An Overview of Multi-Cloud Computing

Jiangshui Hong, Thomas Dreibholz, Joseph Adam Schenkel, and Jiaxi Alessia Hu

Abstract The purpose of this paper is to provide a brief overview of cloud computing technologies, particularly with respect to multi-cloud networks. First, the basics of cloud computing concepts are discussed. Next we outline some challenges facing cloud computing, and discuss how multi-cloud systems including multi-clouds, hybrid clouds, federated clouds, and cross-clouds may be used to deal with some of these issues. Finally, multi-cloud systems may also be used in conjunction with new developing technologies such as Big Data and Machine Learning, leading to exciting innovations. These are reviewed in brief. Our goal is to provide a modern look at the state of the art in multi-cloud computing and review open issues in the field. The goal is that this paper will help the reader to understand challenges facing cloud computing, how multi-cloud computing addresses some of these issues, and inspire community excitement at the future integration of multi-cloud platforms with other novel technologies.

1 Cloud Computing Basics

The National Institute of Standards and Technology (NIST) defines cloud computation as a "model for granting users access to a shared pool of configurable computational resources including networks, servers, storage, applications, and services that can be rapidly provisioned (adapted for new services) with little management or service provider interaction" [25]. Cloud computing is not a new technology in itself, per se, but rather a new way to incorporate existing technologies [26]. In other words, cloud computing is simply a way to remotely access and manage computing resources [34]. This technology has now been integrated into many businesses, as third-party cloud networks allow firms to focus on core business elements without

Jiangshui Hong, Thomas Dreibholz
SimulaMet, Pilestredet 52, 0167 Oslo, Norway, e-mail: jiangshui@simula.no,dreibh@simula.no

Joseph Adam Schenkel, Jiaxi Alessia Hu
Durham University, Stockton Road, Durham, DH1 3LE, United Kingdom,
e-mail: j.a.schenkel@durham.ac.uk,jiaxi.a.hu@durham.ac.uk

the need to worry about maintenance or the computer infrastructure, essentially turning computation into the “fifth utility” after water, electricity, gas, and telephony [6]. A major advantage of this technology is that computational resources can be used far more efficiently, resulting in quicker results due to economies of scale. [6] illustrates this point by comparing the costs of operating one server for 1000 hours versus 1000 servers for one hour. Though the monetary costs may be the same, the speed of 1000 servers operating in unison to solve a problem greatly reduces the time needed for the task.

Clouds have five essential qualities that make them useful to consumers [25]. First of all, they can be remotely accessed at any point by a single individual without requiring any additional human interaction (on-demand self-service). Additionally, these networks can be used through various interfaces, including mobile phones, laptops, and desktop computers (broad network access). Cloud Service Providers (CSP) also usually rely on multi-tenant models, where resources are pooled to serve multiple customers (resource pooling). Another feature of cloud computing is that these networks can be provisioned easily and released to suit consumer needs (rapid elasticity). Finally, cloud systems feature metering capabilities for their services that can be used for transparent feedback for both provider and user (measured service). These characteristics make cloud computing attractive for small businesses as no upfront investment is needed. Additionally, they also reduce operation costs, are easily accessible and decrease the pressure of having to maintain company networks by outsourcing the infrastructure to providers, allowing a company to focus on core business elements [26].

Currently there are three cloud computing service models that dominate the industry: Software as a Service (SaaS), Platform as a Service (PaaS), and Infrastructure as a Service (IaaS). In SaaS, the consumer can operate a software on the provider’s cloud system, typically over the Internet [25, 26]. In PaaS, the provider allows the consumer to use the cloud network as a platform for their own developed or acquired programs. Finally, in IaaS, the consumer is provided with virtual storage and machines.

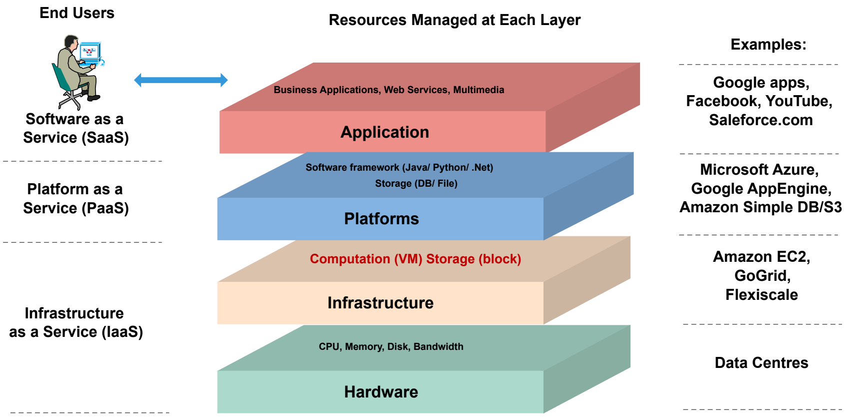
End Users
Resources Managed at Each Layer
Examples:
Business Applications, Web Services, Multimedia
Application
Software as a Service (SaaS)
Google apps, Facebook, YouTube, Saleforce.com
Software framework (Java/ Python/ .Net) Storage (DB/ File)
Platforms
Platform as a Service (PaaS)
Microsoft Azure, Google AppEngine, Amazon Simple DB/S3
Computation (VM) Storage (block)
Infrastructure
Infrastructure as a Service (IaaS)
Amazon EC2, GoGrid, Flexiscale
CPU, Memory, Disk, Bandwidth
Hardware
Data Centres

