2010

Computing Clouds on the Horizon? Benefits and Risks from the User's Perspective

Roger Clarke
Xamax Consultancy Pty Ltd, Australian National University, University of N.S.W., roger.clarke@xamax.com.au

23rd Bled eConference
eTrust:
Implications for the Individual, Enterprises and Society
June 20 - 23, 2010; Bled, Slovenia

Computing Clouds on the Horizon?
Benefits and Risks from the User's Perspective

Roger Clarke

Xamax Consultancy Pty Ltd, Canberra, Australia
and Visiting Professor at the Australian National University
and the University of N.S.W., Sydney

Roger.Clarke@xamax.com.au

Abstract

The term 'cloud computing' has shot into prominence recently, driven, as most buzz-phrases are, by marketing interests. The term encompasses multiple, pre-existing services, but an analysis of the phenomenon's key features shows that it includes some new developments. Those developments have potential user benefits. They also embody risks. Those risks will be almost entirely borne by users and their customers, and their impacts could be considerable. Where cloud computing is being considered for non-trivial applications, careful risk assessment and risk management are essential, for organisational and individual users alike..

Keywords: virtualisation, *aaS, cloud manager, cloud broker, reliability, integrity, security, privacy, compliance, survival

1 Introduction

The image of a cloud has long been used as a means of depicting the means whereby the Internet somehow provides connectivity between distant devices. The origin of the term 'cloud computing' is commonly ascribed to Google's CEO in a conference presentation in 2006. It was quickly applied to announcements by Amazon and IBM of intentions to release services claimed to be rather different from anything currently on offer. Soon afterwards it was retro-fitted to the SaleForce service (Fowler & Worthen 2009). The first articles using the term that are visible in Google Scholar were published during 2008.

Broadly speaking, cloud computing involves customer data and processing being delegated to a service-provider, which uses hosts whose whereabouts are decided by the service-provider without reference to the customer. The approach is claimed to enable pay-for-usage, rather than the customer having to invest in their own technology, which proponents claim will in turn result in enhanced affordability and convenience.

The term was created by marketers, and is intended to convey a sense of excitement and difference, even revolution. On the other hand, it can be regarded as an evolution from

several previous threads of development, and even as an attempt to recover from previous failed attempts. A balanced interpretation might be that it represents a further step in the commoditisation of IT services.

The services currently on offer exhibit somewhat varying characteristics. Some are applications such as SalesForce, Clarizen (project management), gmail, and Google Apps, most visibly Google Docs for Business and Google Docs for Consumers. Others represent underlying infrastructure or platforms, such as Google Code, Amazon's Elastic Compute Cloud (EC2), Microsoft Azure and Sun Grid.

As is the case with all new, marketing-motivated terms, questions arise as to what the term means, what benefits such services confer, to whom, and under what circumstances; what negative impacts and implications they entail, and for whom; and what risks arise, and who bears them. Analyses have been undertaken from the perspective of service-providers (e.g. Armbrust et al. 2009). The need exists for analyses from the viewpoints of prospective organisational users and individual users.

This paper is exegetic in nature, taking the form of a critical examination of the topic. The approach adopted has been the conduct of a substantial review of academic, commercial and popular literatures, complemented by the application of prior bodies of theory and practice to the subject-matter, and re-visits to the terms of contract and privacy policies of one key service-provider and its performance against previously-published normative templates.

In November 2009, Google Scholar disclosed a few dozen articles that had attracted citations, all of which were evaluated, while the AIS eLibrary disclosed no articles of relevance. During December 2009, however, four further conference papers were posted, three concerned with the technology, and the other with its economics. To date, very few formal publications have adopted the perspective of users. This research therefore appears to be a pioneering foray in this particular area.

The article commences by considering the meaning of the term, culminating in a working definition, statement of scope, and architectural model. The potential benefits to users are presented, in structured form. Disbenefits and risks are then considered. Although specific elements have been discussed in various prior publications, few sources were unearthed that attempted a comprehensive review of the issues that arise. It was accordingly necessary for that section to present an original analysis. Implications are drawn for organisational and individual users, and opportunities for researchers are identified.

2 The Nature of Cloud Computing

This section reviews the origins of the term, and key features that represent candidates for a definition, in order to establish a working definition, a statement of scope, and an architectural model.

2.1 Origins

Since the term 'cloud computing' was coined, apparently in 2006, a great many definitions have been offered. The review below shows that some commonalities exist, but that no authoritative definition has yet been enunciated.

Articles published during 2009 generally place cloud computing in the early 'Inflated Expectations' phase of the Gartner 'hype-cycle'. This model was first published in 1995 (Fenn 1995, Linden & Fenn 2003). The hype-cycle notion embodies the presumptions that each new term reflects a meaningful phenomenon, that each new phenomenon is usefully differentiated from predecessors, and that each new phenomenon will survive rather than fail. Allowance also needs to be made for the possibilities of outright failure, and absorption by some subsequent and perhaps better-conceived notion.

One of the difficulties in defining cloud computing is that the descriptions provided encompass a great many pre-existing categories of service, some of which date as far back as the 1960s, while others that are of more recent coinage. Exhibits 1 and 2 identify respectively predecessor terms and related concepts.

Exhibit 1: Predecessor Terms

- Computing as a utility, and 'computer service bureaux', later called 'data centres'. These were established concepts in the 1960s and through the 1970s
- Application Service Providers (ASPs), a 1980s notion
- working from home / tele-work, which has always involved dependence on remote databases, possibly hosted by service-providers
- working on the move, in 'road warrior' mode, which similarly involves dependence, in this case not only on remote service-providers, but also on varying network-access-providers, and possibly also multiple access devices
- docking of portables into corporate networks, and portable-with-desktop replication / synchronisation / mirroring
- Internet Service Providers (ISPs) – as distinct from what are properly referred to as Internet Access Providers (IAPs) – since the late 1980s. An ISP provides remote services for content-owners, such as bulk email-download using the POP protocol, and selective email-download using the IMAP and HTTP protocols, and web-servers. To varying degrees, the copies of emails on the ISP's hosts are the authoritative version, and the copies on the user's device(s) are the secondary versions
- Web Services. This dates from about 2000, and refers to a coherent and cohesive framework for Web-based collaboration among back-end processing and database servers. The intention has been to achieve clearly defined and reliable inter-operability through defined interfacing mechanisms
- Service-Oriented Architecture (SOA). This is a vague notion dating from the early-to-mid-2000s, which appears to be intended to extend the Web Services notion to enable interoperability among heterogeneous systems, through loose coupling

