Cloud computing – what is it and what does it mean for education?

John Powell
Principal Lecturer in Strategy & Management, Leicester Business School, DMU.

Cloud computing is the buzz phrase of the moment, approaching the top of Gartner's hype curve. Even The Economist is rather enthusiastic; for instance, in their special report on Corporate IT:
"The rise of the cloud is more than just another platform shift that gets geeks excited. It will undoubtedly transform the information technology industry, but it will also profoundly change the way people work and companies operate" (Economist, 2008).

Cloud computing is a term you can see being used a lot, but there is a lack of clarity about precisely what cloud computing is (one article cited by a recent McKinsey presentation – Vaquero et al (2009) – suggested that 22 definitions currently exist in the literature.

However, most of us are probably making use of the cloud without realising that this is the case; whenever we access our Gmail or Hotmail accounts, or upload a photo to Facebook, we are using the cloud. The potential benefits and risks, however, are more apparent. I will try and shed some light on defining cloud computing and then explore the opportunities and risks that adoption poses, with particular focus on (higher) education institutions.

Defining the cloud

Broadly, the cloud can be described as on-demand computing, for anyone with a network connection. Access to applications and data anywhere, any time, from any device is the potential outcome. The consumer-level cloud is a good starting point for this – sites like Flickr and Facebook act as digital repositories for data and we can access this data from any internet-enabled device, from our iPhones to our desktop computers. In the case of Flickr and the like, storage of our digital images is, from the consumer point of view, somewhere in the cloud. We don't need to know where specifically, we just need our Flickr login credentials and a web connection. We can see this model as evident in web-based email too.

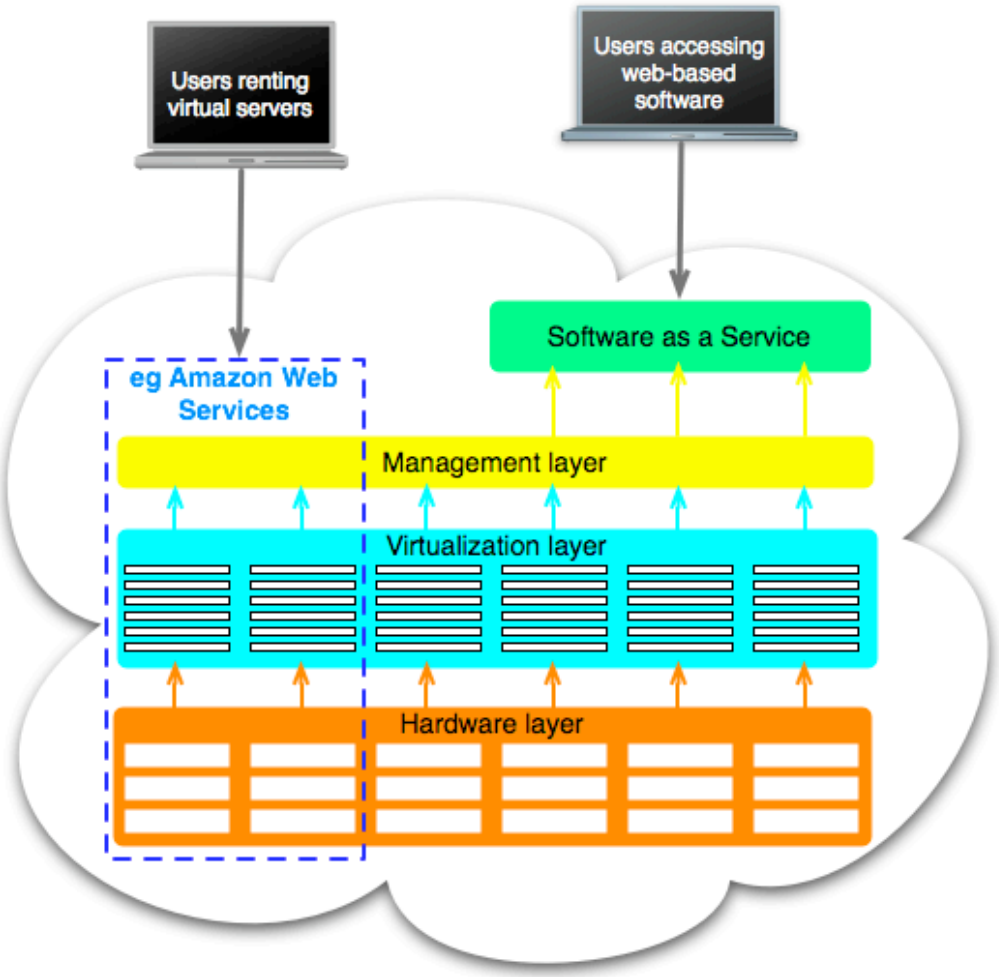
In practice, cloud computing as implemented is substantially more complex than the user perspective of it suggests, and many of the potential benefits of the cloud actually stem from this. Many of the perspectives on the cloud adopt a 'layers' view to describe it (for instance, in the recent MIT Technology Review briefing – such as Naone, 2009a). We also need to note a distinction between 'private clouds' (which exist *within* an organisation) and 'public clouds' which are used to provide services to users outside an organisation (see Armbrust et al, 2009).