Fig. 1 Cloud Computing Architecture, based on [26]

The Cloud computing architecture, as shown in Figure 1, consists of four layers, beginning with the Hardware Layer. The Hardware Layer, as the name suggests, contains the physical hardware including servers, switches, power, cooling systems, etc.. Usually, the Hardware Layer takes the form of data centers harbouring thousands of servers. The next level up is the Infrastructure Layer, which operates by using virtualisation technologies to partition compute and storage resources. The Platforms Layer is made up of operating systems and application frameworks, and is mostly used to decrease the strain of deploying applications directly into virtual machine containers. Finally, the Application Layer consists of the cloud computing applications, which refers to the applications (SaaS) that are built on the lower levels (IaaS and PaaS) of the cloud network. For example, the applications that are available on Google Play or other app stores deliver applications that run some of their features in clouds, allowing both ease of access and lowering the computational cost of the app on the device, as operations can be performed in the cloud.

It is important to discuss deployment models, and overall there are four ways in which cloud computing can be deployed. In the private model, the cloud is provisioned for use by a single organisation with multiple users (e.g. business consumers). The organisation may own and oversee the system, rely on a third party for service or some combination of the two. In the community cloud model, multiple organisations may be responsible for cloud provisioning. As with private cloud computing, one or more of the organisations using the community cloud may own and operate it, rely on third parties, or again, some combination of them. A public cloud is one that is accessible by the general public and, unlike the previous two types of cloud computing, they are only available on the premises of the cloud provider (e.g. university campus), while the others may be accessible off premises. Finally, hybrid cloud systems incorporate two or more of the above deployment models which remain separate entities. However, they are unified by an underlying technology that allows for data application portability between the units [25].

2 Issues with Cloud Computing

Now that the basics of cloud computing have been reviewed, it is possible to discuss some of the issues facing this technology. Though cloud computing is useful to businesses as it removes the need for planned provisioning and grants small businesses the ability to scale their computing needs growing in line with their business requirements, there are some issues and challenges facing cloud systems that could be the subject of future research [26]. Though still in its infancy, the widespread use of cloud computing has allowed many researchers and businesses to begin to weigh in on potential issues facing the technology.

Security: One subject of concern cited by many authors is that of data security and privacy [6,15,26,27]. In a study by [15], security concerns were flagged as a key issue in 66 research papers reviewed, with the next highest issue being infrastructure at 46, followed by data management at 15. In this study, practitioners were also interviewed, expressing additional concerns, however still indicated that security was a practical issue, highlighting that this is a pressing matter for many researchers and businesses.

Security issues can be subdivided into several subcategories, including: safety mechanisms, cloud server monitoring, data confidentiality, and avoiding malicious operations [27]. The authors describe problems pertaining to data storage in cloud computing networks, separating it into the components of data integrity, data confidentiality, data availability, and data privacy. Maintaining data integrity involves ensuring that data present within a system cannot be deleted, modified, or fabricated without authorisation. Data confidentiality is another issue as currently it can be difficult to protect against both insider threats and external breaches limiting the utility of cloud networks for sensitive data storage, such as medical records or government files [33]. Another dimension of security is data availability, or the degree to which data can be recovered. Finally, data privacy refers to the ability of an individual or group to privately and selectively share information only amongst themselves.

Despite these concerns, [6] argues that – given the proper preparation – cloud computing may actually become more secure than other methods. The authors explain that any obstacles facing cloud computing are also present in traditional systems and can be solved using technologies that are already used, such as data encryption, virtual local area networking, firewalls, and more.

Legal Concerns: Legal issues, while related to security issues, are another distinct concern [15, 27]. Data storage rules and laws vary depending on the location [6]. Although service level agreements between providers and consumers have been established, there are currently no standards in place. Adopting cloud computing storage rules to fit local law may be a challenge in the future, and will require careful planning and coordination when designing these systems.