Exhibit 2: Related Concepts

- Software as a Service (SAAS), since the late 1990s. This provides access to business software, hosted remotely, and paid for on a subscription basis, e.g. Salesforce. A range of variants exist (Woloski 2008)
- cluster computing, in which inter-connected stand-alone computers are managed as a single integrated computing resource
- grid computing, in which computational resources are assigned dynamically, depending on such factors as their availability, capability, performance, cost, and quality-of-service requirements
- peer-to-peer (P2P) architectures. These have taken advantage of the considerable capacity of user devices, by exploiting all devices to act as both a client for the device's user and a server for remote users. Cloud computing sustains the longstanding 'client-server' architecture between the user's device and the service-provider's infrastructure, but may adopt P2P architecture within the infrastructure run by the service-provider and its strategic partners
- server-virtualisation. The enormous range of meanings of the term 'virtualisation' is reflected in the size of the Wikipedia disambiguation page. The key requirement in this case is that the user has no need to be aware of which server running on which host is delivering the service, nor where (in network, physical or perhaps even jurisdictional terms) the hosting device is located
- Infrastructure as a Service (IaaS) and Platform as a Service. Both of these appear to date from 2006, and indicate rental of facilities at lower layers than is the case with Software as a Service (SAAS) and the even earlier Application Service Provider (ASP) notion.
- Anything as a Service. As a summary depiction of the various alternative levels at which computing facilities can be bought in, some authors have adopted the rather unattractive acronym *aaS.

2.2 Key Features

A wide range of definitions is available, e.g. CCJ (2009). A commercial provider offers this: "[cloud computing is] the notion of providing easily accessible compute and storage resources on a pay-as-you-go, on-demand basis, from a virtually infinite infrastructure managed by someone else" (Crandell 2008). An attempt at a tighter definition, from the computer science literature, is "a large-scale distributed computing paradigm that is driven by economies of scale, in which a pool of abstracted, virtualized, dynamically-scalable, managed computing power, storage, platforms, and services are delivered on demand to external customers over the Internet" (Foster et al. 2008).

In Armbrust et al. (2009), the three distinguishing features are argued to be "the illusion of infinite computing resources available on demand ..., the elimination of an up-front commitment, ... [and] the ability to pay for use of computing resources on a short-term basis as needed". In Mell & Grant (2009), on the other hand, five 'essential characteristics' were nominated. See also ISACA (2009, p. 6). These were:

- On-demand self-service (i.e. automated response by servers to direct requests by clients)
- Broad network access (i.e. from anywhere, using any device)
- Resource pooling (i.e. the provider allocates resources according to demand, rather than assigning resources to particular clients)
- Rapid elasticity (i.e. resources are scalable according to demand)
- Measured service (i.e. resource usage is metered)

Several authors have attempted to sift the literature for an emergent, authoritative definition. An analysis by Vaquero et al. (2009) found no common features, but the authors suggested that the key aspects were virtualized resources that can be dynamically re-configured to the scale needed by the user at the time.

In Buyya et al. (2009a, 2009b), it is asserted that it is inherent to cloud computing that services are subject to a Service Level Agreement (SLA). Armbrust et al. (2009) take the opposite view, lauding the emergence of "pay-as-you-go computing with no contract" (p. 7). The analysis that follows needs to take into account the extent to which the cloud computing service-provider offers warranties and indemnities, and abides by them, or operates on an 'all-care, no-responsibility' basis.

2.3 A Working Definition and Scope Statement

In order to support a coherent analysis of benefits, downsides and risks, a working definition is essential. In this paper,

Cloud computing refers to a service that satisfies all of the following conditions:

- *the service is delivered over a telecommunications network;*
- *users rely on the service for access to and/or processing of data;*
- *the data is under the legal control of the user;*
- *some of the resources on which the service depends are 'virtualised', by which is meant that the user has no technical need to be aware which server running on which host is delivering the service, nor where the hosting device is located; and*
- *the service is acquired under a relatively flexible contractual arrangement, at least as regards the quantum used.*

Several circumstances are expressly declared to be out-of-scope. The term, as used in this paper, does not relate to access to, or processing of, data that is not under the user's control. Hence a user is not using cloud computing if they are accessing someone else's web-site, or using an outsourced analytical service that processes census data. For clarity, the organisation that outsourced that web-site or those analytical services may be dependent on a cloud computing services provider.

In addition, if the organisation retains control over the primary copy of its data, and has the capability of processing it, then it is not operationally dependent on cloud computing, but is only using it as an adjunct to its operations, e.g. for peak processing power, or backup.

On the other hand, a wide range of circumstances are defined to be within-scope for the purposes of the analysis that follows. Applications may offer computational resources, data storage or communications. The scope extends to emails, the organisation's (or individual's) own web-pages, and data-formats typical of 'business applications', including primarily textual ('word-processed') documents, spreadsheets, images and databases.

Cloud computing applications encompass circumstances in which:

- the sole copy is held 'in the cloud', with any copies on the user's own device(s) being temporary;
- the authoritative copy is held 'in the cloud', with secondary copy/ies on the user's own device(s); and
- a secondary copy exists 'in the cloud', with the primary version remaining on the user's own-device (cf. backup or device-storage synchronisation/mirroring); but with the qualification that cloud computing is in that case not being used operationally, but only for backup, or perhaps peak-demand access or processing.

Asynchronous communications are within-scope, as follows:

- communications stored for the medium or long term (e.g. email and attachments accessed using the IMAP or HTTP protocols – in the latter case, commonly referred to as 'webmail' – or using an MS Exchange server); and
- communications stored remotely on a temporary basis (e.g. email and attachments downloaded using the POP protocol).

Some forms of synchronous communications (e.g. IM, telephone/VoIP, tele- and video-phone and -conferencing) may be dependent on cloud computing from the outset (as, it might be argued, is the case with Skype VoIP), or may come to be so.

Further, a cloud computing service may be openly accessible, (a 'public cloud'), or limited to particular users, as in an Intranet or an Extranet (a 'private cloud'). It may be depended upon by an organisation (in which case the relationship with the service-provider is 'B2B') or by an individual (in a 'B2C' relationship).

