Introducing Cloud Computing

1

What the reader will learn:

- That 'cloud computing' is a relatively new term, and it is important to clearly define what we mean
- Cloud computing is a new delivery model for IT but that it uses established IT resources
- That the concept of abstraction is critical to the implementation of cloud architectures
- Businesses will adopt cloud computing because it offers financial benefits and business agility, not because the technology is inherently 'better'

1.1 What Is Cloud Computing?

Everybody seems to be talking about cloud computing. As technology trends go, cloud computing is generating a lot of interest, and along with that interest is a share of hype as well. The aim of this book is to provide you with a sophisticated understanding of what cloud computing is and where it can offer real business advantage. We shall be examining cloud computing from historical, theoretical and practical perspectives, so that you will know *what* to use, in *which* situation, and *when* it will be most appropriate.

So first of all, just what is cloud computing? This isn't such a silly question. That many things now attract the cloud computing badge, that it is difficult to understand what cloud actually means.

In a nutshell, cloud computing is a means by which computational power, storage, collaboration infrastructure, business processes and applications can be delivered as a utility, that is, a service or collection of services that meet your demands. Since services offered by cloud are akin to a utility, it also means that you only pay for what you use. If you need extra processing power quickly, it is available for use in an instant. When you've finished with the extra power and revert back to your nominal

R. Hill et al., *Guide to Cloud Computing: Principles and Practice*, Computer Communications and Networks, DOI 10.1007/978-1-4471-4603-2_1,
© Springer-Verlag London 2013

usage, you will only be billed for the short time that you needed the extra boost. So you don't need to invest in a lot of hardware to cater for your peak usage, accepting that for most of the time it will be underutilised. This aspect of the cloud is referred to as *elasticity* and is an extremely important concept within cloud computing.

That's the short answer and not necessarily the key to becoming an expert in cloud computing; for some extra information, read on. To understand what makes a cloud different from other established models of computing, we shall need to consider the conceptual basis of computing as a utility and how technology has evolved to date.

1.2 Utility Computing

Utility computing was discussed by John McCarthy in the 1960s whilst working at Massachusetts Institute of Technology (McCarthy 1983), and the concept was thoroughly expanded by Douglas Parkhill in 1966 (The Challenge of the Computing Utility, Parkhill 1966). Parkhill examined the nature of utilities such as water, natural gas and electricity in the way they are provided to create an understanding of the characteristics that computing would require if it was truly a utility. When we consider electricity supply, for example, in the developed world, we tend to take it for granted that the actual electrical power will be available in our dwellings. To access it, we plug our devices into wall sockets and draw the power we need. Every so often we are billed by the electricity supply company, and we pay for what we have used. In the summer time, the daylight hours are longer and we place less demand on devices that provide lighting, hot water or space heating. During the winter months, we use electric lighting and space heating more, and therefore, we expect our bills to reflect the extra usage we make of the utility. Additionally, we do not expect the electricity to 'run out'; unless there is a power cut, there should be a never-ending supply of electricity.

So the same goes for computing resources as a utility. We should expect the resource to be available where we want, by plugging into or by accessing a network. The resource should cater for our needs, as our needs vary, and it should appear to be a limitless supply. Finally, we expect to pay only for what we use. We tend to consider the provision of utilities as services.

1.3 Service Orientation

The term *service orientation* refers to the clear demarcation of a function that operates to satisfy a particular goal. For instance, businesses are composed of many discrete services that should sustainably deliver value to customers now and in the future. Utility companies offer their services in the form of energy supply, billing and perhaps, as energy conservation becomes more widespread, services that support a customer's attempt to reduce their energy consumption. The services that are offered to the consumer are likely to be aggregations of much finer-grained services that operate internally to the business. It is this concept of *abstraction*,

combined with object-oriented principles such as *encapsulation* and *cohesion*, that helps define services within an organisation.

Service-oriented architecture (SOA) utilises the principle of service orientation to organise the overall technology architecture of an enterprise. This means that technology is selected, specified and integrated to support an architectural model that is specified as a set of services. Such an approach results in technologically unique architectures for each enterprise, in order to realise the best possible chance of supporting the services that the business requires. However, whilst the overall architecture may appear bespoke, the underlying services are discrete and often reusable and therefore may be shared even between organisations. For instance, the processing of payroll information is common to most enterprises of a certain size and is a common choice for service outsourcing to third-party suppliers.

From an organisation's perspective, SOA has some key advantages:

- The adoption of the principles of service orientation enables commonly utilised functionality to be reused, which significantly simplifies the addition of new functionality, since a large portion of the existing code base is already present. Additionally, the emergence of standard protocols for service description and invocation means that the actual service is abstracted away from the implementation program code, so it doesn't matter if the constituent parts of a newly composed service are implemented in different ways, as long as their specification conforms to a commonly declared interface contract.
- Changes in business demand that require new services to be specified can be accommodated much easier, and it is quicker to react to business market forces. This means that an SOA is much more fleet of foot, enabling new business opportunities to be explored quickly with less cost.
- The abstraction of service also facilitates consideration of the enterprise's performance at the process level; quality of service (QoS), lead times and defect rates become more obvious measures to observe and therefore targets to specify, since the underlying complexity is shrouded behind a service declaration.
- Tighter integration along value chains is enabled, as a particular functionality can be made available as a service between an enterprise and its satellite suppliers. A supplier may deal with several business customers, and it might not be practical to adopt a number of different systems to integrate with. SOA simplifies this by the publication of services that suppliers can 'hook into' with their own systems. This has the added advantage that any changes to a customer's system are encapsulated behind the service description, and therefore, no other modifications will be required from those that consume that service.

Service orientation and its architectural model SOA are key concepts for the realisation of utility computing. Now, we shall consider some technological developments that can support this realisation. Later, in Chap. 5, we shall encounter SOA again, where you will be building a Google App as an exemplar use of web services.

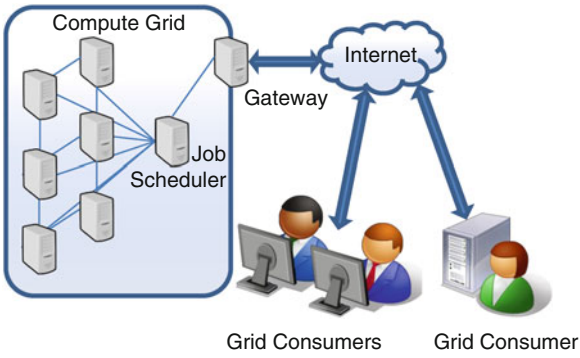

Fig. 1.1 Overview of grid computing architecture

1.4 Grid Computing

Grid computing emerged in the 1990s, as Ian Foster and Carl Kesselman suggested that access to compute resources should be the same as connecting to a power grid to obtain electricity (Foster and Kesselman 1999). The need for this was simple: Supercomputers that could process large data sets were prohibitively expensive for many areas of research. As an alternative, the connection and coordination of many separate personal computers (PC) as a grid would facilitate the scaling up of computational resources under the guise of a virtual organisation (VO). Each user of the VO, by being connected to the grid, had access to computational resources far greater than they owned, enabling larger scientific experiments to be conducted by spreading the load across multiple machines. Figure 1.1 gives a brief overview of a grid architecture. A number of separate compute and storage resources are interconnected and managed by a resource that schedules computational jobs across the grid. The collective compute resource is then connected to the Internet via a gateway. Consumers of the grid resource then access the grid by connecting to the Internet.

As network speeds and storage space have increased over the years, there has been a greater amount of redundant computational resource that lays idle. Projects such as the Search for Extraterrestrial Intelligence (SETI@HOME, http://setiathome.berkeley.edu/) have made use of this by scavenging processing cycles from PCs that are either doing nothing or have low demands placed upon them. If we consider how computer processors have developed in a relatively short time span, and then we look at the actual utilisation of such computational power, particularly in the office desktop environment, there are a lot of processor cycles going spare. These machines are not always used during the night or at lunch breaks, but they are often left switched on and connected to a network infrastructure. Grid computing can harness this wastage and put it to some predefined, productive use.

One characteristic of grid computing is that the software that manages a grid should enable the grid to be formed quickly and be tolerant of individual machines (or nodes) leaving at will. If you are scavenging processor cycles from someone

else's PC, you have to be prepared for them turning their machine off without prior warning. The rapid setup is required so that a grid can be assembled to solve a particular problem. This has tended to support scientific applications, where some heavy analysis is required for a data set over a short period, and then it is back to normal with the existing resources when the analysis is done. Collaboration and contribution from participants has been generally on a voluntary basis, which is often the basis of shared ventures in a research environment.

Whilst grid computing has started to realise the emergence of computing resources as a utility, two significant challenges have hindered its uptake outside of research. Firstly, the ad hoc, self-governing nature of grids has meant that it is difficult to isolate the effect of poorly performing nodes on the rest of the grid. This might occur if a node cannot process a job at a suitable rate or a node keeps leaving the grid before a batch job is completed. Secondly, the connection of many machines together brings with it a heterogeneous collection of software, operating systems and configurations that cannot realistically be considered by the grid software developer. Thus, grid applications tend to lack portability, since they are written with a specific infrastructure in mind.