Data Management: Perhaps an obvious concern to many is what happens to the data on cloud computing systems in the event of a catastrophic loss of data. If, for example, a cloud provider company were to go bankrupt, the data may become irretrievable. This would present a huge problem for large organisations that are dependent on data storage [15,33].

Interoperability: [23] considers interoperability, or the "ability of diverse systems and organisations to work together (inter-operate)" to be the second-greatest challenge to cloud computing systems after data security and trust issues. This is crucial, because it would allow users to avoid vendor lock-in, whereby the user becomes dependent on the CSP because they cannot extract their services to other platforms or clouds. The lack of standardisation means that it can be difficult to transfer data between cloud systems.

High Latency: Another issue facing cloud computing is the problem of high latency, or a lag from when the transfer of data begins following an instruction for its transfer. In the context of cloud computing, this is caused by the necessity of the various nodes in the cloud to communicate with each other. Two solutions have emerged as potential solutions to this problem: fog computing and edge computing. Fog computing may be defined as "a scenario where a huge number of heterogeneous (wireless and sometimes autonomous) ubiquitous and decentralised devices communicate and potentially cooperate among them and with the network to perform storage and processing tasks without the intervention of third parties." [31]. Another suggested fog definition is that proposed by [35]. These authors state that fog computing is "a geographically distributed computing architecture with a resource pool that consists of one or more ubiquitously connected heterogeneous devices at the edge of network and not exclusively seamlessly backed by cloud services. This network collaboratively provides elastic computation, storage and communication in isolated environments to many clients within proximity". To combat latency [17] suggested a fog-based opportunistic spatio-temporal event processing system which predicts future query region for moving consumers. It is also important to note that there are several types of fog networks. Edge computing is somewhat similar and refers to distributed computing paradigms where computation is largely performed on *edge devices* (such as smartphones, and laptops) as opposed to on the central cloud network itself. Computations are shifted to the periphery of the system.

Vendor Lock-in: Numerous authors have also discussed the problem of vendor lock-in, the situation where the client becomes dependent on the provider, making shifting provider costs expensive, as it is complicated by legal issues, or the technical incapability of transfer due to incompatibility issues. This issue is related to the interoperability problem, due to the presence of several cloud systems, many of which are incompatible with one another. [21] explains that a large part of this issue stems from the lack of standardisation with reference to operating platforms, Application Programming Interfaces (API), Service Level Agreements (SLA) and cloud semantics causing interoperability (see above) issues. This makes it difficult for a company to shift its resources from one cloud to another, causing a situation where a company may become dependent on a single provider. Although some businesses have already shifted towards cloud technology, the vendor problem poses a considerable barrier to others due to a lack of trust in a single CSP [20]. For instance, if a business were to hire a CSP, and that provider were to go bankrupt, then the business could potentially lose all the data on that cloud, making it a considerable risk. Solutions to this problems are varied, but one example from recent literature is the standardisation of edge computing [28]. In 2017, the EdgeX Foundry project was launched with the intent of developing a vendor-neutral framework for Internet of Things (IoT) computing. Meanwhile the Open-Fog consortium has also been launched as a collaboration between Princeton University and several notable tech companies including Dell and Microsoft in an attempt to standardise fog computing. Goal of MELODIC (Multi-cloud Execution-ware for Large-scale Optimised Data-Intensive Computing) [11, 12] is to provide a vendor-independent middleware to base cloud applications on.

3 Multi-Cloud Solutions

To help deal with some of the issues mentioned above, the next logical progression in cloud computing is in multi-cloud computing, or cloud systems that utilise numerous cloud networks and services simultaneously [13]. In short, multi-cloud systems simply use more than one CSP, though they come in many subcategories that will be described below. This is similar to the definition of cross-cloud architectures, defined as systems that span across multiple provisioning boundaries [14]. Though the exact distinction between multi-clouds and cross-clouds is lacking (as described in further detail below), it can be understood that in multi-clouds, the user or business in question utilizes different cloud services for different applications in their business. For example, they may store data on a private cloud, share documents on the Google Cloud platform, and perform data analysis on yet another cloud. On the other hand, cross-cloud architectures are designed to make the transfer of data and utilisation of apps across the clouds more streamlined and cohesive [14].

