ASSESSMENT OF CLOUD STORAGE AND THE RELATED COMPONENTS: THE MERITS AND DEMERITS

BY

Joel M. PETERS Ph.D
Department of Library and Information Science
Faculty of Education
University of Rochester
Rochester
New York City

And

Theresa Nse UDO-OKON, Ph.D
Department of Educational Technology and Library Science
University of Uyo, Uyo.

ABSTRACT

The study assessed cloud storage, the related components, the merits and demerits. Cloud storage is a cloud computing model in which data is stored on remote servers and accessed over the internet or through a dedicated private network connection. Cloud storage has evolved into a crucial part of how people and businesses store, access, and manage their data. It has a number of benefits, such as scalability remote accessibility and cost-effectiveness, and remote accessibility which have completely changed how data is managed. The study concluded that the evaluation of cloud storage has highlighted both its many merits and demerit, providing a comprehensive view of this important technology. Cloud storage allows users to save important data or media files on remote, third-party servers. Users are empowered to access data from anywhere thanks to its remote accessibility, which improves flexibility and cooperation. Also known as utility storage, cloud storage is maintained and operated by a cloud-based service provider. From greater accessibility to data backup, cloud storage offers a host of benefits. The most notable are the large storage capacity and minimal costs. One of the recommendations made was that robust security measures, including encryption, access controls, and regular security audits, should be implemented to protect sensitive data in the cloud.

Keywords: Cloud Storage, Related Components, Merits and Demerits

INTRODUCTION

Understanding the benefits and downsides of cloud storage in the context of contemporary data management is essential. Cloud storage has evolved into a crucial part of how people and businesses store, access, and manage their data. It has a number of benefits, such scalability and accessibility, but it also has some drawbacks, like financial implications and worries about data security. As businesses increasingly depend on cloud-based solutions to store, manage, and access their data, understanding the advantages and disadvantages of cloud storage is crucial. Scalability, cost-effectiveness, and remote accessibility are just a few benefits of cloud storage that have completely changed how data is managed. However, it also has its own set of difficulties, including worries about data security, privacy, and possible unintended expenses. This evaluation intends to explore both the benefits and drawbacks of cloud storage, offering a thorough grasp of how it affects contemporary data management practices. It is especially essential in the current digital environment, where data has grown to be a priceless resource for both people and businesses. Due to their capacity to offer adaptable, scalable, and affordable data storage alternatives, cloud storage systems have become increasingly popular. To make wise choices about data management and security, it is necessary to weigh the benefits and drawbacks of cloud storage (Marz, & Warren, 2015).

In order to provide quicker innovation, adaptable resources, and scale economies, cloud computing is the distribution of computer services via the internet ("the cloud"), including servers, storage, databases, networking, software, analytics, and intelligence. Typically, you only pay for the cloud services that you really use, which lowers operational expenses, improves infrastructure management, and enables you to scale as your company's needs change (Microsoft.com, 2023). Cloud storage, as defined by Neha & Chiradeep (2021), is a data deposit paradigm in which digital data, including documents, images, videos, and other types of media, is kept on virtual or cloud servers that are maintained by outside parties. You can do this to move data to an offsite storage system where you can access it anytime you need to. The fundamentals of cloud storage are covered in this article. The cloud computing approach of "cloud storage" enables customers to store crucial data or media files on distant, outside servers. These servers are always available to users through the internet. Cloud analytics, according to York, (2023), is the practice of leveraging cloud storage and analysis to gather useful business information. The application of cloud analytics algorithms to big data collections to find patterns, forecast outcomes, and provide other information helpful to business decision-makers is but one variation of data analytics. On-premises analytics, which necessitates that companies buy, house, and manage expensive data centers, is often a less effective

option. On-premises analytics systems offer businesses internal control over data privacy and security, but scaling them is challenging and expensive.

