Chapter 2
Healthcare Data Storage Options Using Cloud

Sandhya Armoogum and Patricia Khonje

2.1 Introduction

In this era of digitalization, information and communication technologies (ICT) are widely adopted and form part of our daily life. Likewise, the healthcare system is embracing ICT to improve healthcare access, the quality of healthcare, and to be more cost-effective. Such smart healthcare system uses technologies such as wearable devices and the Internet of things (IoT), mobile Internet, big data analytics, cloud computing, and artificial intelligence (AI) to provide better diagnostics, better treatment, and intelligent tools and devices (e.g., for ambient assisted living) that improve the quality of life for patients [1]. As the global population is aging, chronic diseases are increasing. With the emergence of new viruses such as the COVID-19 coronavirus, which is causing a worldwide pandemic, smart healthcare solutions are fundamental to improve the efficiency of medical care, enhance the health service experience of citizens, and decrease the operational cost.

Wearable devices or medical devices which are integrated in an IoT network or connected to a smartphone mobile application can now monitor and collect patient information which enables remote real-time alerting, tracking, and monitoring of patients which permits personalized treatments, better accuracy, and more apt intervention by doctors. Such real-time monitoring via connected devices can save lives in event of a medical emergency like heart failure, diabetes, and asthma attacks. Body sensor networks (BSN), which consist of wireless sensor networks designed for monitoring human physiological signals, are rapidly emerging as human-centric cyber–physical systems. Fortino et al. [2] proposed the BodyCloud, which allows BSNs to be integrated on the cloud and supports the development and deployment

S. Armoogum (✉) · P. Khonje
University of Technology Mauritius, La Tour Koenig, Pointe-aux-Sables, Mauritius
e-mail: sandhya.armoogum@umail.utm.ac.mu

P. Siarry et al. (eds.), *The Fusion of Internet of Things, Artificial Intelligence, and Cloud Computing in Health Care*, Internet of Things,
https://doi.org/10.1007/978-3-030-75220-0_2

of cloud-assisted BSN applications. Furthermore, Gravina et al. [3] proposed a cloud-based activity as a service cyber–physical framework for human activity monitoring in mobility built atop the BodyCloud platform. Healthcare IoT devices are interconnected through wireless communication protocols like COAP, MQTT, Zigbee, and Bluetooth [4]. Big data analytics on medical data provide much insight into diseases and treatments [5]. Similarly, AI enables to interpret complex data and act intelligently to achieve a goal. AI and machine learning are commonly used to analyze patient data and medical imaging to make recommendations that can help doctors in their diagnostics, evaluate diseases, and perform epidemiological analysis.

Large amount of data is generated by smart healthcare system, which is referred to as healthcare big data. Such data include traditional data from electronic medical records (EMRs) or electronic health record (EHR) systems (e.g., patient's medical history, physician notes, clinical reports, laboratory reports, medications, insurance information, medical imaging such as CT scans, MRIs, X-ray reports, PET scans, tomography), and data related to computerized provider order entry (CPOE) systems. Moreover, high volume of real-time IoT-based health care and wearable devices as well as mobile applications data is also recorded such as blood pressure, oxygen and blood sugar levels monitoring, pulse oximeter readings, pacemaker readings, and stress level [6]. Research in medical field also involves omics data such as genomics (genome sequence data) and proteomics (proteome is a set of proteins produced in an organism). The size of a single sequenced human genome is approximately 200 GB [7].

Current estimates suggest a single patient generates close to 80 MB each year only in imaging and electronic medical record (EMR) data [8]. It has been estimated that up to 30% of the entire world's stored data is health-related (on the yottabyte scale) [9]. Data from the U.S. healthcare system alone reached 150 exabytes in 2011 [10]. The Global Healthcare Data Storage market is expected to reach $8.11 billion by 2026 growing at a CAGR of 16.4% from 2019 to 2026 [11]. Healthcare data are projected to grow faster than manufacturing data, financial services data, or media. Thus, the amount of healthcare data is increasing exponentially, and the need to effectively store and manage healthcare big data is crucial.

This chapter delves into the storage of healthcare big data. In Sect. 2.2, the healthcare big data management challenges are discussed. Section 2.3 presents the cloud option of storing healthcare data. Section 2.4 introduces the different types of storage that are available for storing data, as well as the different storage systems on the cloud. The storage security aspect on the cloud is described in Sect. 2.5. Finally, Sect. 2.6 summarizes and concludes the chapter by discussing the future directions.

2.2 Healthcare Big Data Management Requirements and Challenges

When storing healthcare data, the healthcare data management requirements need to be considered. Figure 2.1 depicts the main requirements of healthcare data management.

Healthcare systems have gone though the digitalization phase and have increasingly adopted IoT technology to enhance their services and improve their efficiency. This has resulted in increasing amount of data generated which has to be stored. Big data can be primarily defined as having three V characteristics, i.e., volume, variety, velocity. Healthcare data, apart from the sheer volume, are also of a disparate nature i.e., it consists of structured, semi-structured, and unstructured data as the data can be in different forms such as a table, email, image, video, and sensor feeds. Sensor monitoring often produces real-time data that need to be recorded; this depicts the velocity characteristics of healthcare big data. Similarly,

Fig. 2.1 Healthcare data management requirements

with the improving medical devices now available, the quality and size of medical imaging data have also increased. There is thus the need to adopt a scalable architecture for storing the increasingly amount of healthcare big data [12]. Moreover, for large healthcare institutions, which have care centers in different geographical locations, the data are also distributed across different servers in different locations. Often though, such data need to be accessible to different users from different locations. Healthcare data storage should be accessible 24/7, which means that the storage system has to be reliable and available.

The healthcare industry has to invest largely on upgrading and increasing their storage capacity such that it is more flexible and scalable, which can be costly. It can also be challenging to manage the different variety of clinical data generated from EMR, IoT, medical imaging, genome sequencing, etc.

Another important requirement for the data storage system is compliance, as the healthcare industry is highly regulated. In the USA, most of the regulations are represented in the Health Insurance Portability and Accountability Act (HIPAA) of 1996. In 2009, the HITECH Act (Health Information Technology for Economic and Clinical Health Act) was created to promote and expand the adoption of health information technology, specifically the use of electronic health records (EHRs) by healthcare providers, and to also remove loopholes in the HIPAA. The HITECH Act is not a replacement of the HIPAA but rather a supplement to HIPAA, and as such, it introduced tougher penalties for HIPAA compliance failures to ensure that healthcare organizations duly comply with the HIPAA Privacy and Security Rules. A failure to adhere to the complex set of rules governing access, security, and privacy can result in heavy fines for the healthcare organization. For the European Union (EU), the General Data Protection Regulation (GDPR) is enforced since 2018, and it regulates the way companies collect, process, and use personal data of EU citizens [13]. GDPR, similar to HIPAA, also has extraterritorial scope, since a healthcare institution in any country having European patients must comply with GDPR, which effectively makes GDPR another global regulation. GDPR distinguishes between three types of personal data related to health care which are subject to the rights GDPR conveys upon EU citizens. These three types of data are as follows: (1) data concerning health, (2) Genetic data, and (3) Biometric data.

There are some differences between the HIPAA and GDPR such as the requirements for active consent and a declaration of purpose in data collecting/processing, the right of an EU citizen to request a healthcare organization to delete their records under certain circumstances, and the time span for reporting data breaches (72 h under GDPR and 60 days under HIPAA). Thus, compliance to both HIPAA and GDPR requires healthcare organization to make some operational and technological changes.