2.4 Architecture

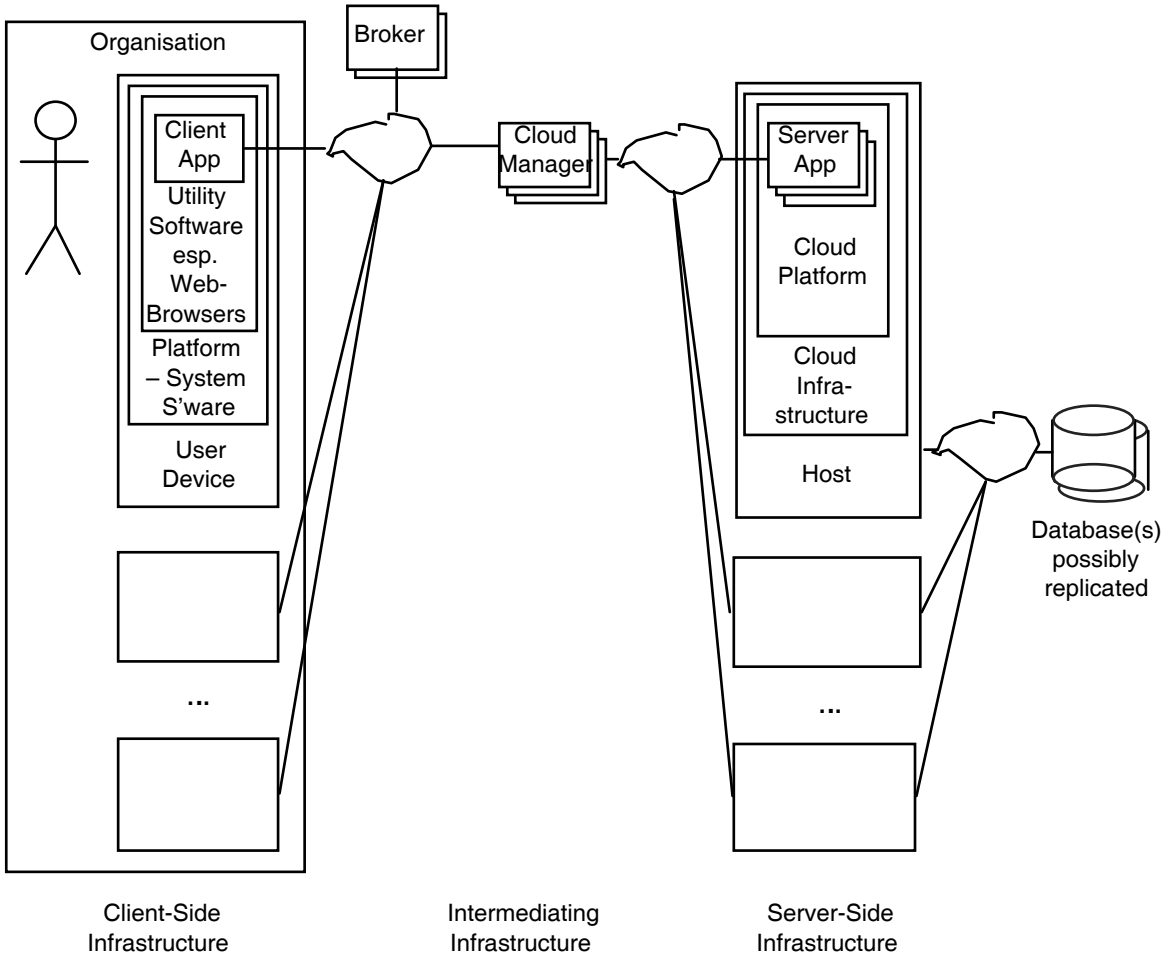
Beyond a mere textual definition and declaration of scope, clarity is needed concerning the elements involved and the inter-relationships among them. Architecture diagrams in the literature to date have focussed almost exclusively on the server-side of the cloud. For example, Youssef et al. (2008) proposed what they referred to as an 'ontology', comprising layers for infrastructure, for software environment or platform, and for applications. Other architectural depictions are in Anderson (2008) and CSA (2009). Maturation is evident, however, in Buyya et al. (2009a) and Buyya et al. (2009b).

To support an analysis of benefits and risks for users, it is essential that the architectural model encompass not only the service-provider side, but also the user side and the intermediating elements. Exhibit 3 reflects all three segments of a comprehensive cloud computing model.

In this model, a user utilises one or more user devices, and may do so within the context of an organisation. The software in each device comprises a client application (which

may be substantial, or merely, say, a few lines of Javascript), which may run within utility software. The presumption is generally made that Web-browsers will be the primary or even sole vehicle, but other possibilities exist. The upper-layer software on the client-side depends, of course, on underlying systems software and hardware.

Exhibit 3: A Sufficiently Rich Architecture for Cloud Computing



On the other side, the cloud application services layer comprises such offerings as all forms of webmail, Google Docs and Microsoft Office Live. Application-servers are dependent on underlying software and hosts. The literature currently distinguishes between 'cloud platforms' (such as Google's App Engine and the SalesForce Apex language), and deeper-nested 'cloud infrastructure' (such as Amazon's Elastic Compute Cloud). A host may run multiple instances of server applications (and conceivably also of platforms or infrastructure). Server applications generally need access to background databases. These may be remote, and hence accessed over the Internet, or replicated locally.

Intermediation is needed between the client and server applications. Communications over the Internet must be facilitated. A function, referred to here as a 'Cloud Manager',

must be available, to at least assign the request to a server. In practice, a range of additional functions must be performed at the gateway to the cloud, including monitoring of the available capacity of the various hosts, load-balancing, and usage accounting. The tariff may be variable, in which case the Cloud Manager needs to administer it, and make it available to clients. If cloud computing matures into a competitive market, Cloud Brokers are likely to emerge, to intercede between vast numbers of clients, on the one hand, and service-providers' Cloud Managers, on the other.

3 Benefits Available from Cloud Computing

Earlier versions of outsourced service provision have already offered a range of potential benefits to both organisational and individual users. The literature generally clusters benefits into business and technical factors (e.g. Loh & Venkatraman 1995, Kremic 2006).

Cloud computing offers upgraded promises, and demands a finer-grained analysis. The catalogue in Exhibit 4 reflects the selling-points attributed to cloud computing in the recent commercial and academic literatures. It is divided into enhanced service accessibility, other technical benefits, and financial benefits.

