-as-you-go basis and benefits from what big data could offer in their businesses.

Database as a Service (DBaaS) adds data storage services as an additional layer to the three predominant cloud service model. This service complement businesses over traditional on-premise database management system which requires huge human and infrastructure resources [26]. The payment scheme of DBaaS is subscription-based and charged on per-usage basis. The DBaaS service model is unique. It involves technological collaboration between the relational data storage providers – MySQL, Oracle database, Microsoft SQL Server etc.; NoSQL data storage providers – MongoDB, Cassandra, DynamoDB etc. and cloud service providers (Amazon, Google, Microsoft, IBM etc.). In other words, organisations, as consumers, have options to choose database management system and cloud platforms from different providers. Nevertheless, the options are bound to the compatibility between platforms. However, the fact that DBaaS provides promising cost savings, management efficiency, operational and maintenance feasibility to the organisations remains debatable as long term investment.

Analytics as a Service (AaaS) adds data analytic services as embedded features to the existing cloud service models, serving as cost-effective analytical tools over traditional licensed business intelligence solutions [16]. Among the popular business intelligence solutions include descriptive data analytics, predictive analytics and interactive visualisations. The availability of analytic tools as a service allows companies to move from web-based to hybrid in the process of adopt big data analytics in business processes. Advancing from batch analytics, the AaaS community is moving rapidly into the line of offering stream processing as services because real-time processing is becoming the next wave in generating business-oriented values especially in the highly competitive e-commerce platforms. As of today, the stream processing services on the cloud requires compatibility between the storage and processing platforms due to technical complexity.

IV. ADVANTAGES OF CLOUD

The integration between big data and cloud computing can bring pragmatic implications. This section discusses the availability and scalability features of cloud that enable deployment of data-intensive applications to drive business analytics.

Cloud computing ensures scalability of resources required for business agility. The high scalability of cloud fulfills the storage, processing, memory and network bandwidth requirements for big data analytics, addresses the main challenge working with big data growing exponentially at volume, velocity and variability dimensions. The cloud enables this feasibility through the elastic load balancing and auto-scaling features of the cloud to deliver on-demand resource provisioning based on the usage requirement. In turns, the scalability characteristic of cloud allows organisation to utilise the resources optimally without up-front capital expenditure as well as simplifies the infrastructural management on big data. Taken together, scalability ensures appropriate resource provisioning based on project based requirement, helps organisations to reduce IT administration cost, leads to improved agility in business.

TABLE II. SUMMARY OF BIG DATA-RELATED CLOUD SERVICE MODELS

Big Data Application	Cloud Computing Service Model	Cloud Deployment Model	Cloud Service Provider	Platform	Framework/Technology/Infrastructure
DBaaS Relational Storage	SaaS	Public, Private, or Hybrid	Amazon	Amazon RDS Amazon RedShift	MySQL, Oracle Database, SQL Server, PostgreSQL
			Google	Google Cloud SQL	MySQL, PostgreSQL, and SQL Server.
			Microsoft	Azure SQL Database	MySQL
			Cloudera	CDP Operational Database	MySQL, Oracle Database, PostgreSQL
DBaaS Non-relational Storage			Amazon	Amazon DynamoDB	Dynamo DB
			Google	Google Cloud Firestore Google Cloud BigTable	Mongo DB HBase
			Microsoft	Azure CosmosDB	CosmosDB
			Cloudera	Cloudera Accumulo	Apache Accumulo
BDaaS Big Data Processing Framework	PaaS		Amazon	Amazon EMR	Hadoop HDFS and MapReduce, Apache Spark, Apache Hive, Apache HBase, Apache Flink, Apache Hudi, and Presto.
			Google	Google Cloud Dataproc Google Cloud Dataflow	Apache Hadoop, Apache Spark
			Microsoft	Microsoft Azure HDInsight	Apache Hadoop, Spark, Hive, Kafka etc.
			Cloudera	Cloudera Distribution of Hadoop (CDH) Cloudera Data Hub	Hadoop HDFS and Map Reduce, YARN, Apache Spark
AaaS Batch Analytics	SaaS		Amazon	Amazon Athena	ETL, interactive query
			Google	Google Cloud BigQuery Google Data Studio	SQL query, machine learning, geospatial analysis interactive query
	PaaS & SaaS		Microsoft	Azure Synapse Analytics Azure Machine Learning Azure Analysis Services	interactive query, machine learning
	SaaS		Cloudera	DXC Analytics Platform	interactive query
AaaS Stream Analytics	SaaS		Amazon	Amazon Kinesis	Apache Flink
			Google	Google DataFlow	In-house proprietory
			Microsoft	Azure Stream analytics	In-house proprietory
			Cloudera	Cloudera DataFlow	Apache Flink