There are various rationales for using multi-clouds or cross-clouds, many of which address some of the issues mentioned in the previous section. According to [24], multi-cloud computing directly addresses 10 key issues, some of which have been mentioned above. Furthermore, it is important to realise that these issues are both shared and specific to the parties involved, including individual customers, company level customers, and the CSPs. The 10 key issues addressed by multi-cloud computing are:

1. Dealing with peaks in service/resource requests using external ones on demand;
2. Optimising costs or improving quality of services;
3. Reacting to changes of the offers by the providers;
4. Following constraints such as new locations or laws;
5. Ensuring the high availability of resources and services;
6. Avoiding the dependence on only one external provider;
7. Ensuring backups to deal with disasters or scheduled inactivity;
8. Acting as intermediary;
9. Enhance own Cloud service/resources offers, based on agreements with others;
10. Consuming different services for their particularities not provided elsewhere.

For example, if a company were to provide an online service, and overestimated how much time the users are spending on their site or application, it may be beneficial to scale down by shifting the infrastructure from lifelong virtual machines to ephemeral virtual machines, which charge by the minute instead of by the hour, and therefore reducing costs. In another case, a company manager may find that different branches or teams within the company rely on different private or public cloud infrastructures. So, consolidating the teams would require substantial overhauls to the underlying logic. In short, changes to the predetermined plan pertaining to service provisioning, use, or management are expected challenges to businesses as they grow, and cross-cloud computing is necessary to deal with the alterations [13, 24].

Another advantage of cross-cloud computing is the avoidance of long-term commitment to a single CSP, tackling issues of interoperability and vendor lock-in. Because cross-cloud platforms often rely on communicating between their different cloud components, they may generate new methods of operability, either through increased standardisation of systems used, or by devising new ways for clouds to share data with one another on a more universal level. Developing technologies that would allow the transfer of services to other CSPs would help with this problem and allow for greater flexibility within the business [13].

Before proceeding, it should be noted that a significant problem in the field of cloud computing models is a degree of ambiguity present in the literature. As mentioned above, standardisation is an issue. However, it also pertains to the definitions surrounding multi-cloud computing. Some authors, such as [13] categorise multi-clouds, hybrid clouds, and federation clouds all under cross-cloud computing. On the other hand, [24] places all of the previous models under "multiple cloud" computation, differentiating between multi-clouds and federation cloud frameworks. Most recently, [32] classified hybrid and federation infrastructures under multi-cloud computing, in line with the classification by [24]. For the purposes of this document, hybrid clouds, multi-clouds, and federation clouds will be treated as distinct methods to avoid confusion. Future work may benefit from more concrete definitions such as those provided by the NIST for cloud computing as a whole [25].

Multi-Cloud: Multi-cloud can be defined as cloud systems in which applications are hosted as chunks among a heterogeneous network of different cloud. However, with multi-cloud the components are all unique cloud systems, not deployment methods, as is the case with hybrid clouds (see below). As discussed above, some have used "multi-cloud" as an umbrella term to describe some of the other multi-cloud systems described below, leading to some ambiguity. So, for the definition presented here, multi-clouds will specifically refer to a cloud system where multiple different cloud networks are combined for different roles, with the aim of reducing the necessary trust requirements among the CSPs [24, 30]. In other words, in multi-cloud networks, unique CSPs are used for a single business or organisation needs, and they may all have differing levels of application as well as differing SLAs. Again, an example of this may be a company that uses a small amount of one cloud to store or send documents, another private cloud for sensitive company data, and yet another for data analysis. All the clouds in this case may have different SLAs, different costs, and different degrees of utilisation by the company.

The reasons for using the multi-cloud approach are varied. For example a business may seek to use a multi-cloud approach for legal considerations, where the business may operate on one cloud, but requires a different cloud to legally store data. An- other reason is to avoid using a single CSP and [9] proposes an architectural framework and principles for Programmable Network Clouds hosting Software Defined Networking (SDN) and Network Function Virtualisation (NFV) for geographically distributed multi-cloud computing environments. Costs and SLA-aware resource provisioning and scheduling that minimised the operating cost without violating the negotiated SLAs were investigated with respect to techniques for autonomic and timely NFV composition, deployment and management across multiple Clouds. The authors strongly suggested for future works to investigate optimisation techniques that simultaneously optimise the VM/container placement and traffic consolidation, maximising both SLA satisfaction and reducing costs [9].

Arguably the most useful and well-researched application of multi-cloud architecture offer enhanced security. [5, 7, 18, 30] describe several multi-cloud architectures with a primary focus on security, specifically with medical records. In this paper, the authors discuss four mechanisms for enhanced security. The first is application imitation, where the clouds within the multi-cloud essentially "double-check" each other, as results from one cloud can be compared with those in a different cloud, ensuring data integrity. Layer-wise application partition allows for a separation between logic and data, protecting against flaws in logic. The application provider cannot recreate logic entirely from user data unless it is entirely executed on the cloud users system. Partitioning into segments refers to splitting the application into smaller segments, which are then run on different clouds, and this process involves two phases: in the first phase, trusted private clouds receive small computational parts, while untrusted public clouds handle higher loads. In the second phase, the computation is shared amongst the untrusted clouds. The final mechanism for increasing security is the distribution of chunks, or simply that data is split amongst the different clouds within the multi-cloud. In this case, different methods for cryptographic data segmentation have been trialed for the storage and retrieval of data.