Exhibit 4: Potential Benefits from Cloud Computing

Enhanced Service Accessibility

Access to Services that are otherwise unavailable. In some circumstances, a service-provider may offer a new or exclusive capability – although this is likely to be the case only during a limited period of time, since most such services are, at least in principle, competable. A more common situation is that some organisational users, and especially many individual users, may be technically or financially incapable of establishing and running a particular service for themselves

Access to Services from multiple desktop devices. Each user, whether within an organisation or acting as an individual, is likely to use multiple desktop devices, in various locations, including at home, at work, at clients' sites, in airport lounges, in Internet cafes, etc. By using authoritative data running on a remote device, the user reduces device-dependence in exchange for increased network-access dependence. In many circumstances, the trade-off may be advantageous

Access to Services from scaled-down devices. A service may perform the vast majority of the function server-side, enabling the client to be run on a device with very limited capacity. This opens up scope for long-promised but little-used 'thin clients', but the primary devices used are more likely to be various forms of handheld computers, and mobile phones. This depends, however, on the service being designed with this aim in mind

Access to Services from multiple device-types. Each user, whether within an organisation or acting as an individual, is likely to use multiple kinds of devices, including desktop PCs, portable PCs, various handhelds, and mobile phones. A suitably-designed service may support convenient access to data and applications on any and each of these device-types, through a variety of user-interfaces

Other Technical Benefits

Professionalised backup and recovery. A service may be designed to provide assured backup of data and software, and assured, simple, efficient recovery. This is because these are core capabilities of a service provider, and that organisation is likely to be more professional, attentive and disciplined than many user organisations and particularly individual users. Backup and recovery services can be provided whether the primary operational service is run on the user's own network, outsourced, or delegated to the cloud

Scalability. Where the transaction and/or data-volumes vary significantly over time, a service may offer assured server-capacity, storage-capacity, and access to the requisite application software. This may apply in a long-term growth curve (or indeed a tailing-off, as occurs with many legacy systems), and in contexts that involve highly-peaked demand, associated with daily, weekly, monthly, annual or even longer cycles, and with events

Collaboration convenience. Collaborative content (including documents and other data which are co-owned and co-maintained) is inherently accessible and amendable by multiple authors. There are advantages in hosting a service such as a Wiki remotely from each of the participants. There may be advantages in the remote host being flexible rather than fixed

Copyright convenience. The service-provider can assume responsibility for all aspects of acquisition, maintenance and licensing of software and of data

Financial Benefits

Lower Investment / up-front cost. The organisation or individual user may be able to avoid investment in hosts, hosting software, applications and people with the requisite expertise to establish the service, because the service-provider takes on that responsibility. The service-provider's costs and profit-margin may be offset through economies of scale and/or scope. On the other hand, effort, time and money still need to be invested by users in determining their requirements, evaluating alternative ways of satisfying them, establishing a strategy and plan, and implementing, monitoring and controlling performance against the plan

Lower Operational Costs. The organisation or individual user may be able to pay-as-they-go, for what they need, rather than paying on an ongoing basis for excess capacity. Savings may arise, provided that the service-provider has significant advantages such as economy of scale, and passes sufficient of those savings on to their customers

Lower IT Staff Costs. Specialist skills are expensive, difficult for business executives to understand, and difficult to manage. Savings may be available by renting such skills from a specialist, rather than sustaining them in-house.

Few of the potential benefits arise solely from the incremental difference between cloud computing and its predecessors, and hence rational users need to consider whether cloud computing or some more conventional form of outsourcing, or indeed insourcing, is appropriate to their needs. Moreover, none of the benefits arise automatically, but rather

are contingent on correspondence between the user's needs, on the one hand, and the service-provider's capabilities, terms of service and pricing, on the other.

Despite the technical benefits, it appears that service-providers perceive the primary driver for adoption as being cost-savings. A secondary driver may be convenience to business divisions arising from the ability to by-pass internal IT departments and contract directly for services. If this transpires to be the case, then the cautious risk assessment conventionally undertaken by IT departments will also be by-passed. It is therefore particularly important for senior executives to appreciate the downsides of cloud computing that are analysed in the following sections. Technical factors are identified first, then business risks.

4 Technical Disbenefits and Risks

In order to provide a sufficiently rich analysis, this section categorises the potential problems arising from cloud computing, firstly into technical and business factors, and then into sub-topics. This section focusses on the technical factors, relating to service operation, contingent risks, and security. Wherever practicable, the analysis is based on well-established business norms and Standards.

4.1 Operational Considerations

This sub-section considers the normal operation of services that are outsourced under cloud computing arrangements. Guidance is provided by relevant standards, in particular AS ISO/IEC 2000-2007 (Information technology - Service management), and by the Information Technology Infrastructure Library (ITIL). But the standards documents are rather diffuse, and ITIL lacks clear statements of the qualities that the processes it describes are intended to assure. Avizienis et al. (2004) summarises the concept of Dependability as being the delivery of trustworthy services free of a number of kinds of defect. The categories adopted below generally reflect the Avizienis definitions, but have been adapted somewhat to better support the analysis of risks from the user's perspective. Once reliance has been placed on cloud computing, risks arise in the following areas.

Fit

There is a fundamental need for a service to fit the needs of the organisation and/or the individual user. In practice, this need is compromised to a greater or lesser degree depending on the cost-savings that can be achieved by accepting a less-good fit. Outsourcing frequently involves a considerable degree of sacrifice of fit. Cloud computing, being a particularly opaque form of outsourcing, may in many cases result in a both a high degree of mis-fit. Worse, there is likely to be uncertainty about what the service's features actually are, and hence a lack of clarity about the details of the mis-fit between the services and the needs they are intended to satisfy.

A further aspect of Fit is the extent to which the software, software versions, protocols and data formats supported by the service-provider are convenient to the organisational or individual user. This can be assessed before a service is adopted; but the change over time in both the user need and the service-provider's offerings may lead to incompatibilities in the future. It is highly disconcerting to an organisation to discover that it is unable to extract a copy of a database on which it depends.

Reliability

The notion of reliability encompasses a cluster of attributes:

- **Availability**. This is commonly used to refer to the readiness and reachability of a service, and hence of the underlying infrastructure of hosts, servers and databases
- **Accessibility**. This refers to the readiness of the networks needed to reach the service, including their speed of access, and the speed of server-response
- **Robustness**. This refers to a low frequency of unavailability, including continuity of service despite the occurrence of a range of threatening events
- **Resilience**. This refers to rapid resumption of services after an outage occurs
- **Recoverability**. A service exhibits Recoverability if the databases that were being operated on, and transactions that were being processed, at the time a service outage occurs, are in a predictable and manageable state following the resumption of services

IT services are commonly subject to planned outages (for such purposes as backup and upgrades), unless they are fully supported by hot-sites and their switchover arrangements are proven and transparent to the user.

All IT services are also subject to unplanned outages. Even at this early stage in the emergence of a cloud computing market, service-providers have suffered unscheduled downtime. For example, Gmail was reported to have had such problems on six occasions in eight months of 2008-09 (Raphael 2009), and Rackspace was reported to have suffered outages in June, July and November 2009 (Brodkin 2009).

