Systems, Services, Solutions of the Public Cloud

Eduardo Correia
Ara Institute of Canterbury, New Zealand

INTRODUCTION

Cloud computing impacts the way people interact with technology, at least for consumers. Users no longer store personal data on a single device, but instead access it from a range of devices. When people purchase a new phone, they need only to link it to the appropriate provider to access their contacts, calendar and other data, whether it be documents, photographs or financial statements, and they can exchange one device for another seamlessly and still be able to do what they wish to do. If an individual device fails, it may be inconvenient, but ultimately it does not involve loss of data or prevent access to the date from another device. As providers hold an up to date copy of the data, a user may for instance begin working on a presentation on a PC, continue working on it on a tablet, and decide to do some light editing on a smart phone.

What is not quite as clear is how these underlying technologies both function and coalesce, and moreover how organizations are to leverage the cloud to bring value to a range of stakeholders and interested parties. This article begins by outlining the fundamentals of cloud computing, concepts like service models and how cloud computing may or may not interact with traditional on-premise infrastructure. Then it considers the at times vexed, contentious questions of risk, security, governance and compliance. Once these questions are addressed and the suitable technologies properly implemented, the cloud is capable of delivering to organizations accelerated innovation and quality. This potential is reflected in the growth of the public cloud. European, Middle Eastern and Asian (EMEA) organizations sourced 35% of their total IT infrastructure from the cloud in 2017 and by 2022 this is forecast to grow to 63% (Scott, 2018).

BACKGROUND

A common view of the cloud is that it is nothing new, that people have been using centralized computing in the form of mainframes for instance. While this sheds light on some of the similarities between mainframes and the cloud, it does not reflect the differences between them, differences that are significant. One useful analogy is to compare the cloud and earlier technologies, with the smart phone and the traditional telephone. While the telephone as such is not new and the smart phone can do the same kinds of things as the traditional phone, it is much more flexible and powerful, and this is down to, in large part, how mobile devices take advantage of the cloud, making them more functional and helping extend battery life (Liu, Chen, Ma, & Xie, 2016). As a term “cloud computing” is problematic because it does not mean one thing, but is “an umbrella term” (Missbach, Stelzel, Gardiner, Anderson & Tempes 2013) that refers not just to a set of technologies but a new model of delivering a wide range of computing resources, services and solutions.

DOI: 10.4018/978-1-7998-3473-1.ch047

Fortunately, the National Institute of Standards and Technology (NIST) has made a deliberate effort to define cloud computing and done this well, so much so that its three-page description of this term is now considered the standard for the industry. NIST defines cloud computing as "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" (Mell & Grance, 2011). This definition is comparable to the four key characteristics of cloud computing as "pay-per-use", "elastic capacity and the illusion of infinite resources", a "self-service interface" and "resources that are abstracted or virtualised" (Voorsluys, Broberg, & Buyya, 2011). The cloud is "ubiquitous", "convenient", and "on-demand" in the sense that users can access applications and data from anywhere and at any time, and because organizations (or "customers" as the literature often calls them) can, without having to talk to anyone, provision computing resources and services from a "shared pool of configurable computing resources," a business model that involves renting multi-tenanted infrastructure.

SERVICE MODELS

Cloud providers, through the use of virtualization, essentially abstract the software, whether it be a server, a database or a network from the underlying physical infrastructure and offer customers logical software-defined computing resources and services that they can customize, secure, provision and manage in a way that meets their needs and requirements (Buyya, Vecchiola, & Thamarai Selvi, 2013). The service models in cloud computing all reflect this principle of abstraction and represent different sets of responsibility depending how much or how little customers wish to outsource. Infrastructure as a Service (IaaS), Platform as a Service (PaaS) and Software as a Service (SaaS) therefore represent progressively lower levels of engagement on the part of customers, as they cede more of their responsibilities to the cloud provider.

With IaaS, customers no longer need to purchase physical equipment and associated infrastructure, and manage it, whether it is a server in the case of a small business or a large datacenter in the case of a large business. They can instead utilize the provider's resources presented to them in the form of *virtual* infrastructure such as virtual machines (VMs), networks, storage, firewalls, and load balancers, all on a granular pay-as-you-go basis, so that the customer pays for example for each hour of use of a small VM, which will be cheaper than a more powerful VM. While the benefits outweigh the constraints, customers need to know not just what the provider can but also what it cannot do and work within the limitations of a particular cloud. For instance, in Microsoft Azure once a virtual network has been created it is not possible to add a virtual machine to it. There are workarounds, for instance keeping the disk with the data and connecting them to newly created VMs in the existing virtual network.

In the case of PaaS, customers have no access to or control over not just to the underlying physical but also virtual infrastructure, such as servers, networks, storage and operating systems. PaaS involves the provision of resources and services, such as a development environment, database, and integration and development tools that can be used in a matter of minutes, all provided so as to reduce the effort and cost for SaaS vendors (Brown & Nyarko, 2013). Customers can instead focus on the functionality and logic of applications. It is the responsibility of the provider to ensure that a database for instance can cope with the increasing demands made on it and that the underlying operating system is properly patched and secure. Well-known examples of PaaS are AWS Elastic Beanstalk and Force.com (part of Salesforce.com). In terms of IaaS and PaaS, Amazon Web Services (AWS) is currently the largest public