1.5 Hardware Virtualisation

Hardware virtualisation is a developing technology that is exploiting the continued increase in processor power, enabling 'virtual' instances of hardware to execute on disparate physical infrastructure. This technology has permitted organisations such as data centres to improve the utilisation and management of their own resources by building virtual layers of hardware across the numerous physical machines that they own. The virtualisation layer allows data centre management to create and instantiate new instances of virtual hardware irrespective of the devices running underneath it. Conversely, new hardware can be added to the pool of resource and commissioned without affecting the virtualised layer, except in terms of the additional computational power/storage/memory capability that is being made available. Figure 1.2 illustrates the key parts of a virtualised architecture. Working from the physical hardware layer upwards, firstly there is a *hypervisor*. The role of the hypervisor is to provide a means by which virtual machines can access and communicate with the hardware layer, without installing an operating system. On top of the hypervisor, *virtual machines* (VM) are installed. Each VM appears to function as a discrete computational resource, even though it does not physically exist. A host *operating system* (OS) is installed upon each VM, thus enabling traditional computing applications to be built on top of the OS.

Virtualisation offers three key advantages for data centre management. Firstly, applications can be confined to a particular virtual machine (VM) appliance, which increases security and isolates any detrimental effect of poor performance on the rest of the data centre. Secondly, the consolidation of disparate platforms onto a unified hardware layer means that physical utilisation can be better managed, leading to increased energy efficiency. Thirdly, virtualisation allows guest operating systems

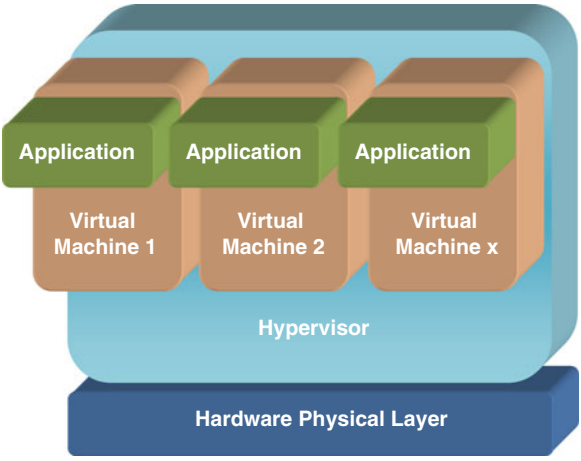

Fig. 1.2 Virtualisation overview

to be stored as snapshots to retain any bespoke configuration settings, which allows images to be restored rapidly in the event of a disaster. This feature also facilitates the user capture of provenance data so that particular situations can be realistically recreated for forensic investigation purposes or to recreate a specific experimental environment. Virtualisation is discussed in more detail in Chap. 4.

1.6 Autonomic Computing

As computing technology becomes more complex, there is a corresponding desire to delegate as much management as possible to automated systems. *Autonomic computing* attempts to specify behaviours that enable the self-management of systems. Self-configuration, self-healing, self-optimising and self-protection (otherwise known as self-CHOP) are the four principles defined by IBM's autonomic computing initiative (IBM Research 2012). If we consider the cloud computing concept of elasticity, we can see that to obtain the 'resource-on-demand' feature will require a variety of computational resources to be configured and, once running, optimised for performance.

If we now consider a grid architecture as a computational resource, then the operations described above will need to take into account some more aspects particular to the technologies involved, including disparate and heterogeneous hardware and software standards. Finally, if we add to the mix hardware virtualisation, there will be a requirement to instantiate and migrate virtual machines (VM) across disparate hardware, dynamically as demand dictates. Such is the complexity of myriad physical and virtualised hardware architectures and software components, that it is essential that this management is automated if true, seamless elasticity is to be realised.

We have now explored the key concepts and technologies that have shaped the emergence of cloud computing, so we shall now explore a more formal definition and observe how this informs the present description of cloud computing architectures.

1.7 Cloud Computing: A Definition

It won't take you long to find a number of 'definitions' of cloud computing. The World Wide Web is awash with attempts to capture the essence of distributed, elastic computing that is available as a utility. There appears to be some stabilisation occurring with regard to an accepted definition, and for the purposes of this book, we'll be persevering with that offered by the National Institute of Standards and Technology (NIST):

> Cloud computing is 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. This cloud model is composed of five essential characteristics, three service models, and four deployment models.
>
> NIST, US Department of Commerce, http://csrc.nist.gov/publications/nistpubs/800-145/SP800-145.pdf

The essential characteristics that NIST's definition refers to are as follows:

- *On-demand self-service*. Traditionally, hosted computing has enabled consumers to outsource the provision of IT infrastructure, such as data storage, so that hardware purchases could be minimised. However, whilst these solutions allowed customers to increase the storage available without purchasing any extra hardware, the request for data storage was typically an order that was fulfilled some time later. The time lag between request and actual availability meant that such increases had to be planned for and could not be depended upon as a reactive resource. Cloud computing should incorporate sufficient agility and autonomy, that requests for more resource are automatically and dynamically provisioned in real time, without human intervention.
- *Broad network access*. As a utility, cloud computing resources must be available over networks such as the Internet, using established mechanisms and standard protocols. Access devices can include (though are not limited to) personal computers, portable computers, mobile phones and tablet devices.
- *Resource pooling*. This characteristic brings together aspects of grid computing (where multiple compute resources are connected together in a coordinated way) and hardware virtualisation. The virtualised layer enables the resources of a cloud computing provider to be pooled together into one large virtual resource, enabling large-scale efficiencies to be achieved by the dynamic management of hardware and virtualised resources. This results in the appearance of homogenous resources to the consumer, without indicating the physical location or granularity of that resource.
- *Rapid elasticity*. Requests for extra resource are self-managed and automatic in relation to demand. From the consumer's perspective, the supply of compute resources is limitless.
- *Measured service*. In the same way that energy usage can be monitored, controlled and reported, cloud computing resource providers dynamically optimise the underlying infrastructure and provide a transparent metering service at a level of abstraction that is relevant to the consumer.

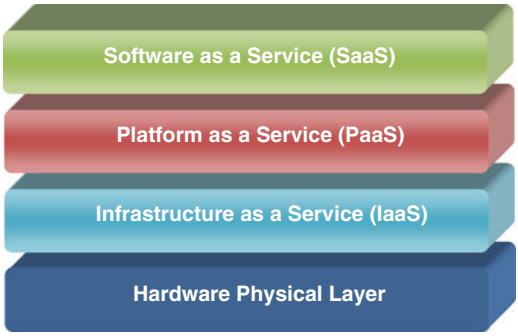

Fig. 1.3 Cloud service models

One theme that is emerging here is that of abstraction; the characteristics above are reliant upon a fundamental architecture of hardware resources that are discrete and varied, upon which there is an abstraction layer of software that realises the characteristics of cloud computing. The physical hardware resource layer includes processor, storage and networking components, and the abstraction layer consists of at least a self-managed virtualisation infrastructure.

1.8 Cloud Computing Service Models

Of course, in cloud-speak we refer to services, and there are three categories of service model described by NIST as illustrated in Fig. 1.3. Working from the physical layer upwards, the first service model layer is known as Infrastructure as a Service (IaaS).

IaaS is usually the lowest level service available to a cloud computing consumer and provides controlled access to a virtual infrastructure upon which operating systems and application software can be deployed. This can be seen as a natural extension of an existing hardware provision, without the hassle and expense of buying and managing the hardware. As such, there is no control over the physical hardware, but the consumer retains control over operating system parameters and some aspects of security. There is a trend emerging for 'bare metal' services, where access to the hardware at its most basic is provided, but this is more akin to traditional data centre or 'hosting' services. For the majority of potential cloud consumers, there is a desire to move away from as much of the detail as possible and therefore progress upwards through the cloud service model stack.

Platform as a Service (PaaS) sits atop IaaS. This layer is ready for applications to be deployed, as the necessary operating system and platform-related tools such as language compilers are already installed and managed by the cloud computing provider. Consumers may be able to extend the existing tool set by installing their own tools, but absolute control of the infrastructure is still retained by the provider. Thus, the consumer has control over application development, deployment and configuration, within the confines of the hosted environment. This situation has most in common with traditional web hosting, where consumers rented remote servers that had existing

development platforms installed upon them. The key difference with cloud computing in this case, however, is the rapid provisioning or elasticity; classic web hosting relied upon manual management of provisioning and therefore required human intervention if demand increased or decreased.

Finally (for the NIST definition), there is Software as a Service (SaaS). This service model abstracts the consumer away from any infrastructure or platform level detail by concentrating upon the application level. Applications are available via thin client interfaces such as internet browsers or program interfaces such as mobile phone apps. Google's Gmail is one popular example of a cloud computing application. An organisation can adopt Gmail and never concern itself with hardware maintenance, uptime, security patching or even infrastructure management. The consumer can control parameters within the software to configure specific aspects, but such interventions are managed through the interface of the application. The end user gets an email service and does not worry as to how it is provided.