Integrity

The term 'integrity' is used here to encompass sustained quality of all relevant service features, including content, content versions, software functionality, software interfaces, and software versions. There are many ways in which data and services can degrade, and loss of integrity is potentially very harmful to the user, and to the user's own customers.

Further, a counterpoint to the transparency of a well-functioning service is opaqueness of problems in a poorly-performing one. The user organisation has no control, and no access, and is entirely dependent on the service-provider for investigation, corrective actions, prioritisation of efforts to achieve corrective action, and information about the cause of the problem.

Maintainability

This term refers to the attribute of a service whereby service deficiencies can be readily discovered, notified, investigated, clarified, and then fixed, and minor modifications can be requested and undertaken, without undue impact on Fit, Reliability and Integrity, and for reasonable Costs.

On the one hand, a specialist outsourced services provider may be expected to have professional capabilities that are superior to those of its customers. On the other hand, the service is opaque, and the service-provider:

- is remote;
- has its own priorities;
- may have many customers using the particular service and whose needs may conflict;
- (particularly in the absence of an SLA) may be under no legal obligation to even respond to notifications of deficiencies and requests for modification, let alone address them; and
- (even with an SLA in place), may not deliver on its undertakings.

Costs

Data processing continues to cost money. Data transmission does as well, particularly for very large volumes of data. Very large volumes may arise because of the size of the data-sets being handled, or because of the need for frequent, medium-sized data transfers as part of replication and synchronisation processes. Associated with this may be delays in data transfer, because large-volume data transfers (whether to enable processing, or to support replication) take time. Despite the vast improvement in delivered bandwidth since the mid-1990s, very large electronic data transfers still take longer than the physical transport of storage volumes.

Even if the service-provider's fees are initially reasonable, they may not stay that way. Outsourcing arrangements commonly result in the user organisation losing corporate knowledge about the application, IT services, and the reasonable costs involved in delivering them. Combined with the inherent lock-in effect achieved by any service-provider, this creates the risk of price-gouging. The outsourcing literature refers to competitive action against the outsourcer by the outsourced service provider as 'opportunism risk' (Loh & Venkatraman 1995, Kremic 2006).

4.2 Contingent Risks

This sub-section considers the contingencies that may arise, and that would have an adverse impact on the organisational or individual user that is depending on a cloud computing service-provider. Additional standards relevant to these risks include ISO/IEC 24762:2008 (Disaster Recovery Services), BS 25999:2006/07 (Business continuity), and BS 25777:2008 (Information and communications technology continuity management – Code of practice). The following major concerns present themselves.

Major Service Interruption

The meaning of 'major' depends on the user's circumstances. That this is not just a theoretical risk is underlined by Metz (2008). Another outage of the same service was reported on 15 July 2009.

Service Survival

A supplier may collapse, or may withdraw from the cloud computing line of business. A number of safeguards need to be provided, including software escrow, data backup, and

assurance that the software and data are not subject to being withheld by a receiver on behalf of creditors.

Data Survival

Even where a supplier continues in business, and even where the service remains operational, the possibility of data loss exists. There have already been several instances reported of outright loss of data by storage providers, e.g. Brodkin (2008b). Microsoft's (aptly-named?) Danger service was reported to have lost data of T-Mobile Sidekick customers in October 2009 (Fried 2009).

Flexibility

A less catastrophic, longer-term risk also requires consideration. It is vital to organisational and individual users that the service-provider provide:

- forward-compatibility of software, protocols and data formats – to enable migration to new levels of service;
- backward compatibility – to protect legacy systems; and
- lateral compatibility – to ensure the freedom to escape to another provider.

4.3 Security Risks

This sub-section considers the security aspects of cloud computing services. Relevant standards include ISO/IEC 27002:2005 – previously 17799:2005 (Information technology - Security techniques - Code of practice for information security management).

An early analysis was presented in Brodkin (2008a). See also Mather et al. (2009). CSA (2009) provides a preliminary analysis by an association of vendors, which was the subject of a scathing attack in Aeon (2009). The risks are largely similar to those to which any in-house operation is subject. However, they are generally more opaque where the service is outsourced, and are likely to be all the more so where cloud computing is adopted. The risks include the following.

Service Security

This refers to resistance against environmental, second-party and third-party threats to any aspect of Reliability or Integrity.

Data Security

This requires resistance to environmental, second-party and third-party threats to content, both in remote storage and in transit. Large organisations may be more highly professional, and are capable of investing more in security. On the other hand, they are highly visible and attractive targets. There has already been a report of Amazon's EC2 cloud service being compromised to the extent of having a botnet command and control module inserted (Prince 2009).

Authentication and Authorisation

This refers to the provision of clients with convenient access to data and processes in the cloud, while denying access to imposters. Despite the efforts of many service providers and industry consortia, this remains a very challenging area, even without the added complexities associated with cloud computing (e.g. Hogben 2009).

Susceptibility to Denial of Service Attacks

The existence of multiple hosts represents a safeguard against blockages on a few of them. There remain vulnerabilities, however, including attacks on single-points-of-failure such as cloud managers, databases and users' connections to the Internet, and attacks via the DNS.

4.4 Risk Management Strategies

Given the operational, contingent and security disbenefits and risks, it is essential that an organisation contemplating the delegation of significant data or processing to a cloud computing service-provider conduct a risk assessment, devise a risk management strategy, and proceed only if the risks are able to be satisfactorily addressed.

For many organisations, some IT services are a business survival factor, in the sense that, after some period of unavailability of the services, the company would lose revenue, market-share, or credibility with key stakeholders, to such an extent that it would not be able to sustain its operations. Where an organisation uses internal IT resources, its survival depends on its own actions. Outsourcing without contingency plans, on the other hand, especially to a cloud of virtual servers, passes control over the organisation's future into the hands of unknown others, with limited legal recourse against malperformance. Investment in fallback mechanisms, and hence in local replication of key data and processing capabilities, appears to be an absolutely essential element of all significant corporate uses of cloud computing.

Even where the consequences of failure would be less than catastrophic, cloud computing adds new layers of barriers between the user organisation and the resources on which they depend. The user organisation accordingly faces much greater difficulties in analysing problems and exercising control.

The conventional approach to managing the operational disbenefits and risks discussed in this section is through commercial contracts and Service Level Agreements (SLAs). However, it appears to be envisaged that cloud computing service-providers will simplify their business model by imposing standard terms, and leaving no scope for negotiation about either the commercial terms or the SLA.