CONCEPT OF CLOUD COMPUTING

According to Richard (2023), cloud computing is defined as the storing and accessing of data and computing services over the internet. It doesn't store any data on your personal computer. It is the on-demand availability of computer services like servers, data storage, networking, databases, etc. The main purpose of cloud computing is to give access to data centers to many users. Users can also access data from a remote server. Cloud computing is the on-demand availability of computer system resources, especially data storage (cloud storage) and computing power, without direct active management by the user. Large clouds often have functions distributed over multiple locations, each of which is a data center. Cloud computing relies on the sharing of resources to achieve coherence and typically uses a pay-as-you-go model, which can help reduce capital expenses but may also lead to unexpected operating expenses for users. Cloud computing is the delivery of different services through the Internet. These resources include tools and applications like data storage, servers, databases, networking, and software. Rather than keeping files on a proprietary hard drive or local storage device, cloud-based storage makes it possible to save them to a remote database. As long as an electronic device has access to the web, it has access to the data and the software programs to run it. Cloud computing is a popular option for people and businesses for a number of reasons, including cost savings, increased productivity, speed and efficiency, performance, and security.

According to Frankenfield (2023), cloud computing is the delivery of different services through the Internet. These resources include tools and applications like data storage, servers, databases, networking, and software. Rather than keeping files on a proprietary hard drive or local storage device, cloud-based storage makes it possible to save them to a remote database. As long as an electronic device has access to the web, it has access to the data and the software programs to run it. Cloud computing is a popular option for people and businesses for a number of reasons, including cost savings, increased productivity, speed and efficiency, performance, and security. Cloud computing is the delivery of computing services, including servers, storage, databases, networking, software, analytics, and intelligence, over the internet ("the cloud") to offer faster innovation, flexible resources, and economies of scale. You typically pay only for the cloud services you use, helping you lower your operating costs, run your infrastructure more efficiently, and scale as your business needs change (Microsoft.com, 2023). Wesley (2023) defines cloud computing as a general term for anything that involves delivering

hosted services over the internet. These services are divided into three main categories or types of cloud computing: infrastructure as a service (IaaS), platform as a service (PaaS), and software as a service (SaaS) (Wikipedia, the free encyclopedia, 2023).

CONCEPT OF CLOUD STORAGE

According to Neha & Chiradeep (2021), cloud storage is defined as a data deposit model in which digital information such as documents, photos, videos, and other forms of media is stored on virtual or cloud servers hosted by third parties. It allows you to transfer data to an offsite storage system and access it whenever needed. Cloud storage is a cloud computing model that allows users to save important data or media files on remote, third-party servers. Users can access these servers at any time over the internet. Also known as utility storage, cloud storage is maintained and operated by a cloud-based service provider. From greater accessibility to data backup, cloud storage offers a host of benefits. The most notable are the large storage capacity and minimal costs. Cloud storage delivers on demand and eliminates the need to purchase and manage your own data storage infrastructure. With "anytime, anywhere" data access, this gives you agility, global scale, and durability. Cloud storage works as a virtual data center. It offers end users and applications a virtual storage infrastructure that can be scaled to the application's requirements. It generally operates via a web-based API implemented remotely through its interaction with in-house cloud storage infrastructure. Cloud storage includes at least one data server to which a user can connect via the internet. The user sends files to the data server, which forwards the message to multiple servers, manually or in an automated manner, over the internet. The stored data can then be accessed via a web-based interface. To ensure the constant availability of data, cloud storage systems involve large numbers of data servers. Therefore, if a server requires maintenance or fails, the user can be assured that the data has been moved elsewhere to ensure availability.

Cloud storage is a cloud computing model in which data is stored on remote servers and accessed over the internet or through a dedicated private network connection. The storage infrastructure is maintained, operated, and managed by the cloud storage service provider. In general, cloud storage operates through a web-based application programming interface (API) that is remotely implemented through its interaction with the client application's in-house cloud storage infrastructure for input/output (I/O) and read/write (R/W) operations. When delivered through a public service provider, cloud storage may also be referred to as utility storage.