Cloud computing ensures 24-by-7-by-365 availability required for business continuity. High availibility does not only ensure the fulfillment of computing resources but also the availability of data, the most essential element for big data analytics. The cloud system eliminates the dependency on a specific end-point device to access to the massive data and its processing framework, enabling services to be accessed through thin and thick client devices from any region through standard internet protocols. Consequently, this convenience promotes collaboration between within business orgasations. Furthermore, cloud system is backed by fault tolerance which increases the resilience and robustness of large scale processing of big data. Moreover, high availability also ensures that data, the real value of big data analytics, is accessible through cloud-based backup. The ability to access to data and run data processing on incidences of fault reduces the impact of downtime thus ensuring business continuity.

IV. Risks and Challenges

As integration of big data into cloud computing are increasingly recognised, it attracts increasing discussion and research in addressing the new challenges. The characteristics of cloud known to promise multiple advantages to businesses can be a two way sword especially in the aspect of security, privacy [28], [29] and management issues [30] arousing from the adoption of cloud for big data analytics [31].

Security of the cloud environment is one of the major aspect in terms of risk and challenges. The term security in cloud concerns additional elements compared to the traditional IT. First of all, the multitenancy nature of cloud environment which allows resources sharing among multiple consumers can increase the security risks to data confidentiality, integrity and availability while securing the network in cloud is beyond the traditional firewall where it requires additional measures like VPN and IDPS to prevent intrusion of the virtual layer as well as trespassing at the data center . Furthermore, standardisation of components in the cloud platforms intended for better performance can be one of the factors that amplify the velocity of attack due to the homogeneity in the cloud environment. In this regard, multitenant public cloud is more vulnerable than that in the private cloud, thus the service level aggrement (SLA) of the public cloud are often stringent as compared to that of the private cloud.

Privacy of the big data contained in cloud prevails as the adoption of cloud for big data analytics further drives the demand for cloud storage. Integration of big data analytics on cloud, requires sensitive data such as personal identifiable information (PII) to be transferred onto the cloud for processing, which exposes the data to risks of data breaches and data loss. Data privacy is bound by laws and regulations in preservation of sensitive user information. One of the most debatable topic concerning data privacy due to cloud adoption is related to employee communication monitoring where it is prohibited in some countries while permitted in certain countries under special circumstances. Notably, any data stored in the data centers resides in US are freely accessible by the government under the National integrity policy. Branching from data privacy issue, data ownership can be a major issue. Considering organisations data migrated to cloud storage, the ownership is prominent. However, considering a circumstance when an organisation employs strategy for collecting and generating data in the cloud environment, data ownership can be debatable, giving rise to further governence issues.

Data governance of big data in cloud concerning about the monetisation of data for business purposes is bound by data-related regulations. Integration of big data and cloud computing extends opportunities for organisations to streams of overwhelming and useful big data from external sources which carries huge monetisation values by selling them raw or translation into insights to the thrid parties. The key challenge is not limited to exposing sensitive information about uninformed party whom data have been used or misused but also the exercise of the monetisation are also exposed to regulatory challenges.

V. Contributions

The world is overwhelmed big data which can be harnessed to gain competitive edge in all scales of businesses. However, the implementation requires huge infrastructures. The main contributions of this paper is to highlight the facilitating role of cloud computing in accelerating small to medium scale businesses in terms of big data adoption. As far as investment into Big Data infrastructures constantly raises as a bottleneck to big data adoption, cloud computing emerges as a successful paradigm shift in computing technology, bringing advanced computing infrastructure available to anyone through virtualisation. Furthermore, the authors provide a bird's eye view of up-to-date cloud deployment and service models provided by renowned cloud service providers – Amazon, Google, Microsoft and Cloudera.

VI. Conclusions and Future Research

In this research, the authors have discussed the suitability of cloud deployment model suited to different business nature and highlights the multicloud deployment as the current trend of adoption. The authors emphasize the advantages over risk and challenges for big data in cloud adoption especially in the small and medium scale enterprises. However, the authors reckoned several areas of improvement for future research. While the availability of cloud services are overwhelmed, there is a lack of standardisation of service quality among and across the cloud providers. The interoperability and security in the adoption of multi-cloud environment may warrant further development of open standards.