Moroever, there appear to be no standards for SLAs, although some guides and checklists exist. The design, negotiation and administration of SLAs are fraught with difficulties in any IT outsourcing activity, but especially so where level of delegation is so substantial. In view of the opaqueness of the services and the commercial and technical conditions, and the user organisation's lack of expertise and lack of information needed to conduct an investigation and to enforce the applicable terms, some form of third-party certification appears to be essential.

A further risk management strategy that can be considered is the use of multiple contracts, or the proven ability to quickly convert to another provider that can deliver the required services at short notice.

5 Business Disbenefits and Risks

The previous section identified a variety of technical factors that may result in harm to the organisational or individual user. This section focusses on business factors relating to acquisition of the service, ongoing usage, privacy and compliance aspects.

5.1 Acquisition Risks

This sub-section considers the factors that arise when an organisational or individual user is in the process of adopting a cloud computing service. This and the following sub-section draw on relevant parts of a normative template for consumer-marketer communications, which has been previously developed and applied by the author and his colleagues (Clarke 2006, Clarke 2008, Svantesson & Clarke 2010). It also reflects reviews of Google's terms of service conducted several years ago (Clarke 2006b) and refreshed as part of the current project.

It is a common feature of outsourcing arrangements that the user organisation delegates decisions about details to the provider, and has only limited understanding of the way in which the services will be performed. The cloud computing service-provider naturally utilises its market power to dictate the contractual terms, with little or no prior input from intended customers, and with little or no scope for negotiation, resulting in considerable advantage to the provider. Particular aspects that are likely to result in problems for outsourcers include:

- the limited amount of information that is available at the time of purchase, and subsequently;
- the limited availability of information that can be used to compare alternative providers' offerings;
- the avoidance of warranties and indemnities in the event of non-delivery, error or failure; and
- the choice of a jurisdiction for any legal actions that suits the provider but not necessarily the user organisation.

5.2 Ongoing Usage Risks

Once a cloud computing service has been acquired, a user or user organisation has an ongoing interest in quality of service. When problems arise, the user needs clarity about the processes for dealing with them. The problems arising at the time of acquiring the services lay a foundation for further problems later. The user is dependent on the provider. The user is ill-equipped to determine whether, when and how to switch to an alternative provider. The user is particularly badly placed to consider bringing the services back in-house. The 'lock-in' effect delivers the provider with an increase in its market power, which it is able to exploit in such ways as:

- price-hikes within the existing tariff;
- tariff changes;

- second-party (service-provider) access to and use of content;
- use of content to the advantage of third parties (e.g. the service-provider's strategic partners and customers); and
- adaptations to the service.

The focus in this paper is primarily on the interests of the user organisation or individual user acquiring cloud computing services. However, some allowance also needs to be made for the interests of customers of the outsourcer. They are concerned about the reliability and quality of services, and recourse in the event of problems.

The terms of the SLA may be of value in addressing these risks, but in many cases the value will be undermined by disparity in organisational size and market power, and jurisdictional distance between the parties.

5.3 Privacy Risks

This sub-section considers privacy impacts of cloud computing. Mere compliance with data protection laws is considered in the following sub-section. The focus here is on public expectations and policy issues that are not, or not yet, reflected in the law. Several early privacy analyses have been published, variously by a Privacy Commissioner (Cavoukian 2009), an industry association (Gellman 2009), a news service (Harris 2009), an IT provider (MS 2009b), a commercial publisher (Mather et al. 2009), and at an academic conference (Creese et al. 2009, Pearson & Charlesworth 2009). At least one privacy advocacy organisation maintains a resource-page (EPIC 2009), and at least one has issued a policy statement on the matter (APF 2009).

A cluster of privacy concerns relates to unauthorised actions by second parties (i.e. cloud computing service-providers and their sub-contractors). Unauthorised actions may include:

- access to content;
- use of content;
- disclosure of content to other parties; and
- retention of content.

A second group of problems relate to unauthorised actions by third parties. These include all of the issues noted immediately above, but importantly also encompass access to content taking advantage of inadequate security in storage or in transit ('hacking', 'cracking' and interception).

Privacy laws generally have very limited enforceability, being of the nature of guidance rather than genuine regulation. Assurance might be sought through the terms of the contract or the SLA, or a Privacy Policy Statement published by the relevant service-provider(s). Previous research has been conducted in this area (Clarke 2005b, 2006a), and updated as part of the current project. In comparison with a template that expresses reasonable consumer expectations (Clarke 2005a), providers generally fall a long way short of reasonable consumer expectations.

Additional problems arise where the cloud crosses jurisdictional boundaries. Even where a user somehow gathers sufficient information about a privacy breach, they are

likely to face difficulties initiating and pursuing actions in the jurisdictional location(s) in which the breach has occurred. Further considerations are that enforcement based on contract is limited to the parties to the contract, and that actions in foreign jurisdictions are slow, expensive and uncertain. This is a major concern, especially in countries that have extremely weak protections in relation to Trans-Border Data Flows, such as Australia.

The user organisation has an obligation to ensure that privacy and security risks are satisfactorily addressed. Given the seriousness of the problem, it is essential that both service-providers and user organisations, when contemplating the application of cloud computing to personal data, undertake Privacy Impact Assessments (PIAs) (Clarke 2009).

5.4 Compliance Risks

Joint et al. (2009) considered legal aspects of the adoption of cloud computing, and concluded that there were serious compliance difficulties, particularly for European companies. For example, data protection laws generally specify, in accordance with EU Directive 95/46 Article 25, that the transfer of personal data across borders can only be to countries that provide "an adequate level of protection". This leads to a fundamental requirement of cloud computing service providers that they be able to assure their customers that personal data will not be passed to, or through, countries other than those approved by the customer. This is far from easy for a service-provider to do, and there are signs that some may prefer to subvert such provisions by contriving the placement of hosts beyond the reach of any jurisdiction (Miller 2008).

In addition to data protection law, Joint et al. drew attention to the law of confidence and financial services regulation including auditability. CSA (2009) offered a similar list, but added evidence discovery. Haslam (2009) further added the risk of exposure of non-US information that happens to be stored in the USA under cloud computing arrangements.

The analysis conducted in the preceding sub-sections suggests that organisations that use cloud computing to handle customer data may also risk being in breach of consumer protection laws. Laws in many countries require information disclosure about the service and the provider, prohibit some kinds of behaviour in relation to contract formation, and embody controls over misleading and deceptive conduct, and misrepresentation. In many jurisdictions, however, such laws are ineffective, and hence a corporation that relies on a cloud computing provider in a foreign jurisdiction might be found to have breached its own, domestic consumer protection laws, for example by failing to apply due care and skill in the provision of the service, and failing to ensure fitness for purpose.