Joel M. PETERS Ph.D
&
Theresa Nse Udo-Okon, Ph.D

CLOUD COMPUTING AND ANALYTICS

Cloud computing and analytics are two fundamental concepts in the world of modern technology and data management. They are intertwined, with cloud computing providing the infrastructure and resources necessary for data storage and processing, while analytics leverages this data to generate insights and drive informed decision-making. Analytics involves the use of data, statistical analysis, and computational techniques to uncover patterns, gain insights, and make data-driven decisions. It encompasses various types, including:

- **Descriptive Analytics**

Descriptive analytics focuses on summarizing historical data to provide an understanding of past events and trends. It forms the basis for further analysis (Provost & Fawcett, 2013).

- **Diagnostic Analytics**

Diagnostic analytics delves into why specific events occurred, aiming to identify the root causes of trends or issues.

- **Predictive Analytics**

Predictive analytics uses historical data to build models and make predictions about future events or trends.

- **Prescriptive Analytics**

Prescriptive analytics goes beyond predictions to recommend actions that organizations can take to optimize outcomes. Analytics is used across various domains, including business, healthcare, finance, and more, to make informed decisions and gain a competitive edge. Cloud computing and analytics are two intertwined technological paradigms that have transformed the way organizations process, store, and analyze data. Cloud computing provides the infrastructure and resources necessary for data storage and processing, while analytics harnesses the power of data to generate insights and drive decision-making. The relationship between cloud computing and analytics;

- **Cloud Computing as an Enabler**

Cloud computing has become a critical enabler for analytics due to its scalability, flexibility, and cost-effectiveness. Organizations can store vast amounts of data in cloud-based storage solutions like Amazon S3 or Azure Blob Storage. This data can then be accessed and processed using cloud-based analytics platforms such as AWS Analytics, Google Cloud's BigQuery, or Microsoft Azure's Power BI.

- ## Big Data Analytics

The advent of big data has necessitated the use of cloud computing for analytics. Big data analytics involves processing and analyzing large volumes of data, which can be challenging to manage on traditional on-premises infrastructure. Cloud computing platforms provide the necessary computational resources and distributed processing capabilities to handle big data analytics workloads efficiently.

- ## Machine learning and AI

cloud-based machine learning and artificial intelligence (AI) services have gained popularity for data-driven decision-making. Services like AWS SageMaker, Google Cloud AI, and Azure Machine Learning allow organizations to build and deploy machine learning models at scale. These services leverage the cloud's computational power to train models on large datasets and make predictions in real-time.

- ## Data Warehousing

Cloud data warehouses, such as Amazon Redshift and Snowflake, offer high-performance analytics capabilities. They can quickly process and analyze large datasets, making them suitable for business intelligence and reporting tasks. Cloud data warehouses are designed to handle complex queries and support integration with analytics tools (Larson, 2019).

- ## Analytics as a Service (AaaS)

The cloud has made it possible to offer analytics as a service. Organizations can leverage cloud-based analytics platforms to perform various types of analytics, including descriptive, diagnostic, predictive, and prescriptive analytics. This allows businesses to gain valuable insights from their data without the need for significant upfront investments in infrastructure (Provost & Fawcett, 2013).

Cloud computing and analytics have become integral components of modern information technology, revolutionizing the way organizations manage and derive value from their data. Cloud computing offers a scalable and cost-effective infrastructure for storing, processing, and accessing data, while analytics leverages this data to provide actionable insights. The synergy between cloud computing and analytics has enabled businesses to harness the power of data-driven decision-making.

- **Data Processing and Analytics**

Cloud platforms offer a wide array of data processing and analytics tools, making it easier for organizations to extract insights from their data. Big data analytics frameworks like Apache Hadoop and Spark can be deployed on cloud infrastructure to analyze large datasets efficiently (Chen 2014). Cloud-based analytics services also provide tools for data visualization, machine learning, and business intelligence.