The starting point (the bottom layer) is the physical hardware – the servers in datacentres, whether owned by a firm or university for internal use or by the likes of Amazon or Google for public access. The next layer is virtualization. Virtualization allows a single physical server to run many independent virtual servers and is a necessary part of gaining the efficiencies (built on economies of scale) that cloud computing can offer those running the datacentres.. Automating the allocation of computing resources amongst the virtual servers and tracking user (customer) resource use requires a management layer. As Naone (2009a) highlights, this allows "true pay-as-you-go" billing which is part of the appeal of the cloud for users.

The management layer enables a set of services that allow users to tap the processing and storage capability of the cloud. For some users, this is where the cloud layers stop piling up. Amazon, for instance, allows users to rent virtual servers by the hour, to do with what they will – run as web servers, process data, whatever the user wants. Some people, such as Nick Carr, have described this kind of activity as utility computing. Carr has, somewhat controversially, pursued the idea of the commoditisation of IT with some vigour over the last six years or so. His most recent book, The Big Switch, draws explicit parallels between how the provision of electricity evolved and how IT is evolving.

The cloud, however, doesn't stop here. The provision of software (services) to end-users mediated by the web is what has come to be known as Software as a Service (resulting in the rather clumsy acronym SaaS). This builds a further layer – the software running in the cloud and being accessed by users through their browser or other web-compliant tool. Firms like Salesforce.com have successfully built their business model entirely around providing web-

based software, with the users' data storage and the processing of that data taking place in the cloud, rather than on the clients' own servers or their local machines.



A simple layers model of the cloud (based on Naone, 2009a)

The benefits of the cloud

In many senses the primary advantages the cloud brings are to do with cost and efficiency, which are closely intertwined. Essentially the capital costs of computing can de done away with if an organisation relies on the public cloud, buying virtual server time and storage space on demand. Expenditure on IT becomes operational, rather than capital. Moreover, the physical space required for racks of servers is no longer necessary and the organisation no longer incurs energy costs for running and cooling its servers.

For many start-up businesses, cloud computing offers access to computing power that would otherwise be beyond their reach. The entry barrier for large-scale computing task is effectively removed by the cloud. As costs are incurred on a per use basis, the risks of committing to large capital purchases are removed. Scalability allows the organisation to add capacity as and when it's needed and to scale down as well as up, driven by demand.

As an example, Amazon's Elastic Cloud Compute (EC2) service allows the rental of virtual servers by the hour, with a variety of processing and OS options. Amazon term a virtual server an 'instance'. So, a small instance comprises:

> 1.7 GB of memory, 1 EC2 Compute Unit (1 virtual core with 1 EC2 Compute Unit), 160 GB of instance storage, 32-bit platform (Amazon, 2009)

and would cost $0.11 per hour of use for a Linux server based in Europe. A further $0.10 is charged per gigabyte of data moved in or out of EC2. Running Windows raises the costs somewhat, to $0.135 per hour for the small instance but the differential between Linux and Windows increases with larger instances.

There is no long term commitment and an organisation is only charged for what it actually uses. If more instances need to be added, it is literally a matter of minutes to configure new server instances.

Google's AppEngine takes a slightly different approach but with some benefits. With Google you are not renting a virtual server as such, but an amount of processing power. If your processing needs increase then Google's AppEngine will scale dynamically with load, rather than requiring manual set-up of new servers.

Even if ditching all physical servers is seen as a step too far (and many University IT managers may well consider this to be the case), building a private cloud with virtualised servers, even if the organisation owns and maintains the physical infrastructure, can deliver large efficiency gains. A McKinsey survey cited by The Economist (2008) suggests that, without virtualization, on average only 6% of server capacity is used. However, the kinds of economies of scale that large cloud providers can take advantage of will typically be absent.

