ARTICLE IN PRESS

Future Generation Computer Systems ■ (■■■■) ■■■–■■■

Contents lists available at ScienceDirect

Future Generation Computer Systems

journal homepage: www.elsevier.com/locate/fgcs

Next generation cloud computing: New trends and research directions

Blesson Varghese [a,*], Rajkumar Buyya [b]

[a] School of Electronics, Electrical Engineering and Computer Science, Queen's University Belfast, UK
[b] Cloud Computing and Distributed Systems (CLOUDS) Laboratory, School of Computing and Information Systems, The University of Melbourne, Australia

HIGHLIGHTS

- Distributed cloud infrastructure will make use of the network edge in the future.
- Two tier applications will be replaced by new multi-tier cloud architectures.
- Next generation cloud computing impacts both societal and scientific avenues.
- A new marketplace will need to be developed for resources at the network edge.
- Security and sustainability are key to architecting future cloud systems.

ARTICLE INFO

Article history:
Received 14 February 2017
Received in revised form 6 September 2017
Accepted 7 September 2017
Available online xxxx

Keywords:
Cloud computing
Fog computing
Cloudlet
Multi-cloud
Serverless computing
Cloud security

ABSTRACT

The landscape of cloud computing has significantly changed over the last decade. Not only have more providers and service offerings crowded the space, but also cloud infrastructure that was traditionally limited to single provider data centers is now evolving. In this paper, we firstly discuss the changing cloud infrastructure and consider the use of infrastructure from multiple providers and the benefit of decentralising computing away from data centers. These trends have resulted in the need for a variety of new computing architectures that will be offered by future cloud infrastructure. These architectures are anticipated to impact areas, such as connecting people and devices, data-intensive computing, the service space and self-learning systems. Finally, we lay out a roadmap of challenges that will need to be addressed for realising the potential of next generation cloud systems.

1. Introduction

Resources and services offered on the cloud have rapidly changed in the last decade. These changes were underpinned by industry and academia led efforts towards realising computing as a utility [1]. This vision has been achieved, but there are continuing changes in the cloud computing landscape which this paper aims to present.

Applications now aim to leverage cloud infrastructure by making use of heterogeneous resources from multiple providers. This is in contrast to how resources from a single cloud provider or data center were used traditionally. Consequently, new computing architectures are emerging. This change is impacting a number of societal and scientific areas. In this discussion paper, we consider *'what future cloud computing looks like'* by charting out trends and directions for pursuing meaningful research in developing next generation computing systems as shown in Fig. 1.

The remainder of this paper is organised as follows. Section 2 presents a discussion of the evolving infrastructure on the cloud. Section 3 highlights the emerging computing architectures and their advantages. Section 4 considers a number of areas that future clouds will impact. Section 5 sets out a number of challenges that will need to be addressed for developing next generation cloud systems. Section 6 concludes this paper.

2. Changing infrastructure

The majority of existing infrastructure hosting cloud services comprises dedicated compute and storage resources located in data centers. Hosting cloud applications on data centers of a single provider is easy and provides obvious advantages. However, using a single provider and a data center model poses a number of challenges. A lot of energy is consumed by a large data center to

* Corresponding author.
E-mail addresses: varghese@qub.ac.uk (B. Varghese), rbuyya@unimelb.edu.au (R. Buyya).
URLs: http://www.blessonv.com (B. Varghese), http://www.buyya.com (R. Buyya).

http://dx.doi.org/10.1016/j.future.2017.09.020
0167-739X/

Please cite this article in press as: B. Varghese, R. Buyya, Next generation cloud computing: New trends and research directions, Future Generation Computer Systems (2017), http://dx.doi.org/10.1016/j.future.2017.09.020.

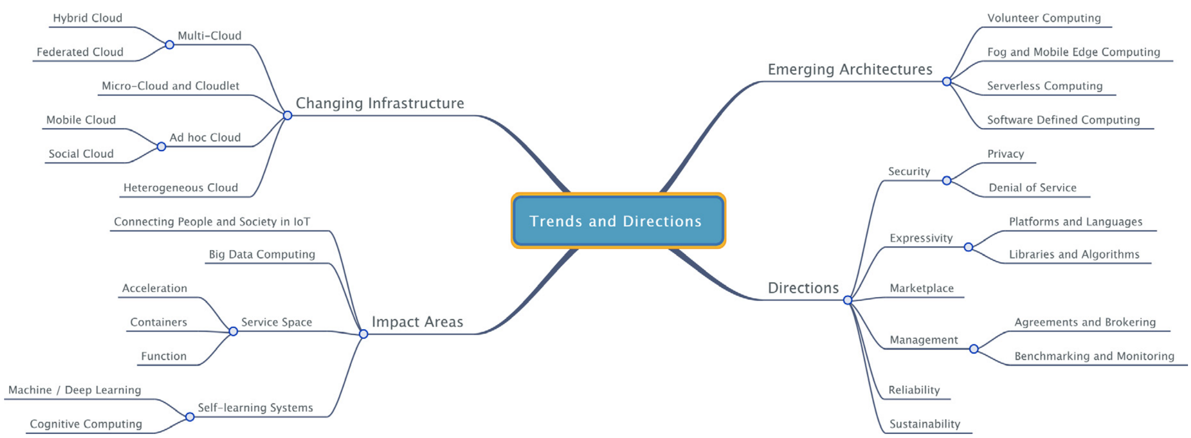

Fig. 1. A snapshot of trends and directions in next generation cloud computing.

keep it operational. Moreover, centralised cloud data centers like any other centralised computing model is susceptible to single point failures. Additionally, data centers may be geographically distant from its users, thereby requiring data to be transferred from its source to resources that can process it in the data center. This would mean that applications using or generating sensitive or personal data may have to be stored in a different country than where it originated.

