Fog Computing: Will it be the Future of Cloud Computing?

Mohamed Firdhous[1], Osman Ghazali[2] and Suhaidi Hassan[3]
[1]University of Moratuwa, Moratuwa, Sri Lanka
[1]Mohamed.Firdhous@uom.lk
[2,3]Universiti Utara Malaysia, Sintok, Kedah, Malaysia
[2]osman@uum.edu.my, [3]suhaidi@uum.edu.my

ABSTRACT
Cloud computing is the newest computing paradigm that makes computing resources available over the Internet on a utility costing basis. Cloud computing offers many advantages to users in terms of reduced cost, elimination of system administrative functions, increased flexibility, better reliability and location independence. Though these are definite advantages, cloud computing also suffers from certain limitations. These limitations arise from the very same reason that is considered an advantage too. Hosting of cloud data centres in the Internet creates large and unpredictable network latencies and undefined security issues as sensitive data is now entrusted to a third party. Also location independence of processing in cloud computing may also not desirable for certain types of networks such as sensor networks and Internet of Things. These services are known as location aware services and require location dependent fast processing. In order to overcome these limitations, researchers have proposed a new cloud computing model called fog computing where the cloud system is located either at the edge of the private network or very close to it. In this paper, the authors take an in depth look at both these technologies to investigate fog computing can reliably overcome the limitations of cloud computing and effectively replace it and become the de facto cloud computing model of the future.

KEYWORDS
cloud computing, fog computing.

1 INTRODUCTION
Cloud computing has gained the attention of both users and service providers as the most promising computing paradigm, no other computing paradigms that came before enjoyed [1]. Cloud computing makes computing resources such as hardware, application development platform and computer applications available as services over the Internet. The services made available this manner are commonly known as Infrastructure as a Service (IaaS), Platform as a Service (PaaS) and Software as a Service (SaaS) [2]. Figure 1 shows the cloud computing business layers as how they are stacked on each other. The bottom two layers namely the physical hardware layer and the virtualized hardware layer provide the required platform along the with necessary security and isolation for multiple systems to run simultaneously.

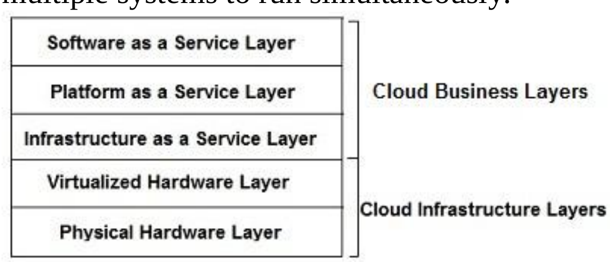

Figure 1 Layers of Cloud Computing

Cloud computing provides many advantages to users compared to traditional purchase, own and run your own computing systems model. The main advantages of cloud computing over the traditional computing paradigm are the economic advantages and the elimination of computer systems administrative tasks and associated costs. Users can access the cloud services and pay for only the services accessed on a utility costing basis [3]. Utilities such as electricity, water, gas and telephony are used by

ISBN: 978-1-941968-00-0 ©2014 SDIWC

customers irrespective of the nature and place of service generation and pay only for the amount of services received (or consumed). Similar to this, cloud computing services are also priced on a per unit time basis irrespective of how much of resources were used. As all the computing hardware and software is hosted in a remote data centre owned and operated by a service provider, the clients can only concentrated on their core business functions. Even the customers involved in computing related businesses such as software development, need not worry about purchasing their own hardware and software such as development platform or project management tools or managing them in house.

Though cloud computing has so many advantages, it also suffers from certain shortcomings too. These shortcomings include the requirement of high capacity (bandwidth) client access link, high latency and security [4],[5],[6]. These limitations have heavy impacts on certain kinds of computing needs such as sensor networks and especially the emerging Internet of Things paradigm that envisages to have every device on the Internet [7]. In order to overcome the limitations of cloud computing to meet the demands of emerging computing models and paradigms, new kind of cloud computing model has been proposed by researchers. in this new model, the devices that respond to and process the client request are hosted either at the edge of the local network or very close to it rather than far away in an unknown location in the middle of the Internet cloud. This kind of cloud computing model has been given the name "Fog Computing". In this paper, we embark on a journey to see how this proposed new model of cloud computing can reduce or eliminate the limitations of the current cloud computing model and respond to the changing demand of the users and emerging computing paradigms.