CLOUD COMPUTING AND BACKUPS

According to Open Text (2023), cloud backup, sometimes referred to as online backup or remote backup, is the process of backing up data to cloud-based servers. When you back up your data to the cloud, you're storing a copy of that data on one or more remote servers, which are owned and managed by a third-party cloud service provider. Typically, cloud service providers charge fees based on things like the amount of storage space or servers required, available server bandwidth, and the number of users who access these servers. Cloud backup, also known as online backup or remote backup, is a strategy for sending a copy of a physical or virtual file or database to a secondary, off-site location for preservation in case of equipment failure, site catastrophe, or human malfeasance. The backup server and data storage systems are usually hosted by a third-party cloud or SaaS provider that charges the backup customer a recurring fee based on storage space or capacity used, data transmission bandwidth, number of users, number of servers, or number of times data is retrieved. Implementing cloud data backup can help bolster an organization's data protection, business continuity, and regulatory compliance strategies without increasing the workload of IT staff. The labor-saving benefit can be significant and enough of a consideration to offset some of the additional costs associated with cloud backup, such as data transmission charges (Bigelow, S. 2023).

CLOUD COMPUTING AND DATA LAKES

According to Dremio (2023), a cloud data lake is a cloud-hosted centralized repository that allows you to store all your structured and unstructured data at any scale, typically using an object store such as Amazon S3 or Microsoft Azure Data Lake Storage (ADLS). Its placement in the cloud means it can be interacted with as needed, whether it's for processing, analytics or reporting. Cloud data lakes can be used to store an organization's data, including data generated from internal and external actions and interactions. The term data lake is traditionally associated with Apache Hadoop-oriented object storage. In such a scenario, an organization's data is loaded into the Hadoop platform and then analyzed as it resides on Hadoop's cluster nodes of commodity computers. While traditional data lakes have been built on HDFS clusters

on premises, the current trend is to move and maintain data lakes in the cloud as an infrastructure-as-a-service. A data lake can include structured data from relational databases (rows and columns), semi-structured data such as CSV, JSON and more, unstructured data (documents, etc.) and binary data such as images or video. The primary utility of this shared data storage is in providing a united source for all data in an organization. Each of these data types can then be collectively transformed, analyzed and more. A cloud data lake is a cloud-hosted storage solution for an organization's data, structured and unstructured and from various sources. A cloud data lake serves as a unified source of truth for an entire organization's data needs, including analysis and developing insights. Cloud data lakes provide near-unlimited capacity and scalability for the storage and computing power you need, combining the power of analytics with the flexibility of big data models and the agility and limitless resources of the cloud. A cloud data lake dramatically simplifies the effort to derive insights and value from all that data and ultimately produces faster business results (Snowflake Inc. 2023).

Cloud computing and data lakes represent two pivotal components of modern data management and analysis. Cloud computing provides scalable, cost-effective infrastructure, while data lakes serve as repositories for diverse and large datasets. The convergence of these two technologies has revolutionized how organizations store, process, and analyze data. The relationship between cloud computing and data lakes;

- **Cloud-Based Data Lakes**

Data lakes, as centralized storage repositories, are often hosted in the cloud. Cloud-based data lakes, such as Amazon S3, Azure Data Lake Storage, and Google Cloud Storage, offer scalable, durable, and cost-efficient options for storing vast volumes of structured and unstructured data (Marz & Warren, 2015).

- **Scalability and Flexibility**

Cloud-based data lakes can easily scale to accommodate growing data volumes, making them highly flexible. Organizations can ingest data from diverse sources, including logs, social media, sensors, and more, into their data lakes, providing a single source of truth for analytics (Armbrust et al., 2010).

- **Cost of efficiency**

Cloud computing's pay-as-you-go model aligns with the cost-efficient nature of data lakes. Organizations pay only for the storage and computational resources they consume, eliminating the need for large upfront investments in infrastructure (Armbrust et al., 2010).

- **Data Processing and Analytics**

Cloud computing provides the computational power necessary for processing and analyzing data within data lakes. With technologies like Apache Spark or cloud-based data analytics services, organizations can glean valuable insights from their data lakes at scale (Chen et al., 2014).