Finally, the interconnected nature of modern health care also makes it vulnerable to online threats and security attacks in general as hackers and cybercriminals continuously look for vulnerabilities to exploit. Accessing patient medical history collected over years possibly from different sources helps to provide better treatment but also puts at risk the healthcare network. The healthcare industry is plagued by a myriad of cybersecurity attacks such as distributed denial-of-service attacks (DDoS)

which can disrupt an organizations ability to provide patient care, theft of data, medical insurance fraud, identity theft, insider misuse and unintentional actions, and ransomware attacks. According to Millard [14], ransomware incidents increased by 300% from 2015 to 2016. The UK's National Health System (NHS) hospitals were held hostage in May 2017 to the WannaCry ransomware which resulted in delayed treatment plans. Universal Health Services, one of the largest healthcare provider in the USA, was hit by the Ryuk ransomware in September 2020, whereby all of the 250 healthcare facilities' computer infrastructure were locked resulting in chaotic conditions which tremendously affected patient treatment [15]. In September 2020, due to a ransomware attack at the Duesseldorf University Hospital, a woman who needed urgent medical care died after being rerouted to a hospital in the city of Wuppertal, more than 30 km away [16]. This incident marked the first reported human death indirectly caused by a ransomware attack.

Going forward, the healthcare industry has to improve its data management capabilities in order to keep pace in a digital landscape and ensure security of data.

2.3 Cloud for Healthcare Big Data Storage

Traditionally, healthcare institutions have been storing their data on-premise as this allows them to have full control over the in-house data, where they minimize risks associated with data breach and maintain their own backup and recovery systems. However, today with the growing complexity of healthcare big data, which includes different geographically located healthcare facilities and the increasing number of smart health applications, there is no "one-size-fit-all" solution. It is becoming more challenging to handle big data storage, the evolving real-time data from IoT devices, Bring Your Own Devices (BYOD) practices, to enforce security, compliance, availability, and to provide ubiquitous access for on-premise healthcare data storage.

On-premise storage is often not very agile. It is required that the storage solution is scalable such that more storage capacity can be added without requiring a complete upgrade and conversion, which may not always be the case. Increasing on-premise storage capacity also implies increasing physical space within the premises which might not be easy to accommodate. Apart from the costs of additional storage solutions, there is also the operational cost incurred such as power supply, cooling of server rooms, IT staff, which makes on-premise data storage a rather costly option. Depending on the backup strategy in place, on-premise storage may also have more latency in recovery from an attack or failure. Healthcare data being critical, longer downtime can significantly impact on the quality of care given to patients when patient data is not readily available.

The cloud has emerged as a most convenient and viable, more global and holistic data storage solution for the healthcare industry [17–22]. Cloud storage is definitely more scalable and more cost-effective in terms of capital expenditure and operational expenditures as well as secure. In Sect. 2.5, the security mechanisms proposed for securing healthcare applications and data are described. More smart

health applications are being deployed on the cloud today and consequently using cloud storage for the related data [23–29]. Aileni et al. [30] propose the use of cloud for storing and processing big data from biosensors. The authors, Kolodner et al. [31], proposed an architecture of a scalable and flexible cloud environment for data-intensive storage cloud services for health care, which enable data mobility across providers and allow computational and content-centric access to storage and deploy new data-oriented mechanisms for QoS and security guarantees. The use of cloud computing has also been proposed for healthcare underground wireless sensor network [32]. Mobile cloud computing (MCC) is also popularly proposed for mobile healthcare (m-healthcare) applications [33–35]. This results in more accessibility, improved performance, better disaster recovery, and more efficiently managed systems which positively impacts the quality of medical care provided to patients ultimately. For healthcare institutions that already have on-premise storage, it is also possible to extend their storage and ICT infrastructure requirements by using the cloud platform, i.e., a hybrid solution.

The NIST [36] defines the cloud as follows "Cloud computing is a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction." There are four cloud deployment models, namely public cloud, private cloud, hybrid cloud, and community cloud. Public cloud is cloud infrastructure owned and operated by a third-party cloud service provider (CSP), delivered over the Internet. The public cloud resources can be used by anyone who subscribes to the cloud offerings. The private cloud is exclusively for the use of a single organization. Often, it is owned and built on the organizations premises, but it could also be outsourced from a third-party organization. Hybrid cloud is the result, when companies partly use their own private cloud and partly use resources from one or several public cloud. In such scenario, usually all sensitive and confidential data are kept on the private cloud. The community cloud is set up from a collaborative effort, for the specific needs of an industry, and shared by the partners. The three-service delivery model of the cloud includes the Software as a Service (SaaS) where software is consumed by users, Platform as a Service (PaaS) which allows developers to build applications, and Infrastructure as a Service (IaaS) which mainly allows to get access to a compute node on the cloud in the form of a virtual machine (VM) with its own virtual disk or even a virtual network of VMs. Cloud customers can use one or more of the service delivery model. The term Storage as a Service mainly refers to the storage of data in the form of files such as document, images, and videos. Common Storage as a Service offering on the cloud includes Google Drive, Microsoft One Drive, Apple's iCloud, and Dropbox. Such storage mechanisms are mainly for storing data or data backups as well as for sharing data among users.

Rajabion et al. [37] state that "Cloud computing is one of the main choices for handling and processing of this type of data (healthcare)." Smart healthcare applications and m-healthcare applications can easily and quickly be deployed on the cloud today where all the required resources are provisioned for the applications including

storage, whether it is in the form of relational databases, NoSQL databases, file storage, or object store.

2.4 Cloud Storage

For the healthcare industry to deliver patient care and coordinate medical information, there is the need for storage that is both reliable and secure. The suitable data storage solution is the cloud storage and the hybrid storage; i.e., healthcare facilities can migrate all their ICT services and storage on the cloud, or they can use both their on-premise systems and storage and extend their storage onto the cloud for more storage capacity.

According to Gartner [38], the leaders for Public Cloud Storage Services are Amazon Web Services (AWS), Microsoft Azure, and Google. All three major cloud providers are proposing new cloud services including cloud storage for hospitals and other healthcare facilities [39]. Recently, Microsoft has launched its new cloud service, Cloud for Healthcare, specifically designed for health care. The new services include scheduling telehealth visits, allowing virtual care via virtual visits, and support the Fast Healthcare Interoperability Resources (FHIR), which is a common data standard for health care, and a tool called Patient 360 which allows health providers to get the full information about the patient [40]. Mayo Clinic has signed a 10-year strategic partnership with Google to accelerate innovation in AI and analytics through digital technologies [41]. Amazon is also proposing solutions to store, protect, and optimize healthcare data in AWS. Cerner, a major company that supplies health information technology services, devices, and hardware, is working with AWS for developing innovative healthcare solutions using AI and machine learning (ML) [42].

2.4.1 Cloud Storage Type

Similar to the on-premise infrastructure, the cloud storage consists of magnetic storage (hard disks) and solid-state storage (flash disks).

Magnetic Storage: Hard Disk Drive (HDD)

HDD are magnetic storage media whereby data is stored in a magnetized medium. They are mechanical hard drives as they have moving parts, namely the central spindle that allows the disk platters of the HDD to rotate and the actuator that allows the read/write head to move. The speed at which the disk spins affects the time taken to read/write data from/to the HDD among others. The higher the revolutions per minute (RPM), the faster data can be read. Magnetic tapes are also another type of

magnetic storage, which is popularly used to store old patient records, and clinical data. Magnetic tapes are also suitable for storing backups as they can store data reliably for a long period of time, but data stored on magnetic tapes cannot be accessed fast online.

HDD can store very large amount of data (high capacity) at relatively cheap price and have better data access times compared to magnetic tapes. Traditionally in the healthcare industry, most data are stored on HDD and backup data on magnetic tapes. Magnetic tapes are preferred despite that they are slow and prove to be expensive when it comes to managing and maintaining them, as they are more secure given that they may not be directly accessible via the network.