Acknowledgment

The work described in this paper was supported by the Asia Pacific University of Technology & Innovation (APU) Malaysia. The authors of this paper are also thankful to the University of Bahrain for providing the opportunity to present this paper free of cost under IEEE platform.

References

[1] A. Sestino, M. I. Prete, L. Piper, and G. Guido, “Internet of Things and Big Data as enablers for business digitalization strategies,” *Technovation*, vol. 98, no. July, p. 102173, 2020, doi: 10.1016/j.technovation.2020.102173.

[2] A. Kobusińska, C. Leung, C. H. Hsu, S. Raghavendra, and V. Chang, “Emerging trends, issues and challenges in Internet of Things, Big Data and cloud computing,” *Futur. Gener. Comput. Syst.*, vol. 87, pp. 416–419, 2018, doi: 10.1016/j.future.2018.05.021.

[3] S. N. Khezr and N. J. Navimipour, “MapReduce and Its Applications, Challenges, and Architecture: a Comprehensive Review and Directions for Future Research,” *J. Grid Comput.*, vol. 15, no. 3, pp. 295–321, 2017, doi: 10.1007/s10723-017-9408-0.

[4] R. H. Tsaih and C. C. Hsu, “Artificial intelligence in smart tourism: A conceptual framework,” *Proc. Int. Conf. Electron. Bus.*, vol. 2018-Decem, pp. 124–133, 2018.

[5] K. A. Shakil and M. Alam, “Cloud computing in bioinformatics and big data analytics: Current status and future research,” *Adv. Intell. Syst. Comput.*, vol. 654, pp. 629–640, 2018, doi: 10.1007/978-981-10-6620-7_60.

[6] B. Seth, S. Dalal, and R. Kumar, *Securing bioinformatics cloud for big data: Budding buzzword or a glance of the future*, vol. 823. Springer International Publishing, 2019.

[7] D. Li, H. Tang, S. Wang, and C. Liu, “A big data enabled load-balancing control for smart manufacturing of Industry 4.0,” *Cluster Comput.*, vol. 20, no. 2, pp. 1855–1864, 2017, doi: 10.1007/s10586-017-0852-1.

[8] Q. Qi and F. Tao, “Digital Twin and Big Data Towards Smart Manufacturing and Industry 4.0: 360 Degree Comparison,” *IEEE Access*, vol. 6, pp. 3585–3593, 2018, doi: 10.1109/ACCESS.2018.2793265.

[9] F. Tao, Q. Qi, A. Liu, and A. Kusiak, “Data-driven smart manufacturing,” *J. Manuf. Syst.*, vol. 48, pp. 157–169, 2018, doi: 10.1016/j.jmsy.2018.01.006.

[10] K. Soomro, M. N. M. Bhutta, Z. Khan, and M. A. Tahir, “Smart city big data analytics: An advanced review,” *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.*, vol. 9, no. 5, pp. 1–25, 2019, doi: 10.1002/widm.1319.

[11] M. Khan, M. Babar, S. H. Ahmed, S. C. Shah, and K. Han, “Smart city designing and planning based on big data analytics,” *Sustain. Cities Soc.*, vol. 35, pp. 271–279, 2017, doi: 10.1016/j.scs.2017.07.012.

[12] Z. Ageed, M. R. Mahmood, M. M. A. Sadeeq, M. B. Abdulrazzaq, and H. Dino, “Cloud Computing Resources Impacts on Heavy-Load Parallel Processing Cloud Computing Resources Impacts on Heavy-Load Parallel Processing Approaches,” *IOSR J. Comput. Eng.*, vol. 22, no. 3, pp. 30–41, 2020, doi: 10.9790/0661-2203043041.

[13] I. A. T. Hashem, I. Yaqoob, N. B. Anuar, S. Mokhtar, A. Gani, and S. Ullah Khan, “The rise of ‘big data’ on cloud computing: Review and open research issues,” *Inf. Syst.*, vol. 47, pp. 98–115, 2015, doi: 10.1016/j.is.2014.07.006.

[14] S. A. El-Seoud, H. F. El-Sofany, M. Abdelfattah, and R. Mohamed, “Big data and cloud computing: Trends and challenges,” *Int. J. Interact. Mob. Technol.*, vol. 11, no. 2, pp. 34–52, 2017, doi: 10.3991/ijim.v11i2.6561.

[15] G. Skourletopoulos *et al.*, “Big Data and Cloud Computing: A Survey of the State-of-the-Art and Research Challenges,” pp. 23–41, 2017, doi: 10.1007/978-3-319-45145-9_2.