This paper consists of five sections as follows: Section 1 introduces the paper while Sections 2 discusses cloud computing especially concentrating on the issues related to cloud computing. Section 3 introduces Fog Computing the emerging cloud computing model along with its features, advantages and disadvantages. Sections 4 presents an in-depth analysis and cloud computing and fog computing based with respect to their capabilities. Finally Section 5 concludes by summarizing the findings and some suggestion for the way forward.

2 CLOUD COMPUTING

Cloud computing has transformed the way computing resources have been procured, used and paid for. Prior to the arrival of cloud computing, computing resources including hardware and software were purchased outright and installed in house at a data centre maintained by the organization or hardware was leased from a public data centre on a fixed charges [8]. The hardware resources leased from data centres were of fixed capacity similar to the hardware installed in house irrespective of the usage. The downside of this kind of arrangement is, most of the time hardware thus purchased or leased idle due to under loading wasting precious financial capital that could be invested on some other resource or the core business operations [9]. If the hardware was undersized, the performance would suffer during high demand periods resulting in unsatisfied customers again affecting the profitability and business in the long run. Hence both over capacity as well under capacity affects the profitability of business operations. On the other hand hardware leased from a cloud service provider does not require any upfront investment and the charges are based on usage. Hence the customer receives nearly 100 percent return on his investments (payments). As there is no initial investment on computing resources and only usage charges need to be paid, cloud computing effectively turns the capital expenditures (CapEx) into operational expenditures (OpEx) releasing the limited financial capital to be invested on core business

ISBN: 978-1-941968-00-0 ©2014 SDIWC

activities improving the operational capacity of businesses [10]. In addition to releasing the capital, cloud computing also protects the users from resource starvation as the cloud based resources are elastic that can respond to the demands dynamically [11]. The dynamic resource provisioning property of cloud computing allows the resource provisioning to follow the demand pattern dynamically even during short term fluctuations by adding and removing virtual systems [12].

The other advantage of cloud computing is its layered architecture that allows customers to purchase services at different levels of abstractions commonly known as IaaS, PaaS and SaaS [2]. Like in traditional leasing of computing resources from a public data centre, customers are not required to lease hardware and install their entire software suits in the leased systems. Customers can purchase computing resources at hardware level, platform level or application level depending on their requirements. Purchasing resources at hardware level as IaaS provides the complete freedom and control of the system to the user including the type of operating system to end user applications to be installed. PaaS provides the application developers a ready-made platform comprising operating system, development and testing tools, project management applications and deployment tools that can be customized to suit their requirements. Hence, the applications developers can get on with their jobs with minimal delay as all the required tools are readily available. On the other hand SaaS provides the customers ready-made web based applications that can be personalized and used almost immediately. PaaS and SaaS also relive the customers from the hassle of purchasing and managing software licenses [13].

Cloud computing also provides location independence as a user can access the system or application from anywhere with an Internet connection and a standard web browser [14]. The location independence provides the additional benefit of widespread collaboration between users across large geographical area irrespective of where they are physically residing or working from. This feature is very handy for application developers as it is possible to have a multinational workforce working on a single project without leaving their place of origin.

Cloud systems can also be considered to be more reliable than in-house maintained systems as the backup and disaster management facilities are generally a part of the service provider's offerings [14]. The availability and operation of backup and disaster management systems are transparent to the ultimate users. Setting up of redundant sites for the purpose of backup and disaster management is relatively easier and less costly as the total cost of this infrastructure can be spread across a large customer base.

2.1 Limitations of Cloud Computing

Though cloud computing provides several advantages to both users and service providers compared to traditional computing paradigms, it has certain limitations as well [6]. The prominent limitations of cloud computing include requirement of high speed reliable Internet connectivity and sometimes multi-homing to avoid link outages, high latency, undefined security etc [15],[16],[17]. The emerging trends in networking such as large distributed Internet connected sensor networks, Internet of Things (IoT), mobile data networks and also real time streaming applications have characteristics that cannot be satisfied by cloud computing [7].