So far, we have described the essential characteristics of cloud computing and then three different service models. As the abstraction concept develops, consumers are finding new ways of using cloud computing to leverage business advantage through the creation of a Business Process as a Service model (BPaaS). Strictly speaking, this sits within SaaS and is not a fourth layer which would fall outside of the NIST definition. We shall revisit this service model later in Chap. 4, so for the time being, we shall consider the models by which cloud computing can be deployed.

1.9 Cloud Computing Deployment Models

A *public cloud*, as its name implies, is available to the general public and is managed by an organisation. The organisation may be a business (such as Google), academic or a governmental department. The cloud computing provider owns and manages the cloud infrastructure. The existence of many different consumers within one cloud architecture is referred to as a multi-tenancy model.

Conversely, a *private cloud* has an exclusive purpose for a particular organisation. The cloud resources may be located on or off premise and could be owned and managed by the consuming organisation or a third party. This may be an example of an organisation who has decided to adopt the infrastructure cost-saving potential of a virtualised architecture on top of existing hardware. The organisation feels unable to remotely host their data, so they are looking to the cloud to improve their resource utilisation and automate the management of such resources. Alternatively an organisation may wish to extend its current IT capability by using an exclusive, private cloud that is remotely accessible and provisioned by a third party. Such an organisation may feel uncomfortable with their data being held alongside a potential competitor's data in the multi-tenancy model.

Community clouds are a model of cloud computing where the resources exist for a number of parties who have a shared interest or cause. This model is very similar to the single-purpose grids that collaborating research and academic organisations have created to conduct large-scale scientific experiments (e-science). The cloud is

owned and managed by one or more of the collaborators in the community, and it may exist either on or off premise.

Hybrid clouds are formed when more than one type of cloud infrastructure is utilised for a particular situation. For instance, an organisation may utilise a public cloud for some aspect of its business, yet also have a private cloud on premise for data that is sensitive. As organisations start to exploit cloud service models, it is increasingly likely that a hybrid model is adopted as the specific characteristics of each of the different service models are harnessed. The key enabler here is the open standards by which data and applications are implemented, since if portability does not exist, then vendor lock-in to a particular cloud computing provider becomes likely. Lack of data and application portability has been a major hindrance for the widespread uptake of grid computing, and this is one aspect of cloud computing that can facilitate much more flexible, abstract architectures.

At this point, you should now have a general understanding of the key concepts of cloud computing and be able to apply this knowledge to a number of common use cases in order to hypothesise as to whether a particular cloud service model might be appropriate. The next part of this chapter will dig a bit deeper into the deployment models and explore some finer-grained challenges and opportunities that cloud computing presents.

1.10 A Quick Recap

Before we proceed, let us just quickly summarise what we understand by cloud computing:

- It's a model of computing that abstracts us away from the detail. We can have broad network access to computing resources without the hassle of owning and maintaining them.
- Cloud computing providers pool resources together and offer them as a utility. Through the use of hardware virtualisation and autonomic computing technologies, the consumer sees one homogenous, 'unlimited' supply of compute resource.
- Computing resources can be offered at different levels of abstraction, according to requirements. Consumers can work at infrastructure level (IaaS) and manage operating systems on virtualised hardware, at platform level (PaaS) using the operating systems and development environments provided, or at application level (SaaS), where specific applications are offered by the provider to be configured by the consumer.
- Cloud computing provides metered usage of the resource so that consumers pay only for what they use. When the demand for more computing resource goes up, the bill increases. When the demand falls, the bill reduces accordingly.
- Cloud computing can be deployed publicly in a multi-tenancy model (public cloud), privately for an individual organisation (private cloud), across a community

of consumers with a shared interest (community cloud), or a mixture of two or more models (hybrid cloud).

1.11 Beyond the Three Service Models

The explanations and discussions so far have allowed us to gain a broad understanding of cloud computing. However, like most things in life, it isn't that simple. When chief executive officers declare that an organisation will embrace 'the cloud', the chief information officer (CIO) may be less enthusiastic. We shall now consider more deeply some of the business drivers and service models for cloud adoption and explore the issues that these drivers can present.

1.11.1 The Business Perspective

Large IT vendors have realised for some time that new technology is sold most successfully on its ability to improve profitability. Grid computing and service-oriented architecture (SOA) are two relatively recent examples. Grid computing has demonstrated massive benefits when disparate compute resources are harnessed together to do supercomputing on the cheap. The problem was that the software and protocols that made these large distributed systems perform were inaccessible to those outside of the grid community. Vendors such as IBM and Oracle have both attempted to sell the advantages to business of grid computing, but the lack of realisation of the concept of utility (which informed the selection of the name 'grid') has meant that insufficient consumers were interested and the ultimate benefits could not be enjoyed.

SOA has had a similar 'reduce your business costs' drive over the years, with many organisations reporting an overall increase in expenditure after the costs of migrating to SOA have been accounted for. So what is different about cloud computing?

One of the attractions of cloud computing is the rapid provisioning of new compute resources without capital expenditure. If the marketing director makes claims about a new market niche, then it is much more cost-effective to experiment with new products and services, since cloud computing removes traditional barriers such as raising capital funds, lengthy procurement procedures and human resource investment. Also, if cloud computing is already part of the organisation's IT infrastructure, then new requests merely become additional demands upon the systems, rather than designing and specifying new systems from scratch. Business agility is therefore one key driver for the adoption of cloud computing.

The other key business driver is the potential reduction in ongoing capital expenditure costs afforded by cloud computing. As the use of IT becomes more sophisticated, greater demands are placed upon IT fundamentals such as data

storage, and if the requirements fluctuate significantly, the pay-per-use model of cloud computing can realise operational savings beyond the costs of the potential extra hardware requirement.

1.12 When Can the Service Models Help?

1.12.1 Infrastructure as a Service

As described earlier, IaaS is about servers, networking and storage delivered as a service. These resources will actually be virtualised, though the consumer wouldn't know any different. The resources may come with or without an operating system. IaaS is a form of computing rental where the billing is related to actual usage, rather than ownership of a discrete number of servers. When the consumer wants more 'grunt', the IaaS management software dynamically provisions more resources as required. Typically, there will be an agreed limit between the consumer and the provider, beyond which further authorisation is required to continue scaling upwards (and thus incur extra cost). IaaS is particularly suited to organisations who want to retain control over the whole platform and software stack and who need extra infrastructure quickly and cheaply. For instance, the research and development department of an organisation may have specific applications that run on optimised platforms. Sporadically, applications are required to process massive data sets. Using a cloud, it would cost the same to have 500 processors run for 1 hour, as it does to run 1 processor for 500 hours, so the research unit opts for speed without having to invest in hardware that would be nominally underutilised.

1.12.2 Platform as a Service

PaaS has parallels with web hosting, in that it is a complete set of software that enables the complete application development life cycle within a cloud. This includes the tools for development and testing as well as the actual execution environment. As with IaaS, the resources are dynamically scaled, and for the most part, this is handled transparently by the cloud provider without making any extra demands upon the developer. For specialist applications that require low-level optimisation, either IaaS or a private cloud is more suitable. One of the potential drawbacks of PaaS is lack of portability and therefore vendor lock-in, as you are developing applications with the tool sets that are supplied by the cloud provider. If, at a later date, you would like to move provider or you want to use another cloud service concurrently, there may be a substantial effort required to port your application across to another vendor's cloud platform. PaaS is a good option if your existing application's development environment is matched by that of a cloud provider or if you would like to experiment with new products and services that can be rapidly composed from pre-existing services that are provided by the platform.

1.12.3 Software as a Service

In some ways, SaaS is the easiest way into cloud computing. You see some software and you try it out for a limited time. If you like it, you continue and start paying to use it, otherwise you look for something else. The software automatically scales to the number of users you have (but you don't notice this), and your data is backed up. You will probably have to invest a bit of time in getting your existing data into the application, and any tweaks to existing systems that you have may also require some work to get them to connect to your new cloud application. SaaS is useful if you are in the situation whereby a legacy application you own has been replicated by a SaaS provider or if a particular SaaS application offers a capability that you don't currently have but can see the business benefit of having it. Customer Relationship Management (CRM) is one example; many organisations operate without CRM systems as they can be expensive and it is impossible to justify the initial investment. Salesforce.com saw the opportunity to bring enterprise-level CRM to the masses via SaaS and has subsequently opened up their own platform, Force.com, as part of a PaaS service model.

Applications like CRM SaaS have enabled organisations to abstract themselves away from the infrastructure headaches, and as a result, they can think more about the actual business workflows that take place. Whilst it would seem that SaaS is all about pre-packaged software, the vendors have realised that consumers should be able to configure these offerings so that the application can be suitably customised to integrate with existing systems. This has led to a new interest in the abstraction of business process management (BPM), whereby organisational units create high-level process descriptions of their operations, within software that interfaces the process descriptions to an underlying, transactional code base. This offers substantial benefits including:

- No knowledge of the underlying program code is required.
- Process descriptions are closer to the real operations and are easier to derive and communicate between business users.
- Process optimisation and waste identification is simplified and easier to implement.
- Process commonality is more visible, and therefore, process reuse is more prominent, both internally within an organisation and outside of the normal process boundaries with suppliers.
- Libraries of process descriptions enables the rapid composition of new processes.