Cloud data storage options also include HDD. The cloud is equipped with a large number of HDD storage arrays which are mounted on racks. The popular vendors of storage arrays include NetApp, Dell EMC, Hewlett Packard Enterprise (HPE), IBM, Huawei, and Hitachi Vantara [43]. Three different kinds of HDD are typically used: Serial ATA (SATA), Serial Attached SCSI (SAS), and Near Line SAS (NL-SAS). SATA drives typically have RPM of 5400 RPM and 7200 RPM; thus, their performance is the lowest among the three types of HDD. However, they are reliable and have large storage capacity and are relatively cheap. SAS drives typically have 10,000 and 15,000 RPM and provide the best performance and low latency, and are typically used to store data for mission critical application as they are also more reliable. The SAS drives have additional commands that control the disks, which contribute to make SAS drives more efficient than SATA. They are also more expensive than SATA drives. NL-SAS disk consists of spinning SATA platters but with the native command set of SAS drives. Due to their low rotational speed, their performance is lower than SAS drives, but overall, they are more efficient than SATA drives. The cloud data centers have all three types of HDD as well as SSD storage. Storage on SAS drives and SSD drives on the cloud provides better performance but are also more expensive. Cloud providers offer storage tiering to allow customers to optimize the use of storage. Thereby, backup data can be stored on SATA drives, for instance, while production data is stored on SAS drives instead of both being stored on SAS drives. This allows the customer to save on the cost storage by making the best use of the storage type for the different kind of data to be stored. Typically, Tier 1 could represent storage which has best performance and reliability, while also being the most expensive. Tier 2 could represent the storage which is slower and less efficient, but which is also usually cheaper. In this regards, SAS drives can be considered as being Tier 1 storage, while SATA drives could be offered as Tier 2 storage on the cloud. Tiering is further addressed in Sect. 2.4.2.

Flash Storage: Solid-State Drive (SSD)

Flash storage is a nonmechanical and nonmagnetic storage whereby data is stored using electronic circuits, i.e., NAND-based flash memory. Given that SSD drives do not have mechanical parts, they are less susceptible to physical shock, are silent, and have overall better performance than HDD and are also much smaller in size for the

same capacity HDD. However, they are also more expensive. Overtime, the cost of flash memory is becoming more affordable and SSD of larger capacity is increasingly available. SSD is the storage of choice for healthcare data that need to be readily accessible especially in cases of emergency. For instance, it would take over 15 s to read an MRI image of the size 1.8 GB stored on a HDD as compared to just 0.15 s from SSD flash storage [44]. Flash storage can also offer built-in encryption which can be particularly useful for securing healthcare data for HIPAA compliance.

Traditional Flash storage uses the SATA III interface, primarily used by HDD, which is not designed to support SSD. SATA III can support 6 GB/s of peak bandwidth theoretically, but due to overhead associated with the interface, the effective bandwidth peaks at about the 555 MB/s. This limits the access speed to the SSD storage. PCIe is one of the fastest storage communication protocols. Nonvolatile memory express (NVMe) is a new storage access and transport protocol that uses PCIe. NVMe flash storage is the next-generation SSD that delivers high bandwidth and low latency for all types of complex workloads. NVMe SSD drives are up to six times faster than the typical SATA SSD. NVMe Over Fabric (NVMe-oF) is a storage networking protocol for NVMe-enabled storage disk arrays. It enables faster and more efficient connectivity between storage disk arrays and the servers. Cloud storage providers are heavily using All-Flash NVMe arrays to provide low latency and high-throughput storage. Facebook uses NVMe cloud storage technology in their data centers around the clock to address the hyperscale issues regarding the data of their 2.5 Billion users [45].

2.4.2 Cloud Storage Tiering

Storage tiering distinguishes between more important or frequently accessed data, which should ideally be stored on the fastest type of storage (but also more costly) with less important and rarely accessed data (e.g., backup data) that can be stored on the slowest and cheapest storage type. The simplest tiering system has two tiers: Tier 1 (also known as hot storage)—provides high performance, usually expensive; and Tier 2 (also known as cold storage)—characterized by lower performance, but is cheaper. Most of the popular public cloud offers more than two tiers of storage, thus offering more choice and granularity for classifying and storing data accordingly onto the different storage tiers.

A four-tier model could involve Tier 0,1,2,3 or mission critical, hot, warm, cold tiers. Here, the tier 3/cold would be used to store rarely accessed data, e.g., old patient health records; tier 2/warm would be used for data which are not constantly accessed, e.g., outpatient health records; tier 1/hot would be data constantly used, e.g., in-patient records; and tier 0/mission critical could be data which is used constantly and which needs to be speedily accessible, e.g., medical imaging data of patient during surgery. Tiering thus allows storage resources to be used more efficiently and is also more cost-effective while at the same time providing good performance access for each class of data.

Because data classification can change over time or in certain situation, e.g., emergency health crisis of a patient, data can be moved from a lower tier to a higher tier as required or vice versa i.e. old patient data can be moved from higher tier to lower tier. Automated storage tiering is also supported whereby data can be moved across storage tiers dynamically, e.g., based on the frequency with which some records are being accessed; they can be moved to a higher tier automatically by the storage system, while data that have been less frequently used from a higher tier can be shifted to a lower tier accordingly.

Typically, the cloud providers offer more storage tier levels for their object storage offerings. Object storage is described in Sect. 2.4.3.3. An object can be stored in one storage class i.e. tier level. Users can configure policies (i.e., object life cycle management) to set up rules describing actions to be taken on objects, e.g., downgrade data not accessed for more than 30 days from the current tier to a lower tier with the aim to reduce long-term storage costs. Google object storage tiering are as follows: (1) *Standard*: Hot data or short storage periods; (2) *Nearline*: Lower cost, good for data accessed once a month or less; (3) *Coldline*: Very low cost, better for data read or modified once per quarter; (4) *Archive*: Lowest cost, for data accessed less than once a year but still with millisecond access times. Microsoft Azure offers three tiering levels for the Azure Blob Storage as follows: (1) *Hot*: Data in active use or expected to be accessed frequently; (2) *Cool*: Lower cost, monthly access or less; (3) *Archive:* Lowest cost, access time is several hours. Typically, data stored in Archive access tier is considered offline, and it takes longer to get the data online and accessible. AWS storage tiering is provided for the object storage Amazon Simple Storage Service (S3) where it has the following storage classes—(1) *Standard*: Low latency and high throughput, used for frequently accessed data and more expensive; (2) *Standard-Infrequent Access (IA)*: Same performance, lower storage charges, with retrieval charges, which is ideal for long-term storage; (3) *Glacier*: Ideal for data archiving and has low cost with a retrieval range from 1 min up to 12 h; (4) *Glacier Deep Archive*: Lowest cost storage but data is restored within 12 h. AWS S3 also supports the object life cycle management allowing users to define policies regarding the shifting of data from one tier to another. However, AWS also has another storage class called ***S3 Intelligent Tiering***, which is a two-tiered storage in a single class (frequent access/hot tier, and infrequent access/cold tier). It is ideal for long-lived data with unknown or unpredictable access patterns. Data objects in this tier are initially placed in the hot tier, and AWS automatically downgrade objects which have not been accessed for 30 consecutive days to the cold tier, and shift them back to the hot tier if they are accessed.

2.4.3 Cloud Storage: File, Block, and Object Storage

Data can be stored on the cloud as files, blocks, or objects.