- **Real-Time Data Ingestion**

Cloud-based data lakes enable real-time data ingestion, supporting streaming analytics and immediate decision-making. Organizations can analyze data as it's generated, uncovering trends and anomalies in real time (Sharma, 2019).

- **Integration with Advanced Analytics**

Data lakes hosted in the cloud integrate seamlessly with advanced analytics and machine learning services provided by cloud providers. This integration enables organizations to build predictive and prescriptive models to extract actionable insights from their data (Géron, 2019).

- **Data Governance and Security**

Cloud providers offer robust data governance and security features. They employ encryption, access controls, auditing, and compliance certifications to protect data in data lakes, addressing concerns related to data privacy and regulatory requirements (Amazon Web Services 2023).

MERITS OF CLOUD COMPUTING

Cloud computing is an emerging technology that almost every company is switching to from its on-premise technologies. Whether it is public, private, or hybrid, cloud computing has become an essential factor for companies to rise above the competition. The advantages of cloud computing over other on-premise technologies;

- ➢ **Cost efficiency**: The biggest reason behind companies shifting to Cloud Computing is that it takes considerably lesser cost than any on premise technology. Now, companies need not store data in disks anymore as the cloud offers enormous storage space, saving money and resources.
- ➢ **High speed:** Cloud computing lets we deploy the service quickly with fewer clicks. This quick deployment lets us get the resources required for our system within minutes. Excellent accessibility Storing information in the cloud allows us to access it anywhere and anytime, regardless of the machine making it a highly accessible and flexible technology of the present time.
- ➢ **Back-up and restore data:** Once data is stored in the cloud, it is easier to get its backup and recovery, which is quite a time-consuming process in on premise technology.
- ➢ **Manageability:** Cloud Computing eliminates the need for IT infrastructure updates and maintenance since the service provider ensures timely, guaranteed, and seamless delivery of our services and also takes care of all the maintenance and management of our IT services according to the service-level agreement (SLA).
- ➢ **Sporadic batch processing:** Cloud computing lets us add or subtract resources and services according to our needs. So, if the workload is not 24/7, we need not worry about the resources and services getting wasted, and we won't end up stuck with unused services.
- ➢ **Strategic Edge:** Cloud computing provides a company with a competitive edge over its competitors when it comes to accessing the latest and most mission-critical applications that it needs without having to invest its time and money in its installations. It lets the company focus on keeping up with the business competition by offering access to the most trending and in-demand applications and doing all the manual work of installing and maintaining the applications for the company.
- ➢ **Easy implementation:** With cloud hosting, businesses can retain the same applications and processes without having to deal with the backend technicalities. Cloud infrastructure can be accessed easily. No hardware required since the cloud hosts everything, there is no requirement for a physical storage center. However, a backup is recommended for disaster management and business continuity.

- ➤ **Automatic Software Integration:** In the cloud, software integration occurs automatically. No additional efforts are necessary for the customization and integration of applications to meet preferences.
- ➤ **Reliability:** Cloud hosting comes with the biggest advantage of reliability. One doesn't have to worry about changes due to instant updates.
- ➤ **Mobility:** Employees working on-premises or from remote locations can easily access all the cloud services as long as they have an internet connection.
- ➤ **Unlimited storage capacity:** The cloud comes with an almost limitless storage capacity, which can be expanded at any time for a very nominal monthly fee.
- ➤ **Collaboration:** Cloud computing platforms help employees located in different locations to conveniently and securely collaborate with each other.