Since cloud computing is basically Internet based computing, it is a must to have reliable Internet connectivity with sufficient bandwidth to access the services. If the link outage occurs due to any reason the total system would be unreachable making a total blackout. Multi-homing through multiple Internet connectivity can reduce the effect of link outages, but it is very expensive and technically more involved in setting up multi-homing computer networks [18].

Since cloud systems are located within the Internet, which is a large heterogeneous network with numerous types, topologies, speeds and technologies with no central control [19]. Because of the non-homogeneous and loosely controlled nature of the Internet, there are many issues especially quality of service related ones remain unresolved. One such issue that affects the quality of service severely is network latency [20]. Real time applications with which users directly interact with are badly affected by delay and delay jitter caused by latency in networks [21]. It is very difficult to control the delay and delay jitter arising from latency in a network of Internet scale. Most importantly, the Internet architecture has been originally designed to be a quality of service and security agnostic one as ensuring availability was of paramount importance those days [22]. Still major portion of the original design principles of the Internet stays intact and not going to change in the future unless there is a wholesale change with a design and implementation of a new Internet. Hence it is safe to say that latency and the issues caused by latency are not going to be resolved in the near future.

The other major issue confronted with cloud computing is security and privacy [5]. Since the cloud systems have been located with the Internet, user requests, data transmission and system responses need to traverse a large number of intermediate networks depending on the distance between the users and systems. When customer data is out there in a public cloud, there is a risk of them being compromised of their integrity and confidentiality [23]. Deeper the data inside the Internet, higher the risk as the data has to travel a long distance to and from the user's computer to the cloud system, even if the data is encrypted. Similarly the availability of the cloud systems can also be attacked using various methods [23]. Thus it can be seen that cloud systems at present face various security threats due to very nature of their implementation within the Internet coupled with location independence.

3 FOG COMPUTING

The term "Fog Computing" was introduced by the Cisco Systems as new model to ease wireless data transfer to distributed devices in the Internet of Things (IoT) network paradigm [7],[24]. According to Bonomi et al., [7] the rationale for coining this term to identify this model is that fog is nothing but cloud that is closer to the ground. Hence cloud computing carried out closer to the end users' networks is thus identified as *fog computing*. Fog computing is a virtualized platform that is typically located between end user devices and the cloud data centres hosted within the Internet [7]. Thus fog computing can provide better quality of service in terms of delay, power consumption, reduced data traffic over the Internet etc [25]. The main feature of fog computing is its ability to support applications that require low latency, location awareness and mobility. This ability is made possible by the fact that the fog computing systems are deployed very close to the end users in a widely distributed manner. Fog computing nodes thus hosted must possess sufficient computing power and storage capacity to handle the resource intensive user requests. Other similar concepts where the computing resources have been proposed to be located closer to the users to overcome the limitations of cloud computing include cloudlets and edge computing [20],[26].

Cloudlets are resource rich computer or a cluster of them with the virtualization capability and located closer to mobile users, so that they can respond to their requests fast while maintaining a strong Internet connectivity [20]. Cloudlets have been specifically designed to provide services to mobile users with limited local resources that can act as a thin client and access the cloudlet resources that are just one hop away through a wireless network. Satyanarayanan et al., [20] claim that mobile devices will always remain resource poor compared to stationary devices such as desktop, laptop and servers due to their other requirements such as smaller size, less weight and longer battery life. But on the contrary, the emerging applications and paradigms such as interactive media, augmented reality, natural language processing, speech recognition etc., requires large amount of resources for processing with minimum latencies. Hence locating the high capacity resources as closest as possible to the end user is the only solution. Thus cloudlet is positioned at the right place to handle this kind of requirements and demands.