File Storage

File Storage is one of the old styles of storage that are file system-based methods of storing and retrieving data. File storage organizes and represents data as files. These files are, in turn, organized in folders, and these folders are then arranged into directories and subdirectories in a hierarchical fashion. To access a file, computers need the path from directory to subdirectory to folder to file. Limited amount of metadata is stored for each file such as the name of the file and size of the file. The files can be stored locally on a directly attached storage (DAS) or remotely on network-attached storage (NAS). The NAS storage uses a network file system (e.g., NFS, CIFS, SMB) to expose data stored. AWS's Elastic File System is an NFS file system-based storage, while Azure also offers cloud-based file storage access via SMB. Google Cloud Filestore is the network file storage solution of the Google Cloud. File storage is scalable as using scale-out NAS devices such as the Dell EMC Isilon; the storage capacity can be increased. Data can be easily shared between users in cloud file storage for collaboration. Administrators can also easily set access rights of files, while security and version control are easily managed. However, file storage is suitable for small organizations. The tree-like file system hierarchy can handle millions of files quite easily, but when the number of files increases further, the file access start to slow down due to the overhead of the underlying file system lookup to find the file location. File storage on cloud can be used to store some type of healthcare data, but it is not the ideal solution for large healthcare facilities. It is also not suitable for storing all the different variety of healthcare data and especially the large volume of real-time healthcare big data generated by health IoT networks.

Block Storage

Block storage chunks data into arbitrarily organized, evenly sized volumes. Each block of data is given a unique identifier, which allows a storage system to store the block of data anywhere on the storage area. No metadata are stored for the block of data. When data are requested, the storage system retrieves and reassembles the blocks of data. Block storage is usually deployed in storage area network (SAN) environments and must be tied to a functioning server (controller). A SAN is a specialized high-speed network that provides block-level network access to storage devices which are typically all-flash disk arrays. SAN environments typically have high performance, consistent low latency, high availability, and resilience. Due to their centralized architecture, consistent methodologies, and tools for security, data protection and disaster recovery can be applied to the SAN. Data can be replicated for reliability and/or stored in a RAID configuration. Block storage and SAN are often used for performance-hungry applications like databases, email servers, virtual disks associated to virtual machines, and large virtual desktop infrastructures (VDIs). Block storage is suited for structured data and is costly as SAN consists

often of a bunch of SSD disks arrays. It can be used for storing structured healthcare data.

AWS offers the Amazon Elastic Block Store (EBS) which is an easy-to-use, high-performance block storage service designed for use with Amazon Elastic Compute Cloud (EC2). It is most suitable for transaction-intensive applications and workloads requiring high throughput at any scale. On EBS, users can store data on five types of storage to optimize storage performance and cost for a broad range of applications, namely three types of SSD (io2, io1, and gp2) and two types of HDD (st1 and sc1) [46, 47]. The Microsoft block storage offering is the Azure Managed Disk where data can be stored on different types of storage media for different performance levels: Premium SSD, Standard SSD, Standard HDD, and Ultra Disk. Google Cloud Persistent Disk provides block storage which is reliable and has high performance; it is used by all virtual machines in Google Cloud (Google Cloud Compute Engine). Data can be stored on three storage types: standard persistent disks, balanced persistent disks, and SSD persistent disks; the latter are designed for enterprise applications and high-performance database needs that require lower latency and more IOPS. SSD persistent disk storage option is also the most expensive. Google also allows choosing where the storage is located (i.e., which region) and the type of availability. Regional persistent disks are replicated in different zones within a region and offer high availability, whereas zonal disks (cheaper option) also provide high availability but with disks replicated only within a single zone [48, 49].

Object Storage

With object storage, data items are not stored in blocks or files but rather in flexibly sized containers called "objects" or "blobs." The object size can vary from a few kilobytes to multiterabytes. Object stores can hold any type of data (e.g., text, image, video, and sensor data) as the object is essentially stored as binary data. Object storage is thus suitable for storing a large volume of unstructured data. Each object is uniquely identified by means of an objectID; they are stored in a flat address space, whereby it is easier to locate and retrieve them, even if they are stored in geographically distributed storage clusters. Each object is linked to associated metadata. With object storage, lots of metadata can be stored about the object, and the objects can be accessed directly through REST APIs or HTTP/HTTPs requests using their identifier. Object storage is potentially good for analytics/searching due to its rich metadata. But it is not convenient for transactional data, e.g., databases. Object storage is cost-efficient and extremely scalable as it can be scaled out and managed simply by adding additional nodes to the storage cluster. Object storage is also very robust and resilient. To protect data stored in an object store from both disk failure (i.e., one node fails, data can still be accessed) and data corruption, replication or erasure coding (EC) or both are applied. Object store is most convenient for hospitals which with the growth in digitalization are required to store large amounts of unstructured data more efficiently and securely.

Traditionally, object storage, due to their highly scalable and reliable storage mechanism for storing large volumes of data, is also suitable for storing archived data, backup data, large datasets of unstructured data for analytics, and Web applications that require large storage capacity which can be accessed via REST APIs. Cloud providers offer different storage tiers based on performance, cost, and frequency of access for object storage which makes it more and more suitable for IoT applications, Web applications, and data analytics. Thus, today, the cloud object storage can be used for storing and accessing a wide variety of applications data. According to Symonds [50], "Modern object storage is cloud-native and suited for containerization and orchestration, making it truly elastic."

Microsoft Azure has Blob storage which is the object storage solution where large amount of unstructured data can be stored. The storage service offers three types of blobs: block blobs, append blobs, and page blobs. They are all geared toward providing better read/write access. Block blobs are optimized for uploading large amounts of data efficiently. Page blobs are a collection of 512-byte pages optimized for random read and write operations. An append blob is comprised of blocks and is optimized for append operations. For cost-effective data storage, Azure block storage supports two performance tiers: (1) Premium: optimized for high transaction rates and single-digit consistent storage latency where data is stored on SSD; and (2) Standard: optimized for high capacity and high throughput [51]. Google object storage solution is the "Google Cloud Storage." Tiering is supported as described in Sect. 2.4.2. Similarly, Amazon offers the Simple Storage Service (Amazon S3) as an object storage service which provides scalability, data availability, security, and performance.

Healthcare Cloud Storage

Healthcare big data consist of large volumes of structured as well as unstructured data. File storage is not scalable enough to be able to store healthcare big data. Block storage is suitable for storing structured transactional healthcare data such as appointment scheduling, invoicing, and insurance data, but it cannot easily scale up and is also a more costly option. Thus, for storing and accessing unstructured clinical data in the healthcare industry such as patients' health records, prescriptions, medical imaging, IoT data, genome data for research, bioinformatics datasets, object storage is most suitable as it is easily scalable, reliable and secure, has good performance, and is extremely cost-effective compared to block storage solutions [52]. Modern healthcare applications communicate through RESTful APIs which is inherently compatible with cloud object storage solutions such as AWS S3. The FHIR (Fast Healthcare Interoperability Resources) Specification, which is a standard for exchanging healthcare information electronically, is also based on RESTful APIs. Cloud healthcare applications can also easily access IoT data stored at the edge of the cloud. The Broad Institute has adopted the Google Cloud Platform for storage and processing of human genome sequence with the aim of finding patterns, which may reveal the origins of diseases. It was revealed that the cloud platform

analyzes human genomes 400% faster, while also ensuring security and privacy of the data [53]. On-premise storage can also easily be leveraged to include cloud object storage.