From a conceptual stance, Business Process as a Service (BPaaS) might be viewed as a fourth layer, above SaaS, but from an architectural perspective, it is clearly a subset of SaaS as Fig. 1.4 illustrates.

BPaaS creates new opportunities for organisations to exploit the cloud, as the abstraction away from technical and integration issues gives organisations a new way to conduct their business. This topic will be explored more fully in Chap. 10, which is all about *enterprise* cloud computing.

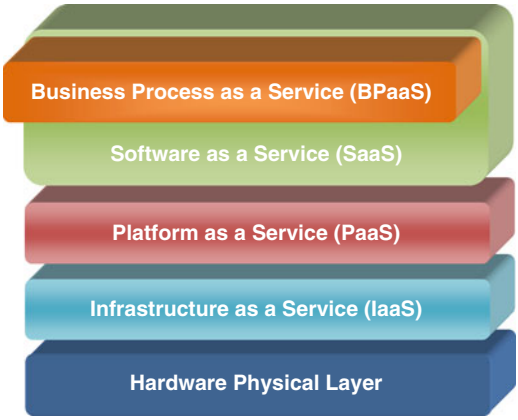

Fig. 1.4 Business process as a service (BPaaS) in the context of the cloud computing stack

1.13 Issues for Cloud Computing

As with any new approach or technology, there are limits by which benefits can be realised, and a new way of working may introduce additional risks. Cloud computing is no different in this respect, particularly as the model is still maturing.

From a consumer's perspective there is a great deal of focus upon security and trust. Many users are ambivalent about where 'their' data is stored, whereas other users (specifically organisations) are more sceptical about delegating the location of the data along with the management processes that go with it. For many smaller organisations, the cloud computing providers will be bringing enterprise-level security to the masses as part of the offering. Most private individuals and small businesses are unaware of the risks of lost data and the adverse impact that it can have upon daily operations. As a consequence, it is likely that they have not put the appropriate security measures in place. In this case, a move towards the cloud can bring real benefits.

However, there may be specific legislation that exists to govern the physical location of data; a multi-tenanted public cloud may place your data in a country that is outside the scope of the jurisdiction that you need to comply with. Additionally, the notion of service as a core component of the cloud leads to new service composition from readily available services. The use of third-party services potentially introduces security and privacy risks, which may therefore require an additional auditing overhead if the services are to be successfully and reliably trusted.

Another concern is that of vendor lock-in. If an organisation utilises IaaS, it may find that the platforms and applications that it builds upon this service cannot be transferred to another cloud computing provider. Similarly, services at PaaS and SaaS can also introduce nonstandard ways of storing and accessing data, making data or application portability problematic.

Quality of service (QoS) is an issue that many organisations already face either as consumers or providers of services. Whilst cloud computing providers offer

measurement and monitoring functions for billing, it might be considered incumbent upon consumers to develop their own monitoring mechanisms to inform any future actions.

Much has been claimed about the potential energy-saving opportunities of organisations moving to the cloud. The ability to pool resources and dynamically manage how these resources are provisioned will of course permit computing resource usage to be more optimised. However, there is an assumption that this occurs at a certain scale, and perhaps less obviously, it is dependent upon the service model required. For instance, an IT department may decide to evaluate the potential of hardware virtualisation as part of a private cloud. The hardware already exists, and the maintenance costs are known. In theory, the more flexible provisioning that cloud architectures offer should release some extra compute resources. In terms of any investment in cooling, for example, then better utilisation of the existing hardware will come cheaper than the purchase of additional air-conditioning units.

Unfortunately, it is only through the provision of compute resources on a massive scale that significant amounts of resource can be redeployed for the benefit of others. The private cloud may be able to scavenge extra processor cycles for heavier computational tasks, but storage management may not be that different from that achieved by a storage area network (SAN) architecture. Thus, significant energy savings can only be realised by using the services of a cloud provider to reduce the presence of physical hardware on premise.

It follows therefore, that it is the massive data centres who offer SaaS that can maximise scalability whilst significantly reducing energy usage. For everyone else, energy reduction might not be a primary motivator for adopting a private cloud architecture.

Of course, as organisations move to the cloud, there is a heightened awareness of measures of availability and the financial impact that a temporary withdrawal of a service might incur. Good practice would suggest that there should be 'no single point of failure', and at first glance a cloud-based system would offer all the resource redundancy that an organisation might want. However, whilst the IaaS, PaaS or SaaS may be built upon a distributed system, the management and governance is based upon one system. If Google or Microsoft went bust, then any reliance upon their comprehensive facilities could be catastrophic. This risk gets greater the higher up the cloud stack that the engagement occurs—if Salesforce.com collapsed, then a great deal of an organisation's business logic would disappear along with the data, all wrapped up in a SaaS application.

Software bugs are a major concern for all software development activity, and many instances of 'undocumented features' occur only when an application is under significant load. In the case of a distributed system, it is not always practical to recreate the open environment conditions, so there remains the potential risk that something catastrophic might occur. Hardware virtualisation can be a way of containing the scope of software bugs, but as many SaaS vendors created their offerings before the widespread use of virtualisation, this form of architectural protection cannot be relied upon. This is clearly a case for open architectural standards for cloud architectures to be established.

As cloud use increases, organisations will place ever-increasing demands that present significant data transfer bottlenecks. Additionally, the distributed architecture of a cloud application may result in a data transfer that would not have occurred had the application been hosted in one physical space. Even though network speeds are getting faster, in some cases the volume of data to be transferred is so large that it is cheaper and quicker to physically transport media between data centres. Of course this only works for data that is not 'on demand' and therefore is relevant when data needs to be exported from one system and imported into another.

With regard to the benefits of scalability, the case for optimising processor cycles across a vast number of units is clear; processors can be utilised to perform a computation and then returned back to a pool to wait for the next job. However, this does not translate as easily to persistent storage, where in general the requirement just continues to increase. Methods for dealing with storage in a dynamic way, that preserve the performance characteristics expected from an application that queries repositories, have yet to be developed and remain a potential issue for cloud computing going forward.

1.14 Summing Up

Cloud computing is a new delivery model for IT that uses established IT resources. The Internet, hardware virtualisation, remote hosting, autonomic computing and resource pooling are all examples of technologies that have existed for some time. But it is how these technologies have been brought together, packaged and delivered as a pay-per-use utility that has established cloud computing as one of the largest disruptive innovations yet in the history of IT. As organisations shift from concentrating on back-office processes, where transactional records are kept and maintained, towards front-end processes where organisations conduct business with customers and suppliers, new business models of value creation are being developed. There is no doubt that the cloud is fuelling this shift.

You've now had a whistle-stop tour of the exciting world of cloud computing. We have covered a lot, and you will probably have some questions that haven't been answered yet. The rest of this book explores a number of important areas in more depth, so that by the end you will not only have a broad understanding of cloud computing, but if you have completed the exercises, you'll be able to implement the technology as well!

1.15 Review Questions

The answers to these questions can be found in the text of this chapter.

1. Explain how energy utility provision has informed the emergence of cloud computing.
2. Briefly discuss the differences between cloud computing service models.

3. Which combination of cloud computing characteristics is the best case for reducing energy consumption?
4. Explain the similarities between grid and cloud computing.
5. Describe the different levels of abstraction that cloud providers can offer.

1.16 Extended Study Activities

These activities require you to research beyond the contents of this book and can be approached individually for the purposes of self-study or used as the basis of group work.

1. You are a member of a team of IT consultants, who specialise in selling IT systems to organisations that have between 100 and 500 staff. Prepare a case for the adoption of cloud computing. Consider the types of IT architecture and systems that might already be in place and whether there are specific business functions made available by cloud computing that an organisation might benefit from.
2. An IT department has decided to investigate the use of cloud computing for application development. What are the issues that they should consider, and how would you advise that they mitigate any risks?

References

Foster, I., Kesselman, C.: The Grid: Blueprint for a New Computing Infrastructure. Morgan Kaufmann Publishers, San Francisco (1999). ISBN 1-55860-475-8

IBM Research: Autonomic Computing. http://www.research.ibm.com/autonomic/ (2012). Last Accessed July 2012

McCarthy, J.: Reminiscences on the History of Time Sharing. Stanford University. http://www-formal.stanford.edu/jmc/history/timesharing/timesharing.html (1983)

Parkhill, D.: The Challenge of the Computer Utility. Addison-Wesley, Reading (1966). ISBN 0-201-05720-4

Business Adoption Models and Legal Aspects of the Cloud

2

What the reader will learn:

- That cloud computing has a number of adoption models
- What is meant by public cloud, and why businesses may choose to adopt this
- What is meant by private cloud, and why businesses may choose to adopt this
- What is meant by hybrid cloud and community cloud, and why businesses may choose to adopt this
- That these new ways of doing business bring with them legal issues that need to be considered as part of any plan to adopt cloud computing

2.1 What Services Are Available?

There are alternative ways a business might adopt cloud computing, and we will be reviewing those approaches in this chapter. As we saw earlier, there are many something-as-a-service options available, and many providers provide all of them, whilst some concentrate on specialist areas like data storage or application platforms.