Strategies implemented to mitigate failures on the cloud include using redundant compute systems in a data center, multiple zones and back up data centers in individual zones. However, alternate models of using cloud infrastructure instead of using data centers from a single provider have been proposed in recent years [2]. In this paper, we consider the multi-cloud, microcloud and cloudlet, ad hoc cloud and heterogeneous cloud to demonstrate the trends in changing infrastructure of the cloud. The feasibility of these have been reported in literature and will find real deployment of workloads in next generation cloud computing. Fig. 2 shows the different layers of the cloud stack where changes need to be accommodated due to the evolving infrastructure. We consider nine layers of abstraction that contribute to the cloud stack, namely network (bottom of the stack), storage, servers, virtualisation, operating system, middleware, runtime, data and application (top of the stack). For facilitating multi-cloud environments and ad hoc clouds, changes will be required from the middleware layer and upwards in the stack. Heterogeneous clouds can be achieved with changes two further layers down the stack from the virtualisation layer. Microclouds and cloudlet infrastructure may require redesign of the servers that are employed and therefore changes are anticipated from the server layer.

We note that there has also been significant changes in the area of data storage on the cloud over the last decade. There are at least three levels of abstraction provided with respect to data storage [3]. At the block level, direct attached storage, such as Amazon EC2, App Engine and Azure VM, and block storage, such as EBS, Blob Storage and Azure Drive, are available for providing quickest access to data for VMs. At the file level, object storage, such as Amazon S3, Google Storage, and Azure Blob, and online drive storage, such as Google and Sky drives, are commonly available services. At the database level, relational data-stores, such as Google Cloud SQL and SQL Azure Blob, and semi-structured data storage, such as Simple DB, Big Table and Azure Table, are available. However, in this paper the focus is on how computing on the cloud has changed over the last decade.

2.1. Multi-cloud

The traditional notion of multi-cloud was leveraging resources from multiple data centers of a provider. Then applications were hosted to utilise resources from multiple providers [4,5]. Rightscale estimates that current businesses use an average of six separate clouds.[1]

The use of multi-clouds are increasing, but there are hurdles that will need to be overcome. For example, common APIs to facilitate multi-cloud need to account for different types of resources offered by multiple providers. This is not an easy given that more resources are rapidly added to the cloud marketplace and there are no unified catalogues that report a complete set of resources available on the cloud. Further, the abstractions, including network and storage architectures differ across providers, which makes the adoption of multi-cloud bespoke to each application rather than using a generic platform or service. Along with the different resources, hypervisors, and software suites employed, the pricing and billing models are significantly different across providers, all of which results in significant programming effort required for developing a multi-cloud application. All management tasks, such as fault tolerance, load balancing, resource management and accounting need to be programmed manually since there are no unifying environments that make these possible. Examples of APIs that alleviate some of these challenges include Libcloud[2] and jClouds.[3] However, further research is required for enabling adoption of clouds across multiple providers.

Hybrid Cloud: A multi-cloud can take the form of a hybrid cloud - a combination of public and private clouds or a combination of public and private IT infrastructure [6,7]. These clouds cater for bursty demands or resource demands known beforehand. The benefit of using hybrid clouds for handling sensitive data is known [8]. It is estimated that 63% of organisations using the cloud have adopted a hybrid cloud approach[4] with use-cases reported in healthcare[5] and energy sectors.[6] The key challenge in setting up a hybrid cloud is network related. Bandwidth, latency and network topologies will need to be considered for accessing a public cloud from a private cloud [9]. Network limitations can result in an ineffective hybrid cloud. Dedicated networking between clouds may enable more effective infrastructure, but requires additional

1 http://www.forbes.com/sites/joemckendrick/2016/02/09/typical-enterprise-uses-six-cloud-computing-services-survey-shows/#e2207a47be31.

2 https://libcloud.apache.org/.

3 https://jclouds.apache.org/.

4 http://www.cloudpro.co.uk/cloud-essentials/hybrid-cloud/6445/63-of-organisations-embracing-hybrid-environments.

5 https://usa.healthcare.siemens.com/medical-imaging-it/image-sharing/image-sharing-archiving-isa.

6 http://w3.siemens.com/smartgrid/global/en/products-systems-solutions/smart-metering/emeter/pages/emeter-cloud.aspx.

Please cite this article in press as: B. Varghese, R. Buyya, Next generation cloud computing: New trends and research directions, Future Generation Computer Systems (2017), http://dx.doi.org/10.1016/j.future.2017.09.020.