Edge computing envisages to push the applications, data and computing power away from centralized computing nodes to the edge of the network [27]. Edge Computing, in addition to providing sub-second response to end users, it also provides high levels of scalability, reliability and fault tolerance [26]. The advantages of edge computing include a significant reduction in data movement across the network resulting in reduced congestion, cost and latency, elimination of bottlenecks resulting from centralized computing systems, improved security of encrypted data as it stays closer to the end user reducing exposure to hostile elements and improved scalability arising from virtualized systems. The challenges that must be managed when deploying applications and systems at the edge of the networks include sandboxing for security, distributed load balancing and resource management, accounting and billing, testing, debugging and monitoring [26].

Though different researchers have coined different terms to describe their implementations of highly virtualized, powerful computing models located closer to the end users in order to meet the requirements of emerging computing and networking scenarios, they all in fact have the same or very similar characteristics. Hence they all can be considered to be one and the same. So, in this paper, we stick to the term for computing to represent any of the computing model described above.

4 CLOUD VS FOG

From the above discussion, it can be seen that technically fog computing is very much similar to cloud computing in the sense that both are made up of virtual systems providing the flexibility of on demand provisioning of compute, storage and network resources. But compared to cloud, fog is implemented very close to the end users In this section, we take an in depth look at the similarities and dissimilarities of these two technologies with respect to the demands of the emerging trends in networking. Table 1 summarizes the results of the comparison. From Table 1, it can be seen that Cloud Computing characteristics have very severe limitations with respect to quality of service demanded by real time applications requiring almost immediate action by the server. The detailed explanation on the comparison with respect to their impacts is given below.

Table 1 Cloud Computing vs Fog Computing

Requirement	Cloud Computing	Fog Computing
Latency	High	Low
Delay Jitter	High	Very low
Location of server nodes	Within the Internet	At the edge of the local network
Distance between the client and server	Multiple hops	One hop
Security	Undefined	Can be defined
Attack on data enroute	High probability	Very low probability
Location awareness	No	Yes
Geographical distribution	Centralized	Distributed
No. of server nodes	Few	Very large
Support for Mobility	Limited	Supported
Real time interactions	Supported	Supported
Type of last mile connectivity	Leased line	Wireless

IoT envisages to put an IP address on very device connecting them to the Internet. Vehicular networks, a specialized kind of emerging mobile networks is an important part of the envisaged IoT [7]. In vehicular networks, communication would be taking place between vehicles, vehicles to access points and between access points for the purpose of transferring important information between these objects. Some of the important features of this communication include location awareness, wireless connectivity, faster processing requirements, low latency, real time interactions and mobility. In order to meet these requirements there must be a large number of processing nodes located very close to the clients (vehicles) that can communicate with the clients over low delay wireless links. Thus fog computing is the most suitable communication model compared to cloud computing where information processing takes place deep within the Internet.

The emerging high quality multimedia applications including distributed interactive games, video on demand and streaming demand large data transfer rates with low delay, delay jitter and packet loss [28]. In order to achieve this, it is necessary to process these applications closer to the end users. Since cloud data centres are generally located within the Internet, it is difficult to manage these factors. Hence fog computing is the practical solution for this kind of performance sensitive applications.

Wireless sensor networks have been widely deployed in many environment related applications [7]. These networks are generally characterized by low power, low bandwidth and limited processing capability nodes distributed across wide geographical areas. These networks must be supported by low latency, location aware and widely distributed systems for processing and distributing the data. These are typical characteristics of fog computing rather than cloud computing.

Data security and integrity are two most important characteristics demanded by many Internet applications [24]. Longer the data stays en-route, more vulnerable it is for attacks even when encrypted. Hence it is always desirable to have few hops between clients and servers. Fog computing provides the shortest possible distance while providing all the other advantages of cloud computing. Hence fog computing is preferred over traditional cloud computing in such situations.

Even the availability of cloud systems located within the Internet can be attacked by miscreants using various Denial of Service

(DoS) attack methods [23]. The DoS attacks need not be carried out directly on the end systems themselves, even attacks targeted towards the intermediary devices such as routers can also be equally fatal. Hence there are many opportunities for hackers to target cloud computing systems. On the other hand fog computing nodes are highly distributed near the edge of the user networks, in order to attack the availability of these systems, it is necessary to carry out a massive attack on all the systems that are nearby a client. This needs massive resources from the attackers side too. Also there are not many intermediate devices that can be targeted by the attacker as for computing nodes are located very close to the end users. Hence it can be safely state that fog computing system is less prone to DoS attacks than cloud computing systems.