In a 2011 paper, Li et al. (2010) indicated four general types of service that are currently available from leading cloud providers:

1. Elastic compute clusters which include a set of virtual instances that run a customer's application code.
2. Persistent storage services in which application or other data can be stored in a cluster.
3. Intracloud networks, which connect an application's components.
4. Wide-area networks (WANs) connect the cloud data centres, where the application is hosted, with end hosts on the Internet.

This is a useful categorisation of service types. The other things we will need to consider are metrics. We will need to have some understanding of measures such as performance, cost and availability if we are to have any hope of assessing which

R. Hill et al., *Guide to Cloud Computing: Principles and Practice*, Computer Communications and Networks, DOI 10.1007/978-1-4471-4603-2_2,
© Springer-Verlag London 2013

Table 2.1 A summary of the key differences between public and private cloud models

	Public	Private
Network	Internet	Private network
Server and data centre location	Global	In company
Costing	By usage or free	Internal mechanism, often by capacity and processor
Tenancy	Multiple	Single
Scale orientation	Vertical (i.e. user focused)	Horizontal (i.e. application focused)
Key selection rationale	Cost	Security

provider offers the best solution for any of these services. We will examine these in the 'Which Cloud Model?' section (Sect. 2.6) at the end of this chapter.

As we saw in the last chapter, there are many definitions of cloud. Vaquero et al. (2009) attempted to collate these and come up with a single, all-encompassing definition:

> Clouds are a large pool of easily usable and accessible virtualized resources (such as hardware, development platforms and/or services). These resources can be dynamically reconfigured to adjust to a variable load (scale), allowing also for an optimum resource utilisation. This pool of resources is typically exploited by a pay-per-use model in which guarantees are offered by the infrastructure provider by means of customised SLAs.

We must also not forget that to businesses, it matters not how we define cloud computing but rather it matters whether this form of IT supports their business by reducing costs or adding revenue and profit. You will see more of this discussion in Chap. 8. These elements too are reviewed by cloud type.

The three types of cloud adoption we shall review are public, private and hybrid. As the latter is a combination of the other two, it may be worth starting by examining the key differences between typical public and private clouds (Table 2.1).

2.2 What Is Meant by Public Cloud?

The US National Institute of Standards and Technology (NIST) suggests in a recent draft that the definition of a public cloud is as follows:

> The cloud infrastructure is made available to the general public or a large industry group and is owned by an organisation selling cloud services (Mell and Grance 2011).

The authors of this book believe the general public or a large industry group should be replaced with the general public or organisations as there is no evidence that industry groups need to be of any particular size to adopt cloud computing. The key element here is that services are offered by the resource owner (usually referred to as the service provider) to anyone who wants to make use of that service. The service can be any of IaaS, PaaS, SaaS and DaaS (see the previous chapter for definitions). The service provider may charge, usually on a utility basis, but sometimes on a termly basis, or may give the service for free and earn revenue from other income streams, such as advertising.

Table 2.2 Services and estimated number of users of public clouds

Provider	Estimated users (millions, as of 2010)
Hotmail	330
Yahoo	302
Gmail	193
Others	200

2.2.1 Who Is Using Public Cloud?

The short answer is millions of people!

Mail providers can be evasive about the size of their user-base. Specialist email marketing site http://www.email-marketing-reports.com/ gathered some statistics that give us a feel for the scale of the browser-based email usage. These figures are for the 'big 3', and we can safely assume the other providers (such as Excite, AOL, Rediffmail) will amount to >200 million. The dates for these figures are different but all in or after 2010 as illustrated in Table 2.2.

Remember that our definition of cloud services is that a provider owns the resources required to provide a service (such as email) and rents this service to users on a pay-for-use basis. This means there are already at least a billion users of cloud email services worldwide.

We talk about the phenomenon of social networks in the Social, Economic and Political Aspects chapter. Again, the numbers using these services are over a billion. Many will also use email, but nonetheless, when added to other free, privately focused services like image storage and editing, drop boxes for file sharing and presentation tools like Prezi, there is little doubt that public cloud-based services are here to stay. From the business perspective, however, the view is different. As reported in Computerworld (Mearian 2011), some research by TheInfoPro, a market research firm, which approached 247 Fortune 1000 corporations showed that

> 87% of the respondents indicated that they had no plans to use the public cloud for storage-as-a-service. Only 10% said that they would use it.

We should also bear in mind that this sort of large corporation will have been in business for many years and will have invested heavily in IT infrastructure before the cloud existed. They will already have in place their own processes based on internal systems. Heavy investment in enterprise systems like ERP systems such as SAP or PeopleSoft, and RDBMS like Oracle or DB2, not to mention the investment they will have had to make in the specialist people needed to run these business processes, means there is really very little need for them to look elsewhere for solutions. There are, however, two exceptions to this general rule:

- The eternal search for efficiency and cost reduction
- When an innovative solution is only, or primarily, available from a service provider

We have also seen that security and ownership of the data storage are big issues for all potential cloud users. Even if the search for value leads a corporation to begin to use virtualisation to maximise resource usage, they will often prefer to keep that

transformation in-house to keep a tight control of security. Set in this context, the indications that large corporates are not racing to take up public cloud offerings are not surprising. For such organisations, private or hybrid clouds may be more appealing (see sections below).

For small-to-medium businesses (SMEs), the argument for adopting public cloud appears a little easier to win. Especially at the micro end, with less than ten employees, businesses are very unlikely to be able to attain the sorts of economies of scale that the megacorporations can achieve with their large-scale IT systems. However, if they, in effect, 'club together and share', they can achieve significant economies of scale. The fact that this collaboration is enabled by a for-profit-making service provider is not consequential.

When you add to this the ease of access to on-demand services which are paid for on a utility basis, the argument is even stronger. If some service providers are to be believed SMEs need never employ an IT specialist again since all their business needs can be made available after signing up and simply completing a series of online questions which act as setup wizards for this application or the other.

Of course life is not always that simple. Apart from the ever-present concern about security (see below) being just as relevant to SMEs as to large corporations, there is the age-old debate between whether you should adapt your business processes to allow the use of off-the-shelf software or keep your processes but have to build, or at least tailor, the software. In terms of IT spent, the former is usually seen as the cheaper, but if your processes are part of what gives you competitive advantage, you may be willing to pay for the privilege of using unique software.

Most of these IT strategy-type questions are not new. The control and specialisation which comes from in-house IT solutions has always been balanced against the savings that can come from off-the-shelf solutions. What is new to cloud, however, is that the cash-flow improvement, at least in the short term, can be very significant as costs become revenue rather than capital, spreading the load over years rather than needing high-cost up-front payments.

The other advantage of the move to pay-for-use is the flexibility that it gives a small firm. Should your business suddenly begin to take off and you need more in the way of IT infrastructure and services, you just pay more to your service provider. Conversely, if part of your business fails, you can stop the IT costs immediately, as opposed to being left with expensive servers doing nothing. Both ways seem to significantly reduce the risks involved in an SME opting to use an IT service.

As usual with business decisions, the preferred solution will be a balance of risks and expected benefits. For SMEs, the balance may seem slightly more biased towards the benefits outweighing the risks. However, every company will be different, and contextual issues like company culture, national norms, sector best practice and government and legal guidelines will all play important parts in the decision-making process.

2.2.2 Another Easy Win for SMEs

One area traditionally less well attended to by smaller organisations is disaster recovery (DR). Even backup and recovery strategies may be relatively unsophisticated. An occasional take backup stored in a fireproof safe may well keep a company's

vital data safe, but recovering the data after, for example, a catastrophic server failure, can take days as a new server is purchased, commissioned and brought back to the state of its predecessor.

Major corporations have business continuity plans that look to keep their core operations active with as little as a few minutes between disaster and response. But they have to pay—considerably—for this sort of service. For a multinational bank, for example, this expense is almost a no-brain decision. They can't afford to lose the business that would occur whilst their systems were down.

For an SME, however, a DR plan revolving around a multisite fully mirrored server solution can be seen as a nice-to-have extra as the expense is high and what it buys may never be needed. Cloud provides a small business with an easier, less costly way to run at least two live data centres with automatic failover. This dramatically reduces mean time to recovery (MTTR)—the time between system failures and recovery.

With the cloud, backup need never be to slow tapes. It can be easily automated to happen without human intervention by uploading backup data to a cloud data centre. A centre which will itself have built-in redundancy, meaning you automatically get multiple copies of your valuable data.

2.2.3 Who Is Providing Public Cloud Services?

Those who have seen Larry Ellison's 2009 tirade lampooning cloud computing as nothing other than a hyperbole (see YouTube) may be surprised to see that Oracle now provide pay-for-use services in the cloud (http://cloud.oracle.com).