[16] A. Ara and A. Ara, “Cloud for Big Data Analytics Trends,” *IOSR J. Comput. Eng.*, vol. 18, no. 05, pp. 01–06, 2016, doi: 10.9790/0661-1805040106.

[17] J. Davis, “Salesforce Outage: Can Customers Trust The Cloud?,” 2016. https://www.informationweek.com/platform-as-a-service/salesforce-outage-can-customers-trust-the-cloud- (accessed Aug. 21, 2021).

[18] J. Novet, “Salesforce says a ‘major issue’ with cloud service results in downtime,” 2019. https://www.cnbc.com/2019/05/17/salesforce-says-a-major-issue-with-cloud-service-results-in-downtime.html (accessed Aug. 21, 2021).

[19] “Worldwide Public Cloud Services Market Totaled $312 Billion in 2020 with Amazon Web Services and Microsoft Vying for the Top Position Overall, According to IDC,” *IDC*, 2021. https://www.idc.com/getdoc.jsp?containerId=prUS47685521 (accessed Aug. 21, 2021).

[20] “Worldwide Public Cloud Services Spending Forecast to Reach $122.5 Billion in 2017, According to IDC | Business Wire,” *Businesswire*, 2017. https://www.businesswire.com/news/home/20170220005049/en/Worldwide-Public-Cloud-Services-Spending-Forecast-to-Reach-122.5-Billion-in-2017-According-to-IDC (accessed Aug. 21, 2021).

[21] “Public Cloud IT Infrastructure Revenue Growth Remained Strong in Third Quarter of 2020, According to IDC,” *IDC*, 2020. https://www.idc.com/getdoc.jsp?containerId=prUS47279621 (accessed Aug. 29, 2021).

[22] J. Tsidulko, “Google Cloud Wins Multimillion-Dollar Pentagon Cybersecurity Contract,” 2020. https://www.crn.com/news/cloud/google-cloud-wins-multimillion-dollar-pentagon-cybersecurity-contract?itc=refresh (accessed Aug. 21, 2021).

[23] J. McKendrick, “NASA’s Nebula: a stellar example of private clouds in government | ZDNet,” 2010. https://www.zdnet.com/article/nasas-nebula-a-stellar-example-of-private-clouds-in-government/ (accessed Aug. 21, 2021).

[24] V. Phillips, “Nebula Cloud Computing Platform | NASA,” 2017. https://www.nasa.gov/open/nebula.html (accessed Aug. 21, 2021).

[25] M. S. Al-hakeem, “A Proposed Big Data as a Service (BDaaS) Model,” *Int. J. Comput. Sci. Eng.*, vol. 4, no. 11, pp. 1–8, 2016.

[26] B. Iancu and T. M. Georgescu, “Saving Large Semantic Data in Cloud: A Survey of the Main DBaaS Solutions,” *Inform. Econ.*, vol. 22, no. 1/2018, pp. 5–16, 2018, doi: 10.12948/issn14531305/22.1.2018.01.

[27] F. Zulkernine, P. Martin, Y. Zou, M. Bauer, F. Gwadry-Sridhar, and A. Aboulnaga, “Towards cloud-based analytics-as-a-service (CLAaaS) for big data analytics in the cloud,” *Proc. - 2013 IEEE Int. Congr. Big Data, BigData 2013*, no. June, pp. 62–69, 2013, doi: 10.1109/BigData.Congress.2013.18.

[28] B. M. Balachandran and S. Prasad, “Challenges and Benefits of Deploying Big Data Analytics in the Cloud for Business Intelligence,” *Procedia Comput. Sci.*, vol. 112, pp. 1112–1122, 2017, doi: 10.1016/j.procs.2017.08.138.

[29] S. Sahmim and H. Gharsellaoui, “Privacy and Security in Internet-based Computing: Cloud Computing, Internet of Things, Cloud of Things: A review,” *Procedia Comput. Sci.*, vol. 112, pp. 1516–1522, 2017, doi: 10.1016/j.procs.2017.08.050.

[30] T. Choudhury, A. Gupta, S. Pradhan, P. Kumar, and Y. S. Rathore, “Privacy and Security of Cloud-Based Internet of Things (IoT),” *Proc. - 2017 Int. Conf. Comput. Intell. Networks, CINE 2017*, pp. 40–45, 2018, doi: 10.1109/CINE.2017.28.

[31] S. Amir, A. Ahmad, and M. R. Ehsan, “SMEs on Cloud: Challenges and Opportunities,” *Natl. Congr. Electr. Eng. Comput. Inf. Technol.*, pp. 1–6, 2013.