Nonetheless, in this private cloud approach an organisation can still take advantage of the on-tap computing power in the public cloud. 'Cloud bursting' is a service that provides 'overflow computing' for dealing with spikes in web traffic or processing load (Naone, 2009a).

Flexibility, as well as cost, is thus another compelling advantage of the cloud. As Erik Brynojolfsson of MIT states, *"The ability to be agile in your infrastructure is what separates the winners from the losers… cloud computing is one of the most important technologies that affect the ability to maintain that level of flexibility"* (quoted in Cass, 2009).

Risks and barriers

Surely getting apparently infinite on-demand computing resources at low cost can't be all good news?

There are downsides to the cloud, and they are in part to do with the nature of the cloud market and its development. Cloud computing is at an early stage in its development, and one of the consequences of this is a lack of definitive market standards. It also means there is a stream of new entrants into the industry, each trying to gain some market power. Whilst the market is in flux like this, any decisions made now about committing to a particular cloud provider may have unfortunate consequences as the market matures (the risk of backing the wrong standard, and consequently being left behind, or the wrong provider that disappears, taking your data with them).

The lack of market standards leads to issues to do with lock-in (and lack of transferability within the cloud). As Naone (2009b) notes, once you've committed to a particular cloud provider, an organisation is locked in to that provider. This is not a contractual lock-in but a logistical one. Getting data out and moved to a different cloud provider is difficult (but not impossible, and third party firms have entered the market to solve this problem). Thus, there are switching costs if you change cloud provider.

The issue of lock-in also reflects concerns about reliability. There have been several high profile failures of cloud access, though usually temporary. Both Amazon's and Google's cloud services have been offline a few times, for instance. If you can't move your data or applications to an alternative provider, then your systems are down for the duration. Armbrust et al (2009) identify four service failures between Amazon and Google in 2008, ranging from 1.5 to 8 hours. So, typically this may be hours rather than days or weeks, but it may still come at just the wrong point. Of course, running your own servers is no guarantee of uptime at crucial periods – having your admissions database server go down due to cooling problems on A-level results day, for instance. Nonetheless, the advantage of the cloud should be that it is worry-free, and it is not quite there yet.

Open source projects are likely to be the solution to the standards issue, but the dominant firm to date – Amazon – has yet to embrace this route as it is not really in their interest to allow users to switch easily. It may yet come as some of the other key players in the market – namely Google and IBM – have embraced open source elsewhere.

Concerns about security and privacy are frequently mentioned as issues, though Armbrust el al (2009:15) suggest this may be more an issue of perception rather than reality. They argue that

> *"… there are no fundamental obstacles to making a cloud-computing environment as secure as the vast majority of in-house IT environments, and that many of the obstacles can be overcome immediately with well-understood technologies…"*

Confidentiality of data is a potential issue, depending on server location. Servers within the US are within the bounds of mores stringent scrutiny legislation than that which exists in the UK currently (the Patriot Act, for instance). This is one of the reasons Amazon offers a choice of US or European servers.

One of the most surprising limitations of cloud computing is the data transfer costs. This is raised in the Armbrust et al (2009) article. Essentially, the bandwidth required to move large amounts of data in and out of the cloud is just not there. The conclusion Armbrust et al come to is that it is quicker and cheaper to courier external hard drives than to move large volumes of data (ie in multiples of terabytes) over the internet. Amazon itself has also come to this conclusion as they now offer an AWS Import/Export service which is exactly that, using the US Postal Office. Amazon's estimates of data transfer speeds illustrate the issue nicely. Even on a T3 connection (ie 45Mbps) it would take 1-2 days minimum to transfer 1 terabyte of data. Until network infrastructure improves by several orders of magnitude this will continue to be a barrier to cloud adoption for organisations with large volumes of data (video or image processing, for instance).

Implications for education

The timely publication of Educause's collection of essays on computing and education, The Tower and the Cloud (ed. Richard N. Katz, 2009 – download pdf), brings attention to many areas where the cloud may impinge on education. Given the scale of that book and the scope of this article, it is impractical to rehearse all the concerns here. I will direct attention to a few of the essays though. Katz's opening essay, 'The Gathering Cloud: Is this the end of the middle?', is an excellent general starting point. Yanosky's 'From Users to choosers: The cloud and the changing shape of enterprise authority' and Goldstein's 'The Tower, the Cloud, and the IT leader and workforce' both address the impact cloud computing will likely have on IT management and process with HE institutions.