Other corporates with long track records in the IT arena also now have public cloud offerings and are joined by some newer names. Just as examples, these well-known brands all offer some sort of cloud service now: IBM, AT&T, Fujitsu, Microsoft, HP and Rackspace. And there are many smaller, new market entrants too. Competition is already hot, which is a good indicator that the cloud is well on its way to being accepted by the market.

When we see that these different providers are moving in the same immature market, we should perhaps be a little cautious about predicting the future. Many examples exist of one brand of technology winning out over others and not necessarily because of its excellence. Perhaps the most famous marketing war like this was that between Sony's Betamax and JVC's VHS video formats. The public chose VHS and Betamax died. But there were many people who lost money by investing in Betamax before it declined.

The same thing could happen with cloud. These providers of services do not currently abide by any universally accepted standards. Getting tied into one provider is indeed a risk that needs to be considered. There is a fuller review of interoperability issues in the hybrid section.

2.2.4 Security: The Dreaded 'S' Word

As we will see in the Cloud Security and Governance chapter, privacy and security are big concerns for all potential users of cloud. All the anxieties that may be

expressed are most acute with public cloud, where the profitability of the service provider is the key driver to all technology decisions. As Kaufman (2010) puts it,

> To achieve the gains afforded through virtualisation, such providers are colocating virtual machines (VMs) from disparate organisations on the same physical server. From a profit/loss perspective, this matching seems to provide a win-win scenario for both the user and service provider. However, this operational profile introduces a new era of security concerns.

As we have said elsewhere, there isn't much new, in terms of technology, with cloud. There is no real reason why cloud platforms should not be as secure as a traditional platform. Indeed, in some cases, it may be more secure. For example, a server in a locked room may not be as well protected as the Google data centres, as described in this YouTube clip:

http://www.youtube.com/watch?v=1SCZzgfdTBo

In these places, biometrics, multi-checkin and log-in make access to hardware from outsiders virtually impossible—probably far more secure than an average SME's premises.

Of course, one of the aspects about public cloud is that services are accessed through the Internet: an Internet that is available worldwide to both friend and foe. This shared remote access model can potentially allow cyberattacks. All this means that security can be an issue with cloud, but there are issues with current IT infrastructures too.

The perception of insecurity is, however, probably the biggest barrier to cloud adoption. For the non-technically minded amongst business decision-makers, it is not difficult to understand why they may be wary about parcelling up their valuable data and giving it to another company to look after, instead of having it sit on a server behind a locked door on their site. These doubts are compounded when you explain that their data will be multi-tenanting, sharing the same physical resources, perhaps, as their biggest competitor. How could that be seen as a sensible move?

Nor is it just data that can be worrisome. Even IT-literate decision-makers are likely to have grown up in an era when modems went down, when Internet connections broke and when speed of transmission plummeted. How can it be sensible to replace your reliably performing single-purpose system connected to a few clients in a small LAN, all under the control of your network team, with a barely understood worldwide web of entangled connections? Why move ERP from in-house to in-Indonesia or some other foreign domain?

It is not this book's place to counter these concerns. The major service providers will fight that battle, but we do need to be aware that security can be a human problem, rather than a technical one.

2.3 What Is Meant by Private Cloud?

The technology stack need be no different to that used by service providers in public cloud solutions. The US National Institute of Standards and Technology (NIST) suggests in a recent draft that the definition of private cloud is as follows:

> The cloud infrastructure is operated solely for an organisation. It may be managed by the organisation or a third party and may exist on premise or off premise.

The key element here is that the resource owner (known as the service provider in public cloud) is the organisation that is using the services. The service can be any of IaaS, PaaS, SaaS and DaaS (see earlier chapters for definitions), and there may be internal charging mechanisms for these services, but they are not normally made available to anyone outside of the organisation and hidden behind a firewall.

2.3.1 Who Is Using Private Cloud?

Because of the expense involved in creating multi-server operations, early adopters tend to be large organisations with existing infrastructures that lend themselves to the adoption of a cloud platform to increase server efficiency (and thus reduce costs) and allow broader availability to systems within the organisation. We must also remember that organisations have been using some of the building blocks, such as virtualisation and SaaS, for years without calling it cloud.

There is an argument that private cloud is not really that different to the ways large organisations typically manage their infrastructures. Stand far enough away and the technology of a large server farm making good use of virtualisation looks very similar to a cloud. To make matters worse, the organisation doesn't even get the advantages of flexibility, which come from sharing resources, nor do they benefit from the move to revenue costing that is also one of cloud's oft-trumpeted advantages.

Whether or not a move to a private cloud will be beneficial to an organisation depends upon many things, but their existing infrastructure is one of the key ones. A recent big spend in modernising the company data centre can be an indicator that investing in cloud is not an immediate need. If it is time to upgrade anyway, then perhaps internal cloud is a solution worth reviewing.

Especially in the current economic conditions, companies are looking at all their costs to see if they can run more efficiently. IT is no different to any other part of the business in this. Most big organisations depend upon a set of core IT processes. The question being asked is 'are we paying too much for this service?' and that question plays into the hands of those arguing the benefits of cloud computing.

Gartner (2010) suggests that

> … cloud computing has become more material, because the challenges inherent in managing technology based on the principles of previous eras — complex, custom, expensive solutions managed by large in-house IT teams — have become greater, and the benefits of cloud computing in addressing these challenges have matured to become more appropriate and attractive to all types of enterprises.

The question on the lips of many larger organisations' CIOs will not be private versus public but rather legacy versus private. The ability of a cloud infrastructure to flexibly move computing resources to deal with spikes in workload means that cloud-based data centres can run much more efficiently than existing ones, and that may be the biggest single factor in the decision.

For organisations who have taken the decision that cloud will be their preferred technology solution, the question of public versus private is likely to force them to think about the value of security to their business. Private allows, or at least seems to allow, organisations to have greater control over their data. There are, however, many more barriers to private since in-house expertise in virtualisation and operations automation may not currently exist and will be expensive to acquire. Moreover, a move to public cloud can happen much more quickly and allows for maximum flexibility in resource management. The ultimate question, therefore, is likely to be how much are we willing to spend to maintain control over our data?

A whole later chapter is reserved for further investigation into enterprise cloud, and many of the issues which surround the process of adopting a private cloud in a large organisation are covered there.

2.3.2 Who Is Supplying Private Cloud?

Most of the big players are now fully committed to selling products or services badged as cloud. Even Oracle, once more famous for laughing at cloud, sells cloud-related services and products, mostly private cloud solutions. They say

> Cloud computing promises to speed application deployment, increase innovation, and lower costs, all while increasing business agility. It also can transform the way we design, build, and deliver applications....
> (http://www.oracle.com/webapps/dialogue/ns/dlgwelcome.jsp?p_ext=Y&p_dlg_id=9270949&src=7054580&Act=13&sckw=WWMK10058758MPP002.GCM.9322)

IBM has been in cloud from very early days. Lotus Notes has now become iNotes, and one prong of the IBM cloud marketing campaigns is clearly aimed at public, with the catchy strapline of

> Install nothing. Access everything.

But IBM clearly recognises the need for private cloud too. They have a suite of underpinning technologies they call SmartCloud Foundations which they describe as

> an integrated set of technologies for enabling private and hybrid clouds, and the virtualisation, automation and management of service delivery. SmartCloud Foundation capabilities allow organisations to easily build and rapidly scale private cloud environments.
> (http://www.ibm.com/cloud-computing/us/en/)

HP is a big player too, playing heavily on the reputation for cloud to be rapid and flexible; they can deliver private cloud computing services within 30 days (http://www.hp.com/hpinfo/newsroom/press/2010/100830a.html).

On their website, their senior vice president and general manager, Technology Services, HP, uses the concept of an ‘internal provider’:

> To better serve the needs of their enterprises, clients are asking us to help them become internal service providers with the ability to deliver applications through a highly flexible private cloud environment.

Citrix too has been in the market since it really started. Their solutions also play on the speed of change possible from cloud:

> With CloudStack, customers can quickly and easily build cloud services within their existing infrastructure and start realizing the benefits of this transformative service delivery model within minutes—without the overhead of integration, professional services and complex deployment schedules.
>
> (http://www.citrix.com/English/ps2/products/product.asp?contentID=2314749)

An interesting development with Citrix is their CloudBridge technology which tackles the perceived security issues in public cloud head-on and seeks to help create secure hybrid solutions:

> Citrix CloudBridge lowers the risk and reduces the effort and cost for enterprises to move production workloads to the cloud by …. making the cloud provider network look like a natural extension of the enterprise datacenter network.
>
> (http://www.citrix.com/site/resources/dynamic/salesdocs/Citrix_NetScaler_Cloud_Bridge.pdf)

As well as suppliers of hardware and software, consultancies too are very much in the market for helping customers migrate to a cloud solution. And it isn't just Western companies who are pushing cloud. TCS and Infosys in India, for example, are major global players.

Simply type private cloud supplier in a Google search, and (at the time of writing) 95 million hits are reported. There can be no doubt that the cloud market is well and truly active!