DEMERITS OF CLOUD COMPUTING

- ❖ **Downtime:** is often cited as one of the biggest cloud computing disadvantages. Since cloud computing systems are internet-based, service outages are always an unfortunate possibility and can occur for any reason. The financial impact of service disruptions and outages is staggering. A recent survey conducted by Uptime Institute indicates that about 31% of businesses have experienced an IT service incident or outage that significantly impacted their operations within the past three years. On average, an outage and or slowdown can cost a business upwards of $100,000 an hour. Unfortunately, no organization is immune, especially when critical business processes cannot afford to be interrupted. In 2023 several companies and services were hit by outages, including IT Glue, Microsoft, Google Cloud and AWS (a major web services provider), Oracle, and Data dog
- ❖ **Security and privacy:** Although cloud service providers implement the best security standards and industry certifications, storing data and important files on external service providers always opens up risks. Any discussion involving data must address security and privacy, especially when it comes to managing sensitive data. We must not forget what happened at Code Space and the hacking of their AWS EC2 console, which led to data deletion and the eventual shutdown of the company. Their dependence on remote cloud-based infrastructure meant taking on the risks of outsourcing everything. Of course, any cloud computing service provider is expected to manage and safeguard the underlying hardware infrastructure of a deployment. However, your responsibilities lie in the realm of user access management, and it's up to you to

carefully weigh all the risk scenarios. Though recent breaches of credit card data and user login credentials are still fresh in the minds of the public, steps have been taken to ensure the safety of data. One such example is the General Data Protection Rule (GDPR), which was recently enacted in the European Union to provide users more control over their data. Nonetheless, you still need to be aware of your responsibilities and follow best practices.

❖ **Vulnerability to attack:** Going on with the drawbacks of cloud computing, another one concerns vulnerability: in cloud computing, every component is online, which exposes potential vulnerabilities. Even the best teams suffer severe attacks and security breaches from time to time. Since cloud computing is built as a public service, it's easy to run before you learn to walk. After all, no one at a cloud vendor checks your administration skills before granting you an account: all it takes to get started is generally a valid credit card

❖ **Limited control and flexibility:** Since the cloud infrastructure is entirely owned, managed, and monitored by the cloud service provider, it transfers minimal control over to the customer. To varying degrees (depending on the particular service), cloud users may find they have less control over the function and execution of services within a cloud-hosted infrastructure. A cloud provider's end-user license agreement (EULA) and management policies might impose limits on what customers can do with their deployments. Customers retain control of their applications, data, and services, but may not have the same level of control over their backend infrastructure

❖ **Vendor lock-in:** is another perceived disadvantage of cloud computing. Easy switching between cloud services is a service that hasn't yet completely evolved, and organizations may find it difficult to migrate their services from one vendor to another. Differences between vendor platforms may create difficulties in migrating from one cloud platform to another, which could equate to additional costs and configuration complexities. Gaps or compromises made during migration could also expose your data to additional security and privacy vulnerabilities.

❖ **Cost concerns:** The last one of the disadvantages of cloud computing concerns cost. Adopting cloud solutions on a small scale and for short-term projects can be perceived as being expensive. However, the most significant cloud computing benefit is in terms of IT cost savings. Pay-as-you-go cloud services can provide more flexibility and lower hardware costs, but the overall price tag could end up being higher than you expected. Until you are sure of what will work best for you, it's a good idea to experiment with a variety of offerings. You might also make use of the cost calculators made

available by providers like Amazon Web Services and Google Cloud Platform.

CONCLUSION

The evaluation of cloud storage has highlighted both its many merits and demerit, providing a comprehensive view of this important technology. Cloud storage allows users to save important data or media files on remote, third-party servers. Users are empowered to access data from anywhere thanks to its remote accessibility, which improves flexibility and cooperation. Also known as utility storage, cloud storage is maintained and operated by a cloud-based service provider. From greater accessibility to data backup, cloud storage offers a host of benefits. The most notable are the large storage capacity and minimal costs. Cloud storage delivers on demand and eliminates the need to purchase and manage your own data storage infrastructure. It has also drawn attention to the drawbacks of cloud storage. It is clear that even if cloud storage has many benefits, it also has its drawbacks which are security and privacy issue, vulnerability to attack, cost concerns etc.