Several models have been suggested for these mechanisms. For example, a cooperative provable data possession (CPDP) scheme was tested by [36]. However, it had a problem in that the security parameter π' is independent of other parameters, allowing authentication bypass via forgery of π' in the response sequence [18]. More recently, [18] have expanded on integrity checking methods using the Co-Check scheme, which is based on Boneh-Lynn-Shacham (BLS) signature and homomorphic tags [18]. As defined by [8], homomorphism is a property by which a problem in one algebraic system can be converted to a problem in another algebraic system. The problem can then be solved and converted back, allowing a third party to view the encrypted information and use it, if it said party had the right tools. Homomorphic encryption in other words, is the ability to perform operations or calculations on encrypted data without the need to decrypt it first. In the Co-Check framework, users do not have to retrieve the entire data file during the challenge-response and integrity checking stages of the model (stages II and III). Instead the users generate the challenges for audition by using parts of the metadata restored at the client side to prompt the audition efficiency and ensure that malicious CSPs cannot bypass the check [18]. Other authors have also used homomorphic encryption tools for multi-cloud security. One of the future research directions for multi-cloud models will be to experiment with security, as homomorphic encryption is still in infancy with cloud computing application.

Hybrid Cloud: Hybrid clouds are probably the most common type of multi-cloud framework, and are defined as a subset of multi-cloud networks in which private cloud deployment methods are combined with one or more public ones. Unlike a multi-cloud, the deployment methods are not entirely unique clouds. Hybrid clouds are useful to businesses aiming to share some of their data across various user types, yet still desire to keep certain elements confidential [3]. For example, one area where hybrid clouds are currently being explored is in the field of medicine, where big data is becoming increasingly available. Recently, [19] explored security solutions for healthcare data using linear network coding and re-encryption based on ElGamal cryptography in the form of hybrid approach. To provide security and fault tolerance for cloud storage, the authors considered a linear network coding mechanism and used the ElGamal re-encryption scheme to encode the key matrix securely. One of the key distinctions between hybrid clouds and multi-clouds is the nature of the work performed on the systems. Because multi-clouds use distinct clouds, they can often be used to handle a variety of tasks and services for a company. On the other hand, hybrid clouds are more unified in their framework and usually used to handle one task, such storing and distributing medical records, in a secure fashion.

Federated Cloud Computing: The key distinction for federated cloud computing is that the CSPs agree to share resources. In this system, CSPs share the client load under a single umbrella federation. This grants the client access to a potential catalogue of services and resources available and increases the interoperability as well as the portability of applications. [32] gives the example of the EU-based EGI Federated Cloud, which unites more than 20 cloud providers and 300 data centers. Federated clouds can address the vendor lock-in problem in that applications. Furthermore, data can be migrated from one cloud to another more easily than in a pure multi-cloud system, due to the underpinning SLAs. Federated cloud computing likely benefits smaller providers, as they allow to cover each other's weaknesses.

Cross-Cloud Computing: According to [14], a cross-cloud application is one that utilizes more than one cloud API under a single version of the application. Cross-cloud architecture may use either dynamically interchangeable APIs or non-dynamically interchangeable APIs, so long as the APIs in question are divergent. The literature referring to cross-cloud systems is sparse when compared to the frameworks mentioned above, and this may be due to the lack of specificity of definition. Even in the [14] paper referenced above, the author defined hybrid-clouds, multi-clouds, and federated clouds as subclassifications of cross-cloud computing, not multi-cloud computing, raising questions as to what the actual distinction between the two is, and what constitutes a "cross-cloud".

Regardless of this ambiguity, recent work proposes different cross-cloud architectures for different ends, and in some cases merges the idea of cross-cloud systems with some of the other multi-cloud models described above. One idea that appears to pervade cross-cloud is the notion of the "cloud broker" – an entity that manages the use, performance, and delivery of cloud services, and negotiates relationships between CSPs and consumers [25]. [2] suggested a broker-based cross-cloud federation manager as a business model for federated cloud networks. More recently, cross-cloud computing has been highlighted as a potential solution for the problem of Big Data management in the IoT [29].