Broadly we may distinguish between what the cloud means for IT management and how institutions organise and buy IT capability, on the one hand, and on the other the changes the cloud is already bringing to how students and academics (and administrators) actually work. Many of the benefits and risks noted above apply to education institutions as much as business. As Katz (2009:12) notes, *"We are in a time of emergence when the best advice is to observe and to be sensitive to areas from which change is emerging."* Nonetheless, he also concludes that *"t]he emerging networked information economy creates unprecedented opportunities for colleges and universities to rationalize their highly distributed IT resources and to extend their institutional footprint"* (Katz, 2009:32). Similarly, Goldstein (in Katz, 2009: 242) notes that *"[f]inancial pressures alone are likely to lead many IT organizations to turn services over to the cloud out of necessity."*

Some of the change Katz highlights above does come from within, and is being driven by the very ease of accessing and using cloud-based services: *"… students will arrive on campus with their own IT architectures and service arrangements. These students… will have little use for or patience with college or university offerings that underperform or force them to lose precious connections to people and processes that they have accumulated since childhood"* (Katz, 2009:18). The same will become true of staff too.

Katz's essay raises a number of over-arching concerns reflecting how institutions respond to the changes the cloud will engender in the IT environment, the education market and in student and staff behaviour. The significance of IT governance in the institution is a theme that is later elaborated in the book by Yanosky and by Goldstein. There are practical considerations to address – for instance, how to ensure necessary institutional information is stored in perpetuity and is auditable in the cloud – and philosophical – such as the nature of scholarship in the digital realm. It is beyond the scope here to resolve these issues but institutions must recognise that the changes in IT provision the cloud implies will inevitably affect all, although the timescale over which this takes place will vary.

At a prosaic level, the focus on partner relationship management (with cloud service providers) will become a substantive part of IT management within institutions. As Goldstein (2009: 243) suggests,
"the cloud may diminish the set of traditional services that an institution's IT organisation must provide on its own. However it also presents that institution with a more complex set of options to weigh."

Whatever view we take of the changes in IT, there is no doubt that the future is cloudy.

References

Armbrust, M et al (2009), Above the clouds: A Berkeley view of Cloud Computing, UC Berkeley EECS, Feb 10th
http://www.eecs.berkeley.edu/Pubs/TechRpts/2009/EECS-2009-28.pdf

Cass, S (2009), Market watch: Virtual computers, real money, *MIT/Technology Review*, July/August
http://www.technologyreview.com/computing/22608/

Economist (2008), Corporate IT Special Report: Let it rise, *The Economist*, Oct 23rd
http://www.economist.com/specialreports/displayStory.cfm?story_id=12411882

Goldstein, P (2009), The Tower, the Cloud, and the IT leader and workforce, in Katz, R (ed) (2009), *The Tower and the Cloud: Higher Education in the Age of Cloud Computing*, Educause
http://www.educause.edu/thetowerandthecloud

Katz, R (ed) (2009), *The Tower and the Cloud: Higher Education in the Age of Cloud Computing*, Educause
http://www.educause.edu/thetowerandthecloud

McKinsey & Co (2009), Clearing the Clouds [discussion document], Uptime Institute, March

http://uptimeinstitute.org/images/stories/McKinsey_Report_Cloud_Computing/mckinsey_clearing_the%20clouds_final_04142009.ppt.pdf

Naone, E (2009a), Technology Overview: Conjuring clouds, *MIT/Technology Review*, July/August
http://www.technologyreview.com/computing/22606/

Naone, E (2009b), Industry challenges: the standards question, *MIT/Technology Review*, July/August
http://www.technologyreview.com/computing/22611/

Vaquero, LM et al (2009), A break in the clouds: towards a cloud definition, *Computer Communication Review,* v39 i1 pp50-55
http://portal.acm.org/citation.cfm?id=1496091.1496100

Yanosky, R (2009), From Users to choosers: The cloud and the changing shape of enterprise authority, in Katz, R (ed) (2009), *The Tower and the Cloud: Higher Education in the Age of Cloud Computing*, Educause
http://www.educause.edu/thetowerandthecloud

See also:

Golden, B (2009), McKinsey Cloud Computing Report Conclusions Don't Add Up, CIO, April 27
http://www.cio.com/article/490770/McKinsey_Cloud_Computing_Report_Conclusions_Don_t_Add_Up

Hinchcliffe, D (2009), Eight ways cloud computing will change business, Enterprise Web 2.0 blog, June 5th
http://blogs.zdnet.com/Hinchcliffe/?p=488

Young, J (2008), 3 ways that web-based computing will change colleges – and challenge them, The Chronicle of Higher Education, v55 i10
http://chronicle.com/free/v55/i10/10a01601.htm