2.4.4 *Blockchain Technology for Storage of Data in Health Care*

A blockchain is a digital distributed ledger; i.e., it consists of a time-stamped series of immutable records of data also known as "blocks" that are managed and stored in a network connected through peer-to-peer nodes. Each of these blocks of data is secured and bound to each other using cryptographic hashing mechanisms to constitute a "chain." Blockchain is inherently reliable due to its distributed nature and is also highly secure. Access to the blockchain is based on authentication; new blocks are authorized and secured by means of digital signatures. Furthermore, blocks in the blockchain are safeguarded from tampering, thus providing integrity, by means of cryptographic hashing and chaining principles. The use of blockchain was first introduced in the context of cryptocurrency for the Bitcoin protocol. However, today, blockchain technology is being explored and used in various industries such as business, finance, retail, agriculture, manufacturing, real estate, and the IoT. Blockchain technology is also being increasingly adopted in the healthcare industry for a wide range of applications and data storage.

Blockchain technology can effectively be used as a distributed storage mechanism [54]. The authors propose the use of the Interplanetary File System (IPFS) and blockchain for storing patient diagnostic reports in healthcare facilities [54]. Blockchain technology can improve healthcare data sharing and storing due to its decentralized nature, immutability of data stored in the blockchain, and the transparency and traceability features of the blockchain [55, 56]. Shahnaz et al. [57] proposed to implement a scalable EHR using blockchain technology where granular access control is provided to secure the storage of patient data. Similarly, Mayer et al. [58] looked into the use of blockchain for storing and sharing EHR data and which can facilitate trust between healthcare providers for sharing data, EHR interoperability, auditability, and privacy. Usman and Qamar [59] used the permissioned Hyperledger blockchain platform for implementing a prototype of EMR Management System. Zheng et al. [60] proposed the use of blockchain technology for users to securely store and share such personal health data in a GDPR compliant manner using cloud storage. Wearable/fitness devices (e.g., Apple Watch, Fitbit Charge 4, Samsung Galaxy Fit, Garmin Vivosmart 4, Whoop Strap 3.0) and mobile applications also store large amount of personal health-related data in databases. Such data can be useful for research or commercial project and can be shared via a blockchain. Kassab et al. [61], in a review article, suggest that blockchain technology has lots of potential in the sphere of health care. It can not only securely store and manage patient data (the most popular use case scenario for blockchain) but

also track the entire supply chain of medicines and medical equipment such that it facilitates accountability and transparency as well as providing reliable data for clinical research and insurance claims. The four key benefits of using blockchain technology were (1) improved availability of data; (2) improved transparency; (3) improved security; and (4) improved performance. The following challenges were also identified: (1) Scalability and performance; (2) Usability; (3) Secure identification of patients; and (4) Lack of Incentives and willingness to adopt blockchain technology. Tariq et al. [62] recommend that the use of blockchain technology for IoT-enabled smart healthcare systems, which are particularly vulnerable to security breaches, can be beneficial but there are also challenges regarding the deployment of large scale blockchain. Esposito et al. [63] in their paper title "Blockchain: A Panacea for Healthcare Cloud-Based Data Security and Privacy?" discuss the potential of using the Blockchain technology to protect healthcare data hosted within the cloud.

2.5 Cloud Storage Security

Deploying healthcare applications and big data storage on public or hybrid cloud is vulnerable to threats posed by attackers, cloud employees, or vendors subcontracted by the cloud service provider. Healthcare providers have to adhere to storing and sharing data securely whether the data are stored on-premise or on the cloud. For HIPAA and GDPR compliance, healthcare data security and privacy are even more important. Protected Health Information (PHI) relates to identifiable health information that is stored in electronic form such as EHR/EMRs health records, laboratory test results, health histories, diagnoses, treatment information, insurance information, and lists of allergies. Basic security mechanisms such as firewall protection, AES data encryption at rest, SSL security for data in transit, secure key management, strong authentication, policy-based data loss protection (DLP) controls for storage have to be implemented [64]. Some of the security and privacy concerns that are thought to come with the use of the cloud include abuse of privileges, poor encryption, identity theft, data loss, and theft because of the shared and open environment [65]. Issues to do with trust, security attacks, and confidentiality are some of the concerns causing reluctance in embracing cloud storage [66–68]. Zhang and Liu [69] presented the important security concepts related to data sharing in EHR system deployed on the cloud and analyzed the arising security and privacy issues in the access and management of the cloud-based EHR systems. Carter [70] emphasizes the considerations and importance for genomic data privacy and security when working in the cloud to ensure compliance. Much research is being conducted to propose novel ways of securing healthcare data and applications deployed on the cloud to address the security shortcomings.

Khattak et al. [71] present the analysis of security concerns of cloud-based healthcare systems such as availability, trust, confidentiality, compliance, integrity, and audit and concludes that a single-cloud solution is more vulnerable to attack

from malicious cloud employees and denial-of-service attacks due to some component failure. While most focus has been on the security related to single cloud, the authors propose a multi-cloud model also known as cloud-of-clouds, as a more secure cloud computing and storage environment for e-healthcare. Similarly, Ziglari and Negini [72] propose the use of several cloud service providers collaborating to offer security in EHR systems; for instance, the EHR application can be deployed on one cloud provider, while the data storage is implemented on another cloud provider for more resilience and security of the overall EHR system. Likewise, a backup of the data could be stored with a third cloud provider. Tunneled communication (cryptographically secured) is used for communication between cloud providers for the purpose of the application.

Hombal and Dayananda [73] address the security issue of accessing and sharing data among the medical practitioners of different healthcare facilities in an efficient manner, i.e., sharing data among a group. The proposed system involves the group members to perform a key agreement protocol, whereby they obtain a common group key that allows sharing of data securely while also detecting malicious users during the group data sharing.

Tupakula and Varadharajan [74] proposed techniques to establish trust on the tenant virtual machine (VM) by performing attestation using the Trusted Platform Module (TPM) chip as well as securing the tenant communication in the context of healthcare facilities (customer/tenant) using the IaaS on the cloud. Attestation allows the customer to ensure integrity of the VM.

Dean et al. [75] depict the design and implementation of a scalable and secure multitenant cloud environment for storage and processing of PHI data in a compliant manner. Alexander and Sathyalakshmi [76] proposed a privacy-aware system and anonymization techniques for data publishing on cloud for PHI based on k-anonymity algorithm and the Advanced Encryption Standard (AES). Shakil et al. [77] focus on the authentication of users accessing and managing the cloud-based BAMHealthCloud system. A behavioral biometric signature-based authentication is proposed for secure data access and retrieval. Al Hamid et al. [78] proposed a key agreement protocol based on a bilinear pairing cryptography for generating a session key to be used to secure medical big data to be shared among the users with whom the session key is shared. Attackers attempting to access the secure data are redirected to decoy data gallery, while only verified users are allowed access to the medical data. Roy et al. [79] propose a fine-grained data access control over multiple cloud servers in a mobile cloud computing-based healthcare applications.

Moreover, cloud storage providers also provide security for data. AWS S3 provides access control mechanisms to block the public in general from accessing the data stored on the cloud as well as configuring the bucket access control lists to set access rights for users based on the principle of least privilege [47]. Users must be authenticated to get access to the storage. AWS typically uses SSL/TLS for securing data in transit from the cloud to the user. Server-side encryption (SSE) can be opted for the encryption of the data stored. Three SSE options are available as follows: (1) Server-side encryption with Amazon S3-managed keys (SSE-S3); (2) Server-side encryption with customer master keys stored in AWS Key Management

Service (SSE-KMS); and (3) Server-side encryption with customer-provided keys (SSE-C). If data stored on the S3 cloud storage are data which do not require modification, e.g., X-Ray image, by using the "Object Lock" feature of S3, they can be stored using a "Write Once Read Many" (WORM) storage. The Object Lock feature prevents accidental or malicious deletion of data. For data that can be modified, the versioning feature can be used to preserve, retrieve, and restore the different versions of the object. Likewise, versioning builds resilience in the storage system. To provide high availability, data can be replicated in multi regions. Finally, AWS S3 also has features that allow the user to monitor and log access to the data.