4 Research Directions and Considerations for the Future

In the sections above, we have outlined how multi-cloud computing frameworks can address some of the issues facing cloud computing as a whole, such as data security and vendor lock-in avoidance. That being said, there are several other applications for multi-cloud computing networks. Two of the major themes currently being explored in the literature are cloud computing and Big Data as well as Machine Learning. The following, we will describe some of the challenges and current projects in these fields.

Big Data: Though data management has always been an issue with cloud computing, the looming goal of processing vast quantities of data is one of the major topics facing current research. The most obvious application of big data is again, medical, where vast quantities of patient information will need to be stored and processed over large cloud computing networks. This problem is compounded when medical wearables are introduced, gathering and sending even more information into the relevant cloud networks, and decentralising data sources. This new interconnectivity between household electronics, smart devices, computers, vehicle monitoring equipment and any other devices that allow for the collection and exchange of data is referred to as the IoT, and presents a large hurdle for data analysis, as the data will likely come from a multitude of sources. Due to the sheer volume of data that either currently exists or is being generated, traditional analytic methods cannot keep up [29]. This data must be processed efficiently, analysed (horizontally and vertically), and transferable between different cloud systems. For example, in the case of medical records, a centralised system may work well within a nationalised healthcare system. However, they must also be able to account for peripheral services, such as private practices, which may use alternative cloud systems. It is important that these private firms possess the ability to feed medical information into the larger, more mainstream sources.

One solution to the problem of big data storage proposed in 2012 is the idea of the "rain cloud", a multi-cloud model where each member cloud completes an SLA with other member clouds that allow them to work together when data gets too large for any single cloud to handle [9]. Hybrid clouds are also thought to be useful for big data management, as the combination between private and public deployment allows for confidentiality when different parties require differing (unequal) levels of access to the available data. In many cases businesses may require portions of their data and infrastructure to remain behind the corporate firewall due to industry standards, legal regulations, or their own desire for privacy [22]. Other approaches involve divvying up data amongst clouds for rapid data storage. In Hadoop, an open source implementation of Google's method for data storage, large volumes of data are cut into more manageable chunks across thousands of computers. A parallelised programming API is then used to distribute the computations to where the data is located and to aggregate the results. This method has tremendous application for bioinformatics and genomics. However, the technical savvy and lack of bioinformatics tools that run in parallel has proven to be a barrier up until very recently [4]. They reviewed various big data cloud computing frameworks including Hadoop, Spark and Flink. It will be interesting to see what further improvements will be made and how multi-cloud computing architectures will integrate these approaches. Finally [29] also mentions the use of Amazon Elastic Map Reduce, IBM BigInsights, and Microsoft Azure HDInsight as methods for allowing accessible large-scale data processing frameworks, namely Apache Hadoop, and Apache Spark.

Machine Learning: Although related to big data, machine learning presents its own sets of challenges and applications for use with cloud computing technologies. Given the availability of data stored in clouds (e.g. photographs and videos), there is a wealth of information that can be processed by machine learning algorithms. Techniques such as virtual machine optimisation for healthcare services are being trialed and studied, and show promise, but must be trialed with other diseases to demonstrate consistency [1]. [16] has explored the use of hybrid cloud storage and machine learning for use in the oil and gas industry, where technical documents contain valuable information from disciplines like geoscience and engineering and are in general stored in a unstructured format. To improve data extraction and utilisation, the authors propose a machine-learning-enabled platform, consisting of a carefully selected sequence of algorithms, developed as a hybrid cloud container that automatically reads and understands the technical documents with little human supervision. The user can upload raw data to the platform, which is stored on a private local server. Structured data is generated as output, which is pushed through to a search engine that is accessible to the user via the cloud. This allows for a user to quickly identify the most important parts of technical documents, automate the extraction or relevant data from the documents, present the data in a meaningful way for further analysis, and finally, allows the user to share it easily and port it to other platforms.

Future directions for machine learning and multi-cloud computing research include a shift towards the wide-scale adoption of auto-tuners, especially for the SaaS layer of the Cloud. [10] also anticipate the advent of new automated tools for cloud users to benefit from the experience of other users through partially automated application builders, automated database sharers, query optimisers, or smart load balancers and service replicators. In other words, the interface between cloud, application, and user, is expected to become more streamlined and intuitive based on what is learned from the user experience. As security has always been a key concern for cloud networks of any kind, novel machine learning methods may be used for increased security measures. However, the authors did not specify what these may look like or on what they would be based, representing a future area of research application.