From the above discussion, it can be seen that fog computing is more responsive to user needs and emerging new computing and networking paradigms than traditional cloud computing systems. Also fog computing is more resilient, rugged and secure than cloud computing in the face of changing needs and emerging trends.

On the other hand, it must also be noted that cloud computing is not without its advantages. Since fog computing requires massive geographically distributed implementation, the single nodes cannot have large amounts of resources due to financial reasons. However high end business computing such as batch processing jobs, would require large amounts of resources while not being very delay sensitive. These kinds of jobs can be handled using traditional cloud computing systems successfully more than fog nodes. Hence fog computing will never be able to replace cloud computing and become the sole cloud computing model of the future. Thus, it is safe to state that cloud and fog would exist side by side serving two different communities and complement each other where necessary.

5 CONCLUSIONS

In this paper, we took an in depth look at the newly proposed cloud computing model known as fog computing. From the above analysis, it can be seen that fog computing performs better than cloud computing in meeting the demands of the emerging paradigms. But fog computing cannot totally replace cloud computing as it will still be preferred for high end batch processing jobs that are very common in the business world. Hence it can be concluded that fog computing and cloud computing will complement each other while having their own advantages and disadvantages. While fog computing will grow in helping the emerging network paradigms that require faster processing with less delay and delay jitter, cloud computing would serve the business community meeting their high end computing demands lowering the cost based on a utility pricing model.

7 REFERENCES

[1] A.Shawish and M.Salama, *Inter-cooperative Collective Intelligence: Techniques and Applications*. Berlin Heidelberg: Spinger, 2014, ch. Cloud Computing: Paradigms and Technologies, pp. 39--67.

[2] R.Buyya, R.N. Calheiros, and X.Li, "Autonomic cloud computing: Open challenges and architectural elements,"' in *3rd International Conference on Emerging Applications of Information Technology*, Kolkata, India, 2012, pp. 3--10.

[3] R.Buyya, C.S. Yeo, and S.Venugopal, "Market-oriented cloud computing: Vision, hype, and reality for delivering IT services as computing utilities," in *10th IEEE International Conference on High Performance Computing and Communications*, Dalian, China, 2008, pp. 5--13.

[4] J.Gubbi, R.Buyya, S.Marusic, and M.Palaniswami, "Internet of Things (IoT): A vision, architectural elements and future directions",*Future Generation Computer Systems*, vol.29, no.7, pp. 1645--1660, 2013.

[5] K.Hashizume, D.G. Rosado, E.Fernández-Medina, and E.B. Fernandez, "An analysis of security issues for cloud computing," *Journal of Internet Services and Applications*, vol.4, no.5, pp. 1--13, 2013.

[6] M.M. Islam, S.Morshed, and P.Goswami, "Cloud computing: A survey on its limitations and potential solutions," *International Journal of Computer Science Issues*, vol.10, no.4, pp. 159--163, 2013.

[7] F.Bonomi, R.Milito, J.Zhu, and S.Addepalli, "Fog computing and its role in the Internet of Things," in *ACM SIGCOMM Workshop on Mobile cloud Computing*, Helsinki, Finland, 2012, pp. 13--16.

[8] M.Georgescu and M.Matei, "The value of cloud computing in the business environment," *The USV Annals of Economics and Public Administration*, vol.13, no.1, pp. 222--228, 2013.

[9] S.Malkowski, Y.Kanemasa, H.Chen, M.Yamamoto, Q.Wang, D.Jayasinghe, C.Pu, and M.Kawaba, "Challenges and opportunities in consolidation at high resource utilization: Non-monotonic response time variations in n-tier applications," in *Fifth IEEE International Conference on Cloud Computing*, Honolulu, HI, USA, 2012, pp. 162--169.

[10] P.Pazowski and Z.Pastuszak, "Cloud computing – a case study for the new ideal of the IS/IT implementation," in *International Conference on Management, Knowledge and Learning*, Zadar, Croatia, 2013, pp. 855--862.