In addition to express obligations, corporations may need to demonstrate adequate performance against various business and technical standards. At the level of company boards, Directors have obligations in relation to asset protection, due diligence, business continuity and risk management. Various of the technical and business risks discussed in this section may be sufficiently significant to demand attention not only from business managers and senior executives, but also from the Board room.

6 Conclusions

In widely-reported comments made in September 2008, the Free Software Foundation's Richard Stallman said that cloud computing forces people to hand over control of their information to a third party. "One reason you should not use Web applications to do your computing is that you lose control. It's just as bad as using a proprietary program. Do your own computing on your own computer with your copy of a freedom-respecting program. If you use a proprietary program or somebody else's Web server, you're defenseless. You're putty in the hands of whoever developed that software" (Johnson 2008).

The analysis reported on in this paper puts flesh to those claims. It has provided some clarity about the term's meaning, and about the benefits that may be available to user organisations and individual users. Most crucially, it has clarified the very substantial downsides and risks involved in using cloud computing. It lays a foundation for guidance to user organisations in determining the circumstances in which cloud computing is an appropriate approach to adopt.

User organisations need to appreciate the nature, benefits, disadvantages and risks, and carefully consider the extent to which the services may be applicable to their needs. Where cloud computing is adopted, risk management strategies need to be devised and carefully implemented. Because services are unlikely to be subject to clear law, policy and standards, fallback positions and disaster plans are essential. For many organisations, and many business functions within them, it is essential that key data be mirrored within the organisation, and that internal processing capabilities and capacity be sufficient to enable interim operation and subsequent resumption of normal business.

Individual users may use cloud computing, provided that they place no great value on the ongoing availability of the data or the services. They need to assume that the terms of service will be opaque, that they will be advantageous to the provider rather than the user, and that the user will have little or no capability to enforce such limited rights as they may appear to have.

Research is needed in a variety of areas. The generic analysis reported on in this paper needs to be applied to specific categories of cloud computing services. The terms offered by particular providers need to be evaluated against these risks, and against the templates used in this paper, or other checklists of a similar nature. Also of considerable value would be case studies of various successful, and especially unsuccessful, applications.

In Armbrust et al. (2008), it was noted that "past efforts at utility computing failed, and we note that in each case one or two ... critical characteristics were missing" (p. 5). If cloud computing is to be more than just another marketing buzz-phrase that leaves corporate wreckage in its wake, service-providers need to invest a great deal in many aspects of their infrastructure, platforms, applications, and terms of service.

References

Aeon (2009) 'Cloud Security: Want Some Fake Fries With That Vapor Shake?' Aeon Security Blog, 30 December 2009, at http://www.theaeonsolution.com/security/?p=131

Anderson R.W. (2008) 'The Cloud Services Stack — Infrastructure', rwandering.net, July 2008, at http://rwandering.net/2008/07/28/the-cloud-services-stack-infrastructure/

APF (2009) 'Policy Statement re Cloud Computing' Australian Privacy Foundation, November 2009, at http://www.privacy.org.au/Papers/CloudComp-0911.html

Avizienis A., Laprie J.C., Randell B. & Landwehr C. (2004) 'Basic Concepts and Taxonomy of Dependable and Secure Computing' IEEE Trans. Dependable and Secure Computing 1,1 (2004) 11- 33

Brodkin J. (2008a) 'Gartner: Seven cloud-computing security risks' InfoWorld, July 2008, at http://www.infoworld.com/print/36853

Brodkin J. (2008b) 'Loss of customer data spurs closure of online storage service 'The Linkup' Network World, August 2008 , at http://www.networkworld.com/news/2008/081108-linkup-failure.html

Brodkin J. (2009) 'Rackspace apologizes for cloud outage, prepares to issue service credits' Network World, 5 November 2009, at http://www.networkworld.com/news/2009/110509-rackspace-outage-apology.html?fsrc=netflash-rss

Buyya R., Yeo C.S., Venugopal S., Broberg J. & Brandic I. (2009a) 'Cloud computing and emerging IT platforms: Vision, hype, and reality for delivering computing as the 5th utility' (2009) Future Generation Computer Systems 25 (January 2009) 599–616, at http://www.buyya.com/gridbus/papers/Cloud-FGCS2009.pdf

Buyya R., Pandey S. & Vecchiola C. (2009b) 'Cloudbus Toolkit for Market-Oriented Cloud Computing' Proc. 1st Int'l Conf. on Cloud Computing, Beijing, 1-4 December 2009, at http://arxiv.org/pdf/0910.1974

Cavoukian A. (2009) 'Privacy in the clouds: A white paper on privacy and digital identity' Information and Privacy Commissioner of Ontario, 2009, at http://www.ipc.on.ca/images/Resources/privacyintheclouds.pdf

CCJ (2009) 'Twenty-One Experts Define Cloud Computing' Cloud Computing Journal, January 2009, at http://cloudcomputing.sys-con.com/read/612375_p.htm

Clarke R. (2005a) 'Privacy Statement Template' Xamax Consultancy Pty Ltd, 19 December 2005, at http://www.rogerclarke.com/DV/PST.html

Clarke R. (2005b) 'Evaluation of Google's Privacy Statement against the Privacy Statement Template' Xamax Consultancy Pty Ltd, 20 December 2005, at http://www.rogerclarke.com/DV/PST-Google.html

Clarke R. (2006a) 'A Pilot Study of the Effectiveness of Privacy Policy Statements' Proc. 19th Bled eCommerce Conf., Slovenia, 5-7 June 2006, at http://www.rogerclarke.com/EC/PPSE0601.html

Clarke R. (2006b) 'A Major Impediment to B2C Success is ... the Concept 'B2C' Proc. ICEC'06, Fredericton NB, Canada, 14-16 August 2006, at http://www.rogerclarke.com/EC/ICEC06.html

Clarke R. (2008) 'B2C Distrust Factors in the Prosumer Era' Proc. CollECTeR Iberoamerica, Madrid, 25-28 June 2008, pp. 1-12, at http://www.rogerclarke.com/EC/Collecter08.html

Clarke R. (2009) 'Privacy Impact Assessment: Its Origins and Development' Computer Law & Security Review 25, 2 (April 2009) 123-135, at http://www.rogerclarke.com/DV/PIAHist-08.html

Crandell B. (2008) 'Defogging Cloud Computing: A Taxonomy' gigaom.com, June 2008, at http://gigaom.com/2008/06/16/defogging-cloud-computing-a-taxonomy/