5 Conclusion

Cloud computing is a blossoming technology with numerous applications in many industries, including remote computing and storage. Though vendor lock-in and cyber security are major concerns hybrid clouds, multi-clouds, and federation clouds may address some of these problems by providing users with alternatives in the case of scheduled maintenance, breaches, or shut-downs, though each has their own benefits and disadvantages. Hybrid systems are easily customised to a given application, however less transferable, and often used for only one task whereas multi-clouds and federated clouds are suited for businesses that require multiple tasks or services. Future work should incorporate multi-cloud paradigms and combine them with other technologies such as machine learning and big data, as these technologies can either be used to solve some of the current issues with cloud computing, such as increasing security, or can be used for entirely new methods of analysis.

References

1. Abdelaziz, A., Elhoseny, M., Salama, A.S., Riad, A.M.: A Machine Learning Model for Improving Healthcare Services on Cloud Computing Environment. Measurement **119** (2018)
2. Abdo, J.B., Demerjian, J., Chaouchi, H., Barbar, K., Pujolle, G.: Broker-based Cross-Cloud Federation Manager. In: 8th International Conference for Internet Technology and Secured Transactions (ICITST-2013). London/United Kingdom (2013)
3. Aggarwal, R.: Resource Provisioning and Resource Allocation in Cloud Computing Environment. International Journal of Scientific Research in Computer Science, Engineering and Information Technology **3** (2018)
4. Aisling, O., Jurate, D., D., S.R.: Big data, Hadoop and Cloud Computing in Genomics. Journal of Biomedical Informatics **46**(5) (2013)
5. AlZain, M.A., Soh, B., Pardede, E.: A Survey on Data Security Issues in Cloud Computing: From Single to Multi-Clouds. Journal of Software **8**(5) (2013)
6. Armando, F., Rean, G., Anthony, J., Randy, K., Andrew, K., Gunho, L., David, P., Ariel, R., Ion, S.: Above the Clouds: A Berkeley View of Cloud Computing. Tech. Rep. UCB/EECS-2009-28, EECS Department, University of California, Berkeley (2009)
7. Awasthi, P., Mittal, S., Mukherjee, S., Limbasiya, T.: A Protected Cloud Computation Algorithm Using Homomorphic Encryption for Preserving Data Integrity. In: Recent Findings in Intelligent Computing Techniques. Singapore (2019)
8. Benzekki, K., Fergougui, A.E., Alaoui, A.E.B.E.: A Secure Cloud Computing Architecture using Homomorphic Encryption. International Journal of Advanced Computer Science and Applications **7**(2) (2016)
9. Buyya, R., Son, J.: Software-Defined Multi-Cloud Computing: A Vision, Architectural Elements, and Future Directions. In: Proceedings of the 18th International Conference on Computational Science and Applications (ICCSA) (2018)
10. Buyya, R., Srirama, S.N., Casale, G., Calheiros, R., Simmhan, Y., Varghese, B., Gelenbe, E., Javadi, B., Vaquero, L.M., Netto, M.A.S., Toosi, A.N., Rodriguez, M.A., Llorente, I.M., Vimercati, S.D.C.D., Samarati, P., Milojicic, D., Varela, C., Bahsoon, R., Assuncao, M.D.D., Rana, O., Zhou, W., Jin, H., Gentzsch, W., Zomaya, A.Y., Shen, H.: A Manifesto for Future Generation Cloud Computing: Research Directions for the Next Decade. ACM Comput. Surv. **51**(5) (2018)
11. Dreibholz, T.: Big Data Applications on Multi-Clouds: An Introduction to the MELODIC Project. Keynote Talk at Hainan University, College of Information Science and Technology (CIST) (2017)
12. Dreibholz, T., Mazumdar, S., Zahid, F., Taherkordi, A., Gran, E.G.: Mobile Edge as Part of the Multi-Cloud Ecosystem: A Performance Study. In: Proceedings of the 27th Euromicro International Conference on Parallel, Distributed and Network-Based Processing (PDP). Pavia, Lombardia/Italy (2019)