RECOMMENDATIONS

- Robust security measures, including encryption, access controls, and regular security audits, should be implemented to protect sensitive data in the cloud.
- Cloud users should develop a clear understanding of cloud storage costs, including data transfer and retrieval fees. Cost monitoring and optimization strategies should also be implemented to avoid unexpected expenses. Use cloud cost management tools and regularly review usage patterns to identify cost-saving opportunities.
- They should be maintenance of local backups and redundancy plans to ensure data availability. Implement automated backup processes and disaster recovery solutions to mitigate the risks of data loss due to cloud service outages or disruptions.

Joel M. PETERS Ph.D
&
Theresa Nse Udo-Okon, Ph.D

REFERENCES

Amazon Web Services. (2023). AWS Cloud Security Retrieved from https://aws.amazon.com/security/

Armbrust, M., Fox. A., Griffith, R. & Joseph, A. (2010) A View of Cloud Computing Available at: https://www.researchgate.net/publication/220422375_A_View_of_Cloud_Computing

Bryant, L., & Hall, N. (2017) Terraform: Up & Running: Writing Infrastructure as Code Retrieve from: O'Reilly Media.

Chen, M. (2014) Big Data: Related Technologies, Challenges and Future Prospects Retrieve from: Springer.

Davenport, T. H., & Harris, J. (2007) Competing on Analytics: The New Science of Winning Retrieve from: Harvard Business Review Press.

Dremio (2023) What Is a Cloud Data Lake? Available at: https://www.dremio.com/resources/guides/cloud-data-lakes/

Dremio (2023) What Is a Cloud Data Lake? Available at: https://www.dremio.com/resources /guides/cloud-data-lakes/

Frankenfield, J. (2023) What is Cloud Computing? Pros and Cons of Different Types of Services Available at: https://www.investopedia.com/terms/c/cloud-computing.asp

Géron, A. (2019) Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow. Retrieve from: O'Reilly Media

Larson, M. (2019). Cloud Data Warehousing for Dummies Retrieve from: Wiley.

Marz, N. & Warren, J. (2015) *Big Data: Principles and best practices of scalable real-time data systems, a*vailable at: Manning Publications.

Marz, N., & Warren, J. (2015) Big Data: Principles and best practices of scalable real-time data systems Available at: Manning Publications.

Marz, N., & Warren, J. (2015) Big Data: Principles and best practices of scalable real-time data systems Available at: Manning Publications.

Microsoft.com (2023). What is cloud computing? Available at: https://azure.microsoft.com/en-us/resources/cloud-computing-dictionary/what-is-cloud-computing#:~:text=Simply%20put%2C%20cloud%20computing%20is,resources%2C%2

Open Text (2023) What is Cloud Backup? Available at: https://www.microfocus.com/en-us/what-is/cloud backup#:~:text=

Cloud%20backup% E2%80%94sometimes %20referred% 20 to,third%2Dparty%20cloud%20service%20provider

Provost, F., & Fawcett, T. (2013) Data Science for Business: What You Need to Know About Data Mining and Data-Analytic Thinking Available at: O'Reilly Media.

Sharma, P. (2019) Big Data Analytics with Spark: A Practitioner's Guide to Using Spark for Large-Scale Data Processing, Machine Learning, and Graph Analytics Available at: Apress.

Sharma, P. (2019) Big Data Analytics with Spark: A Practitioner's Guide to Using Spark for Large-Scale Data Processing, Machine Learning, and Graph Analytics Available at: Apress.

Snowflake Inc. (2023) Cloud Data Lake Available at: https://www.snowflake.com/guides/cloud-data-lake

Wesley, C. (2023) Cloud Computing Available at: https://www.techtarget. com/search cloud computing/definition/cloud-computing

Wikipedia, the free encyclopedia, (2023) Cloud computing Available at: https://en. wikipedia.org/wiki/Cloud computing

York, T. (2023) Cloud Analytics 101: Uses, Benefits and Platforms Retrieve From: https://www.splunk.com/en_us/blog/learn/cloud-analytics.html