Creese S., Hopkins P., Pearson S. & Shen Y. (2009) 'Data Protection-Aware Design for Cloud Computing' Proc. CloudCom 2009, Beijing, Springer LNCS, December 2009, at http://www.hpl.hp.com/techreports/2009/HPL-2009-192.pdf

CSA (2009) 'Security Guidance for Critical Areas of Focus in Cloud Computing' Cloud Security Alliance, April 2009, at http://www.cloudsecurityalliance.org/csaguide.pdf

EPIC (2009) 'Resources on Cloud Computing' Electronic Privacy Information Center, Washington DC, 2009, at http://epic.org/privacy/cloudcomputing/

Fenn J. (1995) 'When to Leap on the Hype Cycle' Gartner Group, January, 1995

Foster I., Zhao Y., Raicu I. & Lu S. (2008) 'Cloud Computing and Grid Computing 360-Degree Compared' Proc. Grid Computing Environments Workshop, 12-16 Nov. 2008

Fowler G.A. & Worthen B. (2009) 'The Internet Industry Is on a Cloud -- Whatever That May Mean' Wall Street Journal, 26 March 2009, at http://online.wsj.com/article/SB123802623665542725.html#printMode

Fried I. (2009) 'Major outage hits T-Mobile Sidekick users' C-Net News, 6 October 6 2009, at http://news.cnet.com/8301-13860_3-10368709-56.html

Gellman R. (2009) 'Cloud Computing and Privacy' World Privacy Forum, 2009, at http://www.worldprivacyforum.org/pdf/WPF_Cloud_Privacy_Report.pdf

Harris L. (2009) 'Perils in the Privacy Cloud' ABC News, 15 Sep 2009, at http://abcnews.go.com/Technology/AheadoftheCurve/privacy-evaporates-computing-cloud/Story?id=8573715&page=1

Haslam H. (2009) 'Clouding Out Global Reality: Can SaaS Systems Work In An International Company?' The Metropolitn Corporate Counsel (September 2009), at http://www.metrocorpcounsel.com/pdf/2009/September/29.pdf

Hogben G. (2009) 'Privacy, Security and Identity in the Cloud' European Network and Information Security Agency (ENISA), June 2009, at http://www.enisa.europa.eu/act/res/other-areas/cloud-computing/Cloud_Identity_Hogben.pdf

ISACA (2009) ' Cloud Computing: Business Benefits With Security, Governance and Assurance Perspectives' ISACA, October 2009, at http://www.isaca.org/AMTemplate.cfm?Section=Deliverables&Template=/ContentManagement/ContentDisplay.cfm&ContentID=53044

Johnson B. (2008) 'Cloud computing is a trap, warns GNU founder Richard Stallman' The Guardian, 29 September 2008, at http://www.guardian.co.uk/technology/2008/sep/29/cloud.computing.richard.stallman

Joint A., Baker E. & Eccles E. (2009) 'Hey, you, get off of that cloud?' Computer Law & Security Review 25, 2 (2009) 270–274

Linden A. & Fenn J. (2003) 'Understanding Gartner's Hype Cycles' Gartner, Strategic Analysis Report R-20-1971, May 2003, at http://carbon.cudenver.edu/~jgerlach/emergingtechnologyOL/FirstReadings/HypeCycleIntro.pdf

Loh L. & Venkatraman N. (1995) 'An Empirical Study of Information Technology Outsourcing: Benefits, Risks, and Performance Implications' Proc. ICIS 1995, Paper 25, at http://aisel.aisnet.org/icis1995/25

Marshall R. (2008) 'Experts urge caution on cloud computing' Secure Computing Magazine, 14 October 2008, at http://www.securecomputing.net.au/News/125405,experts-urge-caution-on-cloud-computing.aspx

Mather T., Kumaraswamy S. & Latif S. (2009) 'Cloud Security and Privacy: An Enterprise Perspective on Risks and Compliance' O'Reilly Media, 2009

Mell P. & Grance T. (2009) 'The NIST Definition of Cloud Computing' National Institute of Standards and Technology, Information Technology Laboratory, Version 15, October 2009 at http://csrc.nist.gov/groups/SNS/cloud-computing/cloud-def-v15.doc

Metz C. (2008) 'Engineer accidentally deletes cloud' The Register, 9 October 2008, at http://www.theregister.co.uk/2008/08/28/flexiscale_outage/

Miller R. (2008) 'Google Planning Offshore Data Barges' Data Centre Knowledge, 6 September 2008, at http://www.datacenterknowledge.com/archives/2008/09/06/google-planning-offshore-data-barges/

MS (2009a) 'Securing Microsoft's Cloud' Microsoft, May 2009, at http://www.globalfoundationservices.com/security/documents/SecuringtheMSCloudMay09.pdf

MS (2009b) 'Privacy in the Cloud Computing Era – A Microsoft Perspective' Microsoft, November 2009

Pearson S. & Charlesworth A. (2009) 'Accountability as a Way Forward for Privacy Protection in the Cloud' Proc. CloudCom 2009, Beijing, Springer LNCS, December 2009, at http://www.hpl.hp.com/techreports/2009/HPL-2009-178.pdf

Prince B. (2009) 'Amazon EC2 Used as Botnet Command and Control' eWeek Security Watch, 11 December 2009, at http://securitywatch.eweek.com/botnets/amazon_ec2_used_as_botnet_command_and_control.html

Raphael J.R. (2009) 'Gmail Outage Marks Sixth Downtime in Eight Months' PC World, 25 February 2009, at http://www.pcworld.com/article/160153/gmail_outage_marks_sixth_downtime_in_eight_months.html

Svantesson D. & Clarke R. (2010) 'A Best Practice Model for eConsumer Protection' Computer Law & Security Review 27, 1 (January 2010)

Vaquero L.M., Rodero-Merino L., Caceres J. & Lindner M. (2009) 'A Break in the Clouds: Towards a Cloud Definition' ACM SIGCOMM Computer Communication Review 39, 1 (January 2009) 50-55, at http://ccr.sigcomm.org/online/files/p50-v39n1l-vaqueroA.pdf

Woloski M. (2008) 'SaaS Taxonomy Map' M. Woloski, July 2008, at http://blogs.southworks.net/mwoloski/2008/07/10/saas-taxonomy-map/

Youseff L., Butrico M. & Da Silva D. (2008) 'Toward a Unified Ontology of Cloud Computing' Proc. Grid Computing Environments Workshop, 2008, at http://www.cs.ucsb.edu/~lyouseff/CCOntology/CloudOntology.pdf