Microsoft Azure also supports several security features for the cloud storage [80]. Azure role-based access control (Azure RBAC) allows setting up access rights to users. Only authenticated and authorized users are allowed access to the data. Azure Defender, if enabled, detects unusual and potentially harmful attempts to access or exploit storage accounts and trigger security alerts to inform the administrator. The administrator can lock users so that they cannot intentionally or unintentionally modify or delete data. Azure storage can also store blob as WORM state whereby the data are immutable. Azure storage soft delete feature allows recovering blob data after it has been deleted. Blobs can also be encrypted and stored at rest as well as in transit (256-bit AES encryption). Similar security features are also provided by Google Cloud Object Storage [49].

2.6 Conclusion

In this chapter, the healthcare data storage options on the cloud as well as the security of healthcare data are described. The healthcare industry is rapidly evolving and turning into smart healthcare systems with the adoption of new technologies such as cloud computing, mobile computing, IoT, AI and machine learning, and data analytics. New challenges will be faced as the smart healthcare systems mature and become more agile, reliable, secure, and robust. The use of blockchain technology for storing healthcare data in the cloud is also promising due to the security provided by the blockchain technology. The demand for storage capacity for healthcare data will only increase with time. For offloading archived healthcare data from cloud storage, new storage may be adopted such as DNA storage, which involves the use of synthetic DNA for storing data. DNA has very high storage density, has potentially low maintenance cost, is energy efficient, and can store data reliably for a longer period of time as compared to both HDD and SDD. Currently, DNA storage is expensive and not being used on a large scale.

References

1. Aceto, G., Valerio, P., & Antonio, P. (2020). Industry 4.0 and health: Internet of things, big data, and cloud computing for healthcare 4.0. *Journal of Industrial Information Integration, 18*, 100129. ISSN: 2452-414X.
2. Fortino, G., Parisi, D., Pirrone, V., & Di Fatta, G. (2014). BodyCloud: A SaaS approach for community body sensor networks. *Future Generation Computer Systems, 35*, 62–79., ISSN 0167-739X.. https://doi.org/10.1016/j.future.2013.12.015
3. Gravina, R., Ma, C., Pace, P., Aloi, G., Russo, W., Li, W., & Fortino, G. (2017). Cloud-based activity-aaService cyber–physical framework for human activity monitoring in mobility. *Future Generation Computer Systems, 75*, 158–171., ISSN 0167-739X.. https://doi.org/10.1016/j.future.2016.09.006
4. Zeadally, S., Siddiqui, F., Baig, Z., & Ibrahim, A. (2019). Smart healthcare: Challenges and potential solutions using internet of things (IoT) and big data analytics. *PSU Research Review, 4*(2), 149–168. https://doi.org/10.1108/PRR-08-2019-0027
5. Rath, M. (2018). Real time analysis based on intelligent applications of big data and IoT in smart health care systems. *International Journal of Big Data and Analytics in Healthcare, 3*(2), 45–61.
6. Suter-Crazzolara, C. (2018). Better patient outcomes through mining of biomedical big data. *Journal of Frontiers in ICT, 5*, 30.
7. Robison, R. J. (2014). How big is the human genome? *Precision Medicine*. https://medium.com/precision-medicine/how-big-is-the-human-genome-e90caa3409b0. Accessed 24 October 2020.
8. Huesch, D., & Mosher, T. J. (2017). Using it or losing it? The case for data scientists inside health care. *Journal of NeJM Catalyst Innovations in Care Delivery*. https://catalyst.nejm.org/case-data-scientists-inside-health-care/
9. Faggella, D. (2019). Where Healthcare's big data actually comes from. https://www.techemergence.com/where-healthcares-bigdata-actually-comes-from. Accessed 20 June 2018.
10. Raghupathi, W., & Raghupathi, V. (2014). Big data analytics in healthcare: Promise and potential. *Health Information Science System, 2*, 3. https://doi.org/10.1186/2047-2501-2-3
11. Syngene Research. 2020. Global healthcare data storage market analysis 2020. https://www.researchandmarkets.com/reports/5136822/global-healthcare-data-storage-market-analysis. Accessed 24 November 2020.
12. Dash, S., Shakyawar, S. K., Sharma, M., et al. (2019). Big data in healthcare: management, analysis and future prospects. *Journal of Big Data, 6*. https://doi.org/10.1186/s40537-019-0217-0. Article no. 54.
13. Argaw, S. T., Troncoso-Pastoriza, J. R., Lacey, D., et al. (2020). Cybersecurity of hospitals: Discussing the challenges and working towards mitigating the risks. *BMC Medical Informatics and Decision Making, 20*, 146. https://doi.org/10.1186/s12911-020-01161-7
14. Millard, W. B. (2017). Where bits and bytes meet flesh and blood. *Annals of Emergency Medicine, 70*(3), A17–A21. https://doi.org/10.1016/j.annemergmed.2017.07.008
15. Guardian. (2020). US hospital systems facing 'imminent' threat of cyber attacks, FBI warns. https://www.theguardian.com/society/2020/oct/28/us-healthcare-system-cyber-attacks-fbi. Accessed 29 October 2020.
16. Cimpanu, C., (2020). First death reported following a ransomware attack on a German hospital, September 17, 2020. https://www.zdnet.com/article/first-death-reported-following-a-ransomware-attack-on-a-german-hospital/. Accessed 30 October 2020.
17. Ali, O., Shrestha, A., Soar, J., & Wamba, S. F. (2018). Cloud computing-enabled healthcare opportunities, issues, and applications: A systematic review. *International Journal of Information Management* . ISSN: 0268-4012., *43*, 146–158. https://doi.org/10.1016/j.ijinfomgt.2018.07.009