13. Elkhatib, Y.: Mapping Cross-Cloud Systems: Challenges and Opportunities. In: Proceedings of the 8th USENIX Conference on Hot Topics in Cloud Computing. Berkeley/United State (2016)
14. Elkhatib, Y., Blair, G.S., Surajbali, B.: Experiences of Using a Hybrid Cloud to Construct an Environmental Virtual Observatory. In: Proceedings of the 3rd International Workshop on Cloud Data and Platforms. Prague/Czech Republic (2013)
15. Ghanam, Y., Ferreira, J., Maurer, F.: Emerging Issues and Challenges in Cloud Computing – A Hybrid Approach. Journal of Software Engineering and Applications **5**(11A) (2012)
16. Hernandez, N.M., Lucañas, P.J., Graciosa, C.M.J.C., Caezar, I.P.L., Maver, K.G., Yu, C., Maver, M.G.: An Automated Information Retrieval Platform For Unstructured Well Data Utilizing Smart Machine Learning Algorithms Within A Hybrid Cloud Container. In: First EAGE/PESGB Workshop Machine Learning (2018)
17. Hong, K., Lillethun, D., Ramachandran, U., Ottenwälder, B., Koldehofe, B.: Opportunistic Spatio-temporal Event Processing for Mobile Situation Awareness. In: Proceedings of the 7th ACM International Conference on Distributed Event-based Systems (2013)
18. Mao, J., Tian, W., Zhang, Y., Cui, J., Ma, H., Bian, J., Liu, J., , Zhang, J.: Co-Check: Collaborative Outsourced Data Auditing in Multicloud Environment. Security and Communication Networks **2017** (2017)
19. Modi, K.J.: Securing Healthcare Information over Cloud Using Hybrid Approach. In: P.C. Rani, P.A. K., M. Sudip, P. Bibudhendu, L. Kuan-Ching (eds.) Progress in Advanced Computing and Intelligent Engineering. Springer, Singapore (2019)
20. Opara-Martins, J.: Taxonomy of Cloud Lock-in Challenges. In: M. Khatib, N. Salman (eds.) Mobile Computing. IntechOpen, Rijeka (2018)
21. Opara-Martins, J., Sahandi, R., Tian, F.: Critical Analysis of Vendor Lock-in and its Impact on Cloud Computing Migration: A Business Perspective. Journal of Cloud Computing **5**(1) (2016)
22. Ouyang, C., Moura, M.: A Vision of Hybrid Cloud for Big Data and Analytics. Tech. rep., IBM Big Data and Analytic Hub (2017)
23. Petcu, D.: Portability and Interoperability between Clouds: Challenges and Case Study. In: Towards a Service-Based Internet. Springer, Berlin/Heidelberg (2011)
24. Petcu, D.: Multi-Cloud: Expectations and Current Approaches. In: Proceedings of the International Workshop on Multi-cloud Applications and Federated Clouds. Prague/Czech Republic (2013)
25. Peter, M., Tim, G.: The NIST Definition of Cloud Computing. Tech. rep., National Institute of Standards and Technology Gaithersburg, MD 20899-8930 and U.S. Department of Commerce (2011)
26. Qi, Z.: Cloud Computing: State-of-the-Art and Research Challenges. Journal of Internet Services and Applications **1**(1) (2010)
27. Sun, Y., Zhang, J., Xiong, Y., Zhu, G.: Data Security and Privacy in Cloud Computing. International Journal of Distributed Sensor Networks **10**(7) (2014)
28. Taherizadeh, S., Jones, A.C., Taylor, I., Zhao, Z., Stankovski, V.: Monitoring Self-Adaptive Applications within Edge Computing Frameworks: A State-of-the-Art Review. Journal of Systems and Software **136** (2018)
29. Taherkordi, A., Zahid, F., Verginadis, Y., Horn, G.: Future Cloud Systems Design: Challenges and Research Directions. IEEE Access **6** (2018)
30. Thillaiarasu, N., ChenthurPandian, S.: Enforcing Security and Privacy over Multi-Cloud Framework using Assessment Techniques. In: Proceedings of the 10th International Conference on Intelligent Systems and Control (ISCO). IEEE, Coimbatore/India (2016)
31. Vaquero, L.M., Rodero-Merino, L.: Finding Your Way in the Fog: Towards a Comprehensive Definition of Fog Computing. SIGCOMM Comput. Commun. Rev. **44**(5) (2014)
32. Varghese, B., Buyya, R.: Next Generation Cloud Computing: New Trends and Research Direction. Future Generation Computer Systems **79** (2018)
33. Vurukonda, N., Rao, B.T.: A Study on Data Storage Security Issues in Cloud Computing. Procedia Computer Science **92** (2016)
34. Winans, T.B., Brown, J.S.: Cloud Computing: A collection of working papers. Deloitte LLC (2009)
35. Yi, S., Li, C., Li, Q.: A Survey of Fog Computing: Concepts, Applications and Issues. In: Proceedings of the 2015 Workshop on Mobile Big Data. ACM, Hangzhou/China (2015)
36. Zhu, Y., Hu, H., Ahn, G.J., Han, Y., Chen, S.: Collaborative Integrity Verification in Hybrid Clouds. In: 7th International Conference on Collaborative Computing: Networking, Applications and Worksharing (CollaborateCom). IEEE, Orlando/USA (2012)