[11] B.H. Bhavani and H.S. Guruprasad, "Resource provisioning techniques in cloud computing environment: A survey," *International Journal of Research in Computer and Communication Technology*, vol.3, no.3, pp. 395--401, 2014.

[12] T.Lorido-Botran, J.Miguel-Alonso, and J.A. Lozano, "Auto-scaling techniques for elastic applications in cloud environments," University of the Basque Country, Spain, Technical Report EHU-KAT-IK-09-12, 2012.

[13] C.Kerr, P.T. Vu, and S.Davari, "Cloud computing: Business trends and the challenges," *International Journal on Recent Trends in Engineering and Technology*, vol.8, no.1, pp. 76--83, 2013.

[14] K.Lee, S.Lee, and H.D. Yang, "Towards on cloud computing standardization," *International Journal of Multimedia and Ubiquitous Engineering*, vol.9, no.2, pp. 169--176, 2014.

[15] E.O. Yeboah-Boateng and K.A. Essandoh, "Factors influencing the adoption of cloud computing by small and medium enterprises in developing economies," *International Journal of Emerging Science and Engineering*, vol.2, no.4, pp. 13--20, 2014.

[16] R.P. Padhy and M.R. Patra, "Managing IT operations in a cloud-driven enterprise: Case studies," *American Journal of Cloud Computing*, vol.1, no.1, pp. 1--18, 2013.

[17] K.Kaur and A.K. Rai, "A comparative analysis: Grid, cluster and cloud computing," *International Journal of Advanced Research in Computer and Communication Engineering*, vol.3, no.3, pp. 5730--5734, 2014.

[18] A.Akella, J.Pang, A.Shaikh, B.Maggs, and S.Seshan, "A comparison of overlay routing and multihoming route control," in *ACM SIGCOMM Conference on Applications, Technologies, Architectures, and Protocols for Computer Communications*, Portland, OR, USA, 2004, pp. 93--106.

[19] A.Csaszar, G.Enyedi, G.Retvari, and M.H.P. Sjodin, "Converging the evolution of router architectures and IP networks," *IEEE Network*, vol.21, no.4, pp. 8--14, 2007.

[20] M.Satyanarayanan, P.Bahl, R.Caceres, and N.Davies, "The case for VM-based cloudlets in mobile computing," *IEEE Pervasive Computing*, vol.8, no.4, pp. 14--23, 2009.

[21] J.L. Adams, L.G. Roberts, and A.I. Jsselmuiden, "Changing the internet to support real-time content supply from a large fraction of broadband residential users," *BT Technology Journal*, vol.23, no.2, pp. 217--231, 2005.

[22] G.Boddapati, J.Day, I.Matta, and L.Chitkushev, "Assessing the security of a clean-slate internet architecture," in *20th IEEE International Conference on Network Protocols*, Austin, TX, USA, 2012, pp. 1--6.

[23] S.Sudha and V.M. Viswanatham, "Addressing security and privacy issues in cloud computing," *Journal of Theoretical and Applied Information Technology*, vol.48, no.2, pp. 708--719, 2013.

[24] J.K. Zao, T.T. Gan, C.K. You, C.E. Chung, Y.T. Wang, S.J.R. Mendez, T.Mullen, C.Yu, C.Kothe, C.T. Hsiao, S.L. Chu, C.K. Shieh, and T.P. Jung, "Pervasive brain monitoring and data sharing based on multi-tier distributed computing and linked data technology," *Frontiers in Human Neuroscience*, vol.8, no. 370, pp. 1--16, 2014.

[25] D.Kovachev, "Mobile multimedia services in the cloud," Ph.D. dissertation, RWTH Aachen University, Aachen, Germany, 2014.

[26] A.Davis, J.Parikh, and W.E. Weihl, "Edgecomputing: Extending enterprise applications to the edge of the Internet," in *13th international World Wide Web conference*, New York, 2004, pp. 180--187.

[27] B.Sosinsky, *Cloud Computing Bible*. Hoboken, NJ: Wiley, 2011.

[28] A.Erbad, "Real-time support for interactive multimedia applications," Ph.D. dissertation, University of British Columbia, Canada, 2012.