18. Casola, V., Castiglione, A., Choo, K. R., & Esposito, C. (2016, November-December). Healthcare-related data in the cloud: Challenges and opportunities. *IEEE Cloud Computing, 3*(6), 10–14. https://doi.org/10.1109/MCC.2016.139
19. Kumari, A., Tanwar, S., Tyagi, S., & Kumar, N. (2018). Fog computing for healthcare 4.0 environment: Opportunities and challenges journal of. *Computers and Electrical Engineering, 72*, 1–13.
20. Shanmugam, M., & Singh, M. (2017). A comparitive study on traditional healthcare system and present healthcare system using cloud computing and big data. 2017 international conference on signal processing and communication (ICSPC) (pp. 269–273). Coimbatore. https://doi.org/10.1109/CSPC.2017.8305852
21. Yu, W. D., Kollipara, M., Penmetsa, R., & Elliadka, S. (2013). A distributed storage solution for cloud based e-healthcare information system. 2013 IEEE 15th international conference on e-health networking, applications and services (Healthcom 2013) (pp. 476–480). Lisbon. https://doi.org/10.1109/HealthCom.2013.6720723
22. Zhan, H., & Zhang, W. (2014). Cloud storage-based medical data integration technology. In H. Jeong, S. M. Obaidat, N. Yen, & J. Park (Eds.), *Advances in computer science and its applications. Lecture notes in electrical engineering* (Vol. 279). Springer. https://doi.org/10.1007/978-3-642-41674-3_152
23. Candela-Mendoza, E., Cruz-Ipanaque, L., & Armas-Aguirre, J. (2018). Mobile technology model to ambulatory healthcare information online using cloud platform. In *2018 IEEE XXV international conference on electronics* (pp. 1–4). Electrical Engineering and Computing (INTERCON). https://doi.org/10.1109/INTERCON.2018.8526466
24. Hameed, R. T., Mohamad, O. A., Hamid, O. T., & Tapus, N. (2015). Design of e-Healthcare management system based on cloud and service oriented architecture. 2015 E-health and bio-engineering conference (EHB) (pp. 1–4). Iasi. https://doi.org/10.1109/EHB.2015.7391393
25. Ibrahim, L. F., Sadek, S., Hakeem, S., Al-Sabban, L., Ibrahim Mohammed Ahmed, A., & Hassan Al-Sayed, A. (2014). Enhanced hospital information system by cloud computing: SHEFA'A. In A. Marcus (Ed.), *Design, user experience, and usability. User experience design for everyday life applications and services. DUXU 2014. Lecture notes in computer science* (Vol. 8519). Springer. https://doi.org/10.1007/978-3-319-07635-5_6
26. Liu, W., & Park, E. K. (2013). e-healthcare cloud computing application solutions: Cloud-enabling characteristics, challenges and adaptations. 2013 international conference on computing, networking and communications (ICNC) (pp. 437–443). San Diego, CA. https://doi.org/10.1109/ICCNC.2013.6504124
27. Mourya, A. K., Shafqat-Ul-Ahsaan, & Idrees, S. M. (2020). Cloud computing-based approach for accessing electronic health record for healthcare sector. In A. Chaudhary, C. Choudhary, M. Gupta, C. Lal, & T. Badal (Eds.), *Microservices in big data analytics*. Springer. https://doi.org/10.1007/978-981-15-0128-9_16
28. Reddy C.T., Sirisha, G., & Reddy, A. M. (2018). Smart healthcare analysis and therapy for voice disorder using cloud and edge computing. 2018 4th international conference on applied and theoretical computing and communication technology (iCATccT) (pp. 103–106). Mangalore, India. https://doi.org/10.1109/iCATccT44854.2018.9001280
29. Tawalbeh, L. A., & Habeeb, S. (2018). An integrated cloud based healthcare system. 2018 fifth international conference on internet of things: systems, management and security (pp. 268–273). Valencia. https://doi.org/10.1109/IoTSMS.2018.8554648
30. Aileni, R. M., Pasca, S., & Valderrama, C. (2015). Cloud computing for big data from biomedical sensors monitoring, storage and analyze. 2015 conference grid, cloud & high performance computing in science (ROLCG), Cluj-Napoca (pp. 1–4). https://doi.org/10.1109/ROLCG.2015.7367419
31. Kolodner, E. K., et al. (2011). A cloud environment for data-intensive storage services. 2011 IEEE third international conference on cloud computing technology and science (pp. 357–366). Athens. https://doi.org/10.1109/CloudCom.2011.55

32. Zungeru, A. M., Gorata, M., Mangwala, M., & Chuma, J. (2017). Healthcare cloud computing for underground wireless sensor networks. 2017 2nd international conference on frontiers of sensors technologies (ICFST) (pp. 136–140). Shenzhen. https://doi.org/10.1109/ICFST.2017.8210490
33. Hanen, J., Kechaou, Z., & Ayed, M. B. (2016). An enhanced healthcare system in mobile cloud computing environment. *Vietnam Journal of Computer Science, 3*, 267–277. https://doi.org/10.1007/s40595-016-0076-y
34. Jemal, H., Kechaou, Z., Ayed, M. B., & Alimi, A. M. (2015). Mobile cloud computing in healthcare system. In M. Núñez, N. Nguyen, D. Camacho, & B. Trawiński (Eds.), *Computational collective intelligence. Lecture notes in computer science* (Vol. 9330). Springer. https://doi.org/10.1007/978-3-319-24306-1_40
35. Karaca, Y., Moonis, M., Zhang, Y.-D., & Gezgez, C. (2019). Mobile cloud computing based stroke healthcare system. *International Journal of Information Management*., ISSN: 0268-4012., *45*, 250–261. https://doi.org/10.1016/j.ijinfomgt.2018.09.012
36. NIST, Mell, P., & Grance, T. 2011, September. The NIST definition of cloud computing. NIST Special Publication 800–145. Computer Security Resource Center. https://csrc.nist.gov/publications/detail/sp/800-145/final. Accessed 1 November 2020.
37. Rajabion, L., Shaltooki, A. A., Taghikhah, M., Ghasemi, A., & Badfar, A. (2019). Healthcare big data processing mechanisms: The role of cloud computing. *International Journal of Information Management*., ISSN: 0268–4012., *49*, 271–289.
38. Gartner. (2018). Gartner magic quadrant for public cloud storage services, worldwide. ID: G00340206. https://www.gartner.com/en/documents/3884480/magic-quadrant-for-public-cloud-storage-services-worldwi
39. Evans, M. (2019). Google, Amazon and Microsoft in battle to store health data in the Cloud. Wall Street Journal. https://www.wsj.com/articles/google-amazon-and-microsoft-in-battle-to-store-health-data-in-the-cloud-11568122202. Accessed 5 November 2020.
40. Landi, H. (2020a, September). Microsoft makes big play for healthcare cloud business in competition with Google, Amazon, Fierce healthcare. https://www.fiercehealthcare.com/tech/microsoft-makes-big-play-for-healthcare-cloud-business-competition-google-amazon. Accessed 5 November 2020.
41. Landi, H. (2020b, September). Mayo Clinic taps Google cloud as strategic partner to accelerate innovation in AI, analytics and digital tools, Fierce healthcare. https://www.fiercehealthcare.com/tech/mayo-clinic-taps-google-cloud-as-strategic-partner-to-accelerate-innovation-ai-digital. Accessed 5 November 2020.
42. Miliard, M. (2019, December). Cerner expands AWS relationship with new machine learning initiatives. https://www.healthcareitnews.com/news/cerner-expands-aws-relationship-new-machine-learning-initiatives. Accessed 6 November 2020.
43. Hernandez P. (2020). 20 top enterprise data storage vendors: An overview of leading data storage vendors, based on a variety of enterprise storage needs. https://www.enterprisestorageforum.com/products/20-top-enterprise-data-storage-vendors.html. Accessed 5 November 2020.
44. Delkin Industrial. (2020, July). Delkin blog: Medical hardware flash storage: factors to consider. https://www.delkin.com/blog/medical-hardware-flash-storage-factors-to-consider/. Accessed 5 November 2020.
45. Stenfort R. (2019, December). How Facebook leverages NVMe™ cloud storage in the datacenter, NVM express. https://nvmexpress.org/how-facebook-leverages-nvme-cloud-storage-in-the-datacenter/. Accessed 24 October 2020.
46. Amazon (2020a). Amazon EBS volume types. https://aws.amazon.com/ebs/volume-types/. Accessed 6 November 2020.
47. Amazon (2020b). Data protection in Amazon S3. https://docs.aws.amazon.com/AmazonS3/latest/dev/DataDurability.html. Accessed 4 November 2020.
48. Google. (2020a). Block storage performance. https://cloud.google.com/compute/docs/disks/performance. Accessed 29 October 2020.