2.4 What Is Meant by Hybrid Cloud?

NIST definition:

> Hybrid cloud. The cloud infrastructure is a composition of two or more clouds (private, community, or public) that remain unique entities but are bound together by standardized or proprietary technology that enables data and application portability (e.g., cloud bursting for load balancing between clouds).

The key aspect is that hybrid includes some mix of public and private cloud in a non-specified ratio.

2.4.1 Who Is Using Hybrid Cloud?

If an organisation has a steady and quantifiable use of IT resources, they are able to adopt private cloud, gaining the benefits of efficiency and availability, without missing the other strength of cloud—flexible scalability.

If, on the other hand, like many organisations, they have spikes of activity, planned or not, then public cloud's ability to offer unlimited and immediate scalability on an occasional basis may well appeal. Building your systems to cope with

standard workloads in-house and extend outwards when required should allow for the best of both worlds. Sensitive systems can be kept entirely in-house if required.

Some e-commerce organisations can adopt a hybrid approach to help with the activity associated with the front-end during peak shopping periods whilst maintaining secure back-end services in their own private cloud. This prevents them having to invest in many servers which may be idle for long periods just to cope with occasional high loads.

The other likely driver towards a hybrid approach is the organisation's existing infrastructure and their IT strategy. Hybrid may well be an interim approach which means that wholesale in-house architectural changes do not need to happen immediately as some changes are contracted out to service providers and some existing systems continue to function. Interoperability between these different systems here is a key issue (see below).

Another way that hybrid is likely to happen is by accident. An organisation with its own private cloud platform for its main systems may, for example, decide that Google's Gmail email solution is the right one for their organisation. The security risks with noncritical systems like email will seem relatively minor, and the cost-effectiveness of such a solution may attract many organisations. Part of their IT stack then becomes private, part public—de facto a hybrid cloud solution.

2.4.2 What Are the Issues with Hybrid Cloud?

Whilst suppliers, such as Citrix and their CloudBridge, will be keen to suggest that hybrid offers the best of both private and public worlds, it is also arguable that it is the worst of both. After all, as we saw in the private section above, one of the biggest drivers for private solutions is the ability to control your own, independent data centre for security reasons. Claybrook (2011) suggests

> The challenges of building a bridge between private and public clouds are real. (http://www.computerworld.com/s/article/9217158/Cloud_interoperability_Problems_and_best_practices)

The report goes on to quote Joe Skorupa, a Gartner vice president, as saying that

> … users and cloud vendors are in very different places on this issue [interoperability], and true cloud interoperability will likely not occur for some time -- if ever. Standards are nascent and will take years to fully develop.

The lack of standards is indeed likely to be a major stumbling block when it comes to trying to pass data, which will usually be encrypted, between different systems in a hybrid cloud solution. It is not unusual in IT for technology to get so far ahead of standards. And in the absence of standards, there is little reason for the various providers to ensure ease of communications between themselves and other providers. Indeed, the cynical amongst us may even think that these different approaches can help tie in the customer to a provider.

The two key proprietary virtualisation technologies (VMWare and Hyper-V) will be trying to keep their own customers whilst also fighting off open-source alternatives in the PaaS area. As trust is one of the likely decision factors for cloud platform providers' customers, some form of industry-wide standard is being actively sought. Unfortunately, however, there are several agencies keen to seek to take the lead in this area. At the time of writing, these included:

- IEEE, self styled as 'the world's largest professional association advancing technology for humanity'
- Open Grid Forum
- Cloud Security Alliance
- NIST

All these agencies are themselves liable to lobbying from the industry. This lobbying is generally for financial reasons, but it is also true that individual providers naturally believe their particular solutions are the best! It is unlikely that a truly global and agreed standard will happen for a few years yet, so interoperability is likely to remain one of the biggest barriers to hybrid adoption.

2.5 What Is Meant by Community Cloud?

NIST definition:

> The cloud infrastructure is shared by several organisations and supports a specific community that has shared concerns (e.g., mission, security requirements, policy, and compliance considerations). It may be managed by the organisations or a third party and may exist on premise or off premise.

The key aspect here is that of inter-organisational collaboration. Community cloud is just like a private cloud except that several organisations share the responsibility for resourcing the cloud, instead of just one.

2.5.1 Who Is Using Community Cloud?

Trust between companies operating in a competitive marketplace is not a usual phenomenon, and so community is not a realistic option for them. However, organisations which are about care and support have naturally tended to help each other in the past. Charitable organisations, for example, have been coming together to share all sorts of resources, including IT.

One example is the International HIV/AIDS Alliance which is a partnership for '… everyone who works with and for NGOs and CBOs and is involved in community and health system strengthening worldwide'.

Whilst the political advantages which come from small charities coming together as a single pressure group are their reason d'être, the support provided by IT across the partnership can also be important. Working with Cisco, the alliance has implemented online collaboration and SaaS platform:

Fig. 2.1 AIDS Alliance website home page (last accessed 22 May 2012)

http://www.aidsalliance.org/includes/Document/Uploaded/IHAA_CISCO_D1.pdf

The vision expressed by Sam McPherson, associate director, International HIV/AIDS Alliance, is

> We want to exploit the technology available to us and truly become a collaborative organisation. By using the full complement of WebEx solutions, we hope to move closer toward our vision of a world in which people do not die of AIDS.

One major problem is that not all third sector organisations are as forward thinking as the International HIV/AIDS Alliance (Fig. 2.1). Many charitable organisations are small and not cash rich and are therefore afraid of the costs associated with IT systems (Maison 2011). In a recent survey of nearly 160 charities, *the Guardian* found

> Eight of 10 people said that technology could help build the 'big society'. Yet only one in three have the time or confidence to try out new tools like cloud computing.
> http://www.guardian.co.uk/voluntary-sector-network/2011/jun/01/charities-save-money-cloud

Other first movers in the area of community cloud are governmental organisations. Sometimes the key driver here is the need, traditionally difficult to address with different organisations with disparate IT systems, to share information. In the UK, for example, the police service is separated into constabularies, and they have their own budgets and have met their information system needs with different solutions. This can make sharing information about a suspect difficult when they cross boundaries between constabularies. The matter gets yet more complicated should the suspect be apprehended and taken to court, as the court systems will also be different, not to mention prison systems should they be found guilty.

In the USA, firms like IBM have been quick to spot how they can offer a service to governmental organisations. In a recent press release, they say

> IBM has launched a new Federal Community Cloud specifically designed to help federal government organisations respond to technology requirements more quickly. The secure, private cloud environment is part of IBM's established and dedicated Federal Data Centers (FDC) that provide secure and comprehensive certified computing capabilities to federal government clients.

In the UK there is G-Cloud. This is a government-funded initiative to gain the benefits that cloud can give whilst attempting to save the public purse £200m/annum by 2014/2015: http://gcloud.civilservice.gov.uk/ The G-Cloud program is a cross-government initiative; collaboration across departments, and throughout the public sector, being encouraged and enabled by cloud.

Reported in *the Guardian* in January 2012 (Best 2012), Liam Maxwell, the UK Cabinet Office's director of ICT futures, foresees

> "In two or three years' time what we now call IT, the delivery of those disaggregated services like hosting, networking, end user devices, support, all of those, will become core commodity services and will be bought 'like stationery'".

2.6 Which Cloud Model?

Of course, the answer to the question 'which type of cloud' may well be none. Richard Stallman, founder of GNU, argued that cloud was a trap in an article in *the Guardian* (Johnson 2008). He argued

> 'One reason you should not use web applications to do your computing is that you lose control', he said. '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 defenceless. You're putty in the hands of whoever developed that software.'

Before 2010, there were many such warning sirens. Larry Ellison, Oracle's CEO and co-founder, is also famously quoted as saying that cloud is 'nonsense'. And yet, now, Oracle is a leading player in cloud services to corporates.

If we examine the sales statistics from the cloud service providers, there can be little doubt that many CIOs, IT Managers and IT Consultants are now seriously

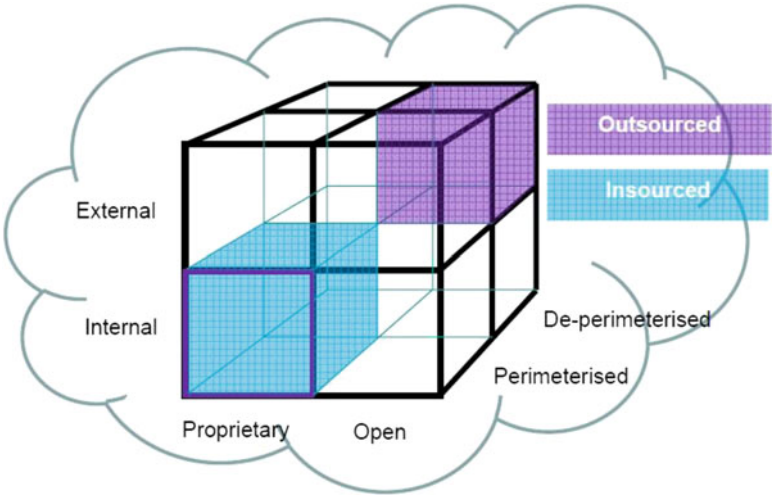

Fig. 2.2 Jericho Cloud Cube Model (2010)

considering cloud platforms as one of their options when looking at how to deliver their IT strategies. So, how do they decide which cloud adoption model to use?