49. Google. (2020b, June). 5 ways to enhance your cloud storage security and data protection. https://cloud.google.com/blog/products/storage-data-transfer/5-ways-to-enhance-your-cloud-storage-security-and-data-protection
50. Symonds J. (2020, February). Why the healthcare industry is adopting high performance object storage. https://blog.min.io/why-the-healthcare-industry-is-adopting-high-performance-object-storage/. Accessed 24 October 2020.
51. Microsoft. (2019). Performance tiers for block blob storage. https://docs.microsoft.com/en-us/azure/storage/blobs/storage-blob-performance-tiers. Accessed 6 November 2020.
52. Tso, M. (2018). Forbes Technology Council, May 2018, health care information and the object storage VNA. https://www.forbes.com/sites/forbestechcouncil/2018/05/17/health-care-information-and-the-object-storage-vna/?sh=43e9453a6744. Accessed 24 October 2020.
53. Broad Institute. (2020). Broad Institute and Google Genomics. https://www.broadinstitute.org/google. Accessed 30 October 2020.
54. Kumar, R., & Tripathi, R. (2020) Chapter 2 -Blockchain-based framework for data storage in peer-to-peer scheme using interplanetary file system. Handbook of research on blockchain technology (pp. 35–59). https://doi.org/10.1016/B978-0-12-819816-2.00002-2
55. Shah, B., Shah, N., Shakhla, S., & Sawant, V. (2018). Remodeling the healthcare industry by employing blockchain technology. 2018 international conference on circuits and systems in digital enterprise technology (ICCSDET) (pp. 1–5). Kottayam, India. https://doi.org/10.1109/ICCSDET.2018.8821113
56. Shi, S., He, D., Li, L., Kumar, N., Khan, M. K., & Choo, K.-K. R. (2020). Applications of blockchain in ensuring the security and privacy of electronic health record systems: A survey. *Journal of Computers & Security*., ISSN: 0167-4048., *97*, 101966. https://doi.org/10.1016/j.cose.2020.101966
57. Shahnaz, A., Qamar, U., & Khalid, A. (2019). Using Blockchain for electronic health records. *IEEE Access, 7*, 147782–147795. https://doi.org/10.1109/ACCESS.2019.2946373
58. Mayer, A. H., da Costa, C. A., & Righi, R. d. R. (2020). Electronic health records in a Blockchain: A systematic review. *Health Informatics Journal, 26*(2), 1273–1288. https://doi.org/10.1177/1460458219866350
59. Usman, M., & Qamar, U. (2019). Secure electronic medical records storage and sharing using blockchain technology. *Procedia Computer Science*., ISSN: 1877–0509., *174*, 321–327. https://doi.org/10.1016/j.procs.2020.06.093
60. Zheng, X., Mukkamala, R. R., Vatrapu, R., & Ordieres-Mere, J. (2018). Blockchain-based personal health data sharing system using cloud storage. 2018 IEEE 20th international conference on e-health networking, applications and services (Healthcom) (pp. 1-6). Ostrava. https://doi.org/10.1109/HealthCom.2018.8531125
61. Kassab, M. H., DeFranco, J., Malas, T., Laplante, P., Destefanis, G., & Graciano Neto, V. V. (2019). Exploring research in blockchain for healthcare and a roadmap for the future. In IEEE transactions on emerging topics in computing. https://doi.org/10.1109/TETC.2019.2936881
62. Tariq, N., Qamar, A., Asim, M., & Khan, F. A. (2020). Blockchain and smart healthcare security: A survey. *Procedia Computer Science, 175*(2020), 615–620. https://doi.org/10.1016/j.procs.2020.07.089
63. Esposito, C., De Santis, A., Tortora, G., Chang, H., & Choo, K. R. (2018, January/February). Blockchain: A panacea for healthcare cloud-based data security and privacy? *IEEE Cloud Computing, 5*(1), 31–37. https://doi.org/10.1109/MCC.2018.011791712
64. Michelson, S. (2019). Key considerations for securing your digital healthcare cloud. https://www.idigitalhealth.com/news/. Accessed 6 November 2020.
65. AbuKhousa, E., Mohamed, N., & Al-Jaroodi, J. (2012). E-health cloud: Opportunities and challenges. *Future Internet, 4*(3), 621–645.
66. Dong, N., Jonker, H., & Pang, J. (2011). Challenges in eHealth: From enabling to enforcing privacy, in foundations of health informatics engineering and systems. In Z. Liu & A. Wassyng (Eds.), *FHIES 2011. Lecture notes in computer science* (pp. 195–206). Springer.

67. ENISA, European Network Information Security Agency. (2009). *An SME perspective on cloud computing*. European Network Information Security Agency.
68. Kuo, A. M.-H. (2011). Opportunities and challenges of cloud computing to improve health care services. *Journal of Medical Internet Research, 13*(3), e67.
69. Zhang, R, & Liu, L. (2010). Security models and requirements for healthcare application clouds. In Proceedings of the 2010 IEEE 3rd international conference on cloud computing (pp. 268–275). Miami, FL.
70. Carter, A. B. (2019, July). Considerations for genomic data privacy and security when working in the cloud. *The Journal of Molecular Diagnostics, 21*(4), 542–552. https://doi.org/10.1016/j.jmoldx.2018.07.009
71. Khattak, H. A. K., Abbass, H., Naeem, A., Saleem, K., & Iqbal, W. (2015). Security concerns of cloud-based healthcare systems: A perspective of moving from single-cloud to a multi-cloud infrastructure. 2015 17th international conference on E-health networking, application & services (HealthCom) (pp. 61–67). Boston, MA. https://doi.org/10.1109/HealthCom.2015.7454474
72. Ziglari, H., & Negini, A. (2017). Evaluating cloud deployment models based on security in EHR system. 2017 international conference on engineering and technology (ICET) (pp. 1–6). Antalya. https://doi.org/10.1109/ICEngTechnol.2017.8308142
73. Hombal, U., & Dayananda, R. B. (2019). An effective security protocol for GDS (group data sharing) in healthcare cloud environment. 2019 international conference on communication and electronics systems (ICCES) (pp. 488–492). Coimbatore, India. https:/doi.org/10.1109/ICCES45898.2019.9002033
74. Tupakula, U., & Varadharajan, V. (2014). Trust enhanced cloud security for healthcare services. 2014 IEEE 13th international conference on trust, security and privacy in computing and communications (pp. 344–351). Beijing. https://doi.org/10.1109/TrustCom.2014.46
75. Dean, D. J. et al. (2017). Engineering scalable, secure, multi-tenant cloud for healthcare data. 2017 IEEE world congress on SERVICES (SERVICES) (pp. 21–29). Honolulu, HI. https://doi.org/10.1109/SERVICES.2017.13
76. Alexander, E., & Sathyalakshmi. (2017). Privacy-aware set-valued data publishing on cloud for personal healthcare records. In *Artificial intelligence and evolutionary computations in engineering systems* (pp. 323–334). Springer.
77. Shakil, K. A., Zareen, F. J., Alam, M., & Jabin, S. (2020, January). BAMHealthCloud: A biometric authentication and data management system for healthcare data in cloud. *Journal of King Saud University - Computer and Information Sciences, 32*(1), 57–64. https://doi.org/10.1016/j.jksuci.2017.07.001
78. Al Hamid, H. A., Rahman, S. M. M., Hossain, M. S., Almogren, A., & Alamri, A. (2017). A security model for preserving the privacy of medical big data in a healthcare cloud using a fog computing facility with pairing-based cryptography. *IEEE Access, 5*, 22313–22328. https://doi.org/10.1109/ACCESS.2017.2757844
79. Roy, S., Das, A. K., Chatterjee, S., Kumar, N., Chattopadhyay, S., & Rodrigues, J. J. P. C. (2019, January). Provably secure fine-grained data access control over multiple cloud servers in mobile cloud computing based healthcare applications. *IEEE Transactions on Industrial Informatics, 15*(1), 457–468. https://doi.org/10.1109/TII.2018.2824815
80. Microsoft. (2020, June). Security recommendations for Blob storage. https://docs.microsoft.com/en-us/azure/storage/blobs/security-recommendations