We have identified already that cloud security is seen as a major concern by many organisations. At least whilst the platform is still quite new, many will adopt a 'wait and see' approach—especially if their existing infrastructure is adequate. Some, seeking to gain some advantage from early adoption, may see the advantages of cloud but still want to be cautious about how they look after their data and internal systems. For them, probably starting with pilot projects to test the water, private cloud may well seem more attractive.

The Jericho Forum proposed a framework Fig. 2.2 (Opengroup 2010) which is intended to help organisations find the most appropriate cloud 'formations' for their own particular business need. 'Formations' is a nice way of describing the many alternative solutions available in a mix-and-match environment. Every organisation is likely to be different.

The Forum describes itself as '…an international IT security thought-leadership association dedicated to advancing secure business in a global open-network environment', so it is not surprising to see that security figures highly in their proposed decision-making process.

The cube usefully expresses the considerations that need to be made when deciding which approach to take. The dimensions are described below.

- Internal/external here is the same as private/public clouds.
- Proprietary/open is, as with other software, whether or not the software or platform is open source or not. Also important in the cloud is how open the data standards adopted by a supplier are. Really we are talking about how much tie-in the supplier has over the customer, and whether that is an issue of concern or not.

- Perimeterised/de-perimeterised is about where the IT services exist. If a company keeps all its data behind a firewall within its own private network, for example, we would call that perimeterised. The Jericho paper interestingly refers to this as a mindset. This is very import as an organisation's culture will impact heavily upon their willingness to expose, or not, their systems to external access.
- Insourced/outsourced is about who does the work in the cloud. Entirely insourced means that the organisation employs the people directly. The use of contractor or specialist consultants allows for a control to be maintained within the organisation whilst certain specialist skills are outsourced, often temporarily whilst in-house staff gain the skills themselves.

This cube is an excellent start, but other important factors in the decision about which cloud adoption model to select are not covered but need reviewing.

2.6.1 Internal Factors

1. *Existing infrastructure and IT portfolio*. 'If it ain't broke, don't fix it.' Cloud has some potential benefits, but as with all new technologies, it has risks too. If the organisation's IT is delivering what it should, as well as it should, then there is probably nothing for a CIO to do other than keep their eye on the cloud space.
2. *Capability*. Rightly or wrongly, CIOs in organisations with a long history of managing their own IT systems with their own employees may feel that some of the marketing hype about the cloud's approachability and ease of use does not apply to them. Their CEOs and CFOs may actually disagree if there is board level dissatisfaction with existing internally supplied services.
   Start-ups, on the other hand, will have none of these prejudices. The ability to implement sophisticated enterprise-style systems with no in-house expertise may well be seen as the single biggest reason for opting into public cloud services.
3. *Emphasis on costs*. It may seem obvious that companies will always look to run as efficiently as possible, but in a time of economic hardship such as most of the world is enduring as we write, it is the case that efficiencies are more aggressively sought. Being new, we have no real evidence as to whether cloud is truly a cheaper alternative long term, but we do know that moving away from big capital expenditure IT projects towards pay-for-use will move costs away from a company's fixed assets and into revenue costs, spreading the cash flow over many years as it does so. This drive to efficiency can point towards public cloud where the nature of the shared capacity leads to significantly more savings than would private cloud.
4. *Performance and scalability*. Again, there are not enough studies carried out to suggest how cloud performs in comparison to in-house client/server technology. The most obvious point is that a reputable cloud provider will always be running on high-performance equipment in order to enable them to support many users. However, how big a 'slice' of that platform a customer gets is variable.

The other aspect of this comparison is that a recently upgraded internal infrastructure will perform better than an ageing one and will therefore be less likely to be outshone by cloud. If performance is paramount to a business, the likelihood is that they would adopt private cloud, where they can manage the performance themselves and ensure that nothing can cause degradation.

It is probably true that a need for scalability is a significant driver towards adopting cloud. If an organisation understands its business well and it is relatively stable, it can plan what capacity is required and purchase as and when required. Many organisations, however, go through unexpected sharp up- and downturns in their OLTP traffic in step with the business performance. Not having to purchase extra capacity 'just in case' in such circumstances can make public cloud more appealing.

2.6.2 External Factors

1. *Publicly available bandwidth.* Cloud computing requires reliable, high-performance access to the Internet to work effectively. In some luckier Western countries, this is not a problem with almost country-wide broadband coverage. In other nations, however, the Internet is only available through mobile telephones or private networks. Organisations which have their own private networks in these countries will be able to decide on a cloud adoption model as described elsewhere, but those with limited or poorly performing access may be constrained to only using public cloud SaaS options, such as email and document sharing.
2. *The competition*. It is the nature of a competitive market that organisations will monitor what each other is doing. They need to ensure that no-one steals a march in adopting some new technology that may give competitive advantage. Sustainable competitive advantage in the IT arena is an impossible dream as every advance can be replicated by the competition given time. However, to not seek at least temporary advantage is, in actual fact, to allow oneself to go backwards, as everyone else in the market will be looking for the next new advance. Of course, caution is needed. Just blindly adopting an approach because a competitor has it is a recipe for disaster. However, if your major competitor suddenly starts using public cloud for some of their IT needs, it may well be the case that you should at least review the potential advantages to your organisation.
3. *Suppliers' and purchasers' expectations*. The balance of power between your organisation and its customers on the one hand and its suppliers on the other will impact your decision-making. When electronic data interchange (EDI) came to the fore in the 1980s, it was seen by adopters as a cost-reducing technology which would speed the order-to-delivery process. Typically the early adopters were large companies in particular markets. The motor trade was one such market, and early adopters were the big automobile manufacturers. In order to ensure that their suppliers would adopt this new technology, some manufacturers began to dictate that all their orders for parts would be delivered electronically. In a market where the customer was king, this meant that part manufacturers had to adopt EDI practices or else face bankruptcy.

Similar pressures will begin to bear on companies dealing with organisations which are using the public cloud to manage all or part of their own supply chain. In those circumstances, the decision to use public cloud might be made for you by default.

There are many other business reasons for and against which model to adopt, and we investigate some more detailed investment appraisal approaches in Chap. 8.

2.7 Legal Aspects of Cloud Computing

The law about cloud computing, because of the relative newness of the concept, is largely uncertain, and, as is often the case in a rapidly moving field like IT, the lawyers and legislators are having difficulty keeping up with the changes. However, there are some elements that are clear.

2.7.1 A Worldwide Issue

In March 2010, in the USA, the ITIF president Robert D. Atkinson said, 'There is no way a law enacted at the dawn of the digital age can guide law enforcement officials and protect privacy rights in an age of cloud computing and the wireless Internet, and as billions of electronic exchanges occur every hour' (ITIF Press Release 2010).

One reason that cloud is going to be problematic to law makers is borne of its very essence—global, shared, distributed and replicated data which may reside anywhere in the world. Several of the leading players in the spread of cloud have formed a pressure group in the USA to try and push legislators to recognise that current legal frameworks are not cloud friendly. They are called the Digital Due Process (DDP) group and their aim is to

> …simplify, clarify, and unify the ECPA [Electronic Communications Privacy Act] standards, providing stronger privacy protections for communications and associated data in response to changes in technology and new services and usage patterns, while preserving the legal tools necessary for government agencies to enforce the laws, respond to emergency circumstances and protect the public. (DDP Website 2011)

Naturally, when organisations like Amazon, Facebook, Google and IBM (all fierce competitors in the cloud market) can agree to come together to lobby government, we can see that there is a lot of commercial interest in getting the legislation changed. We are, however, still at the stage where we will have to wait and see what the law makers do in response. This all sounds very American, but we should acknowledge that in terms of cloud, where the USA goes, so, often, follows Europe and the rest of the world. China is a noticeable exception, having a massive internal market for cloud technology, but with its own particular legal frameworks which do include filtering out certain cloud content before it crosses into China.

Because of the inherently international nature of cloud computing, commentators are suggesting that the world needs international treaties to allow for the free

movement of information across borders, in the same way agreements protecting commercial bank transfers between organisations in different nations allows the globalisation of trade in goods.

Policing, too, is difficult when the cybercrime is so international in nature. There are international agreements already in place. The Budapest Convention, for example, allows police to access servers in other countries. However, cybercriminals can move data and applications from one server to another, across national boundaries, very easily and quickly, which makes the work of the police extremely difficult.

This uncertainty is doubtless adding to the perceived level of risk for organisations thinking of using the cloud. Compared to current service-focused IT provision, they see cloud as less transparent and may legitimately feel less protected by the law. Particularly when organisations are talking about handing over vital or sensitive information to service providers, their concerns are understandable. Moreover, even if the service providers themselves do act as their customers wish, there have been cases where governments and their legal systems have forced service providers to hand over data stored in the cloud.

When this happens, there may well be no impetus for the service provider to fight any subpoena as the information is not theirs and they can blame the state for them having to pass the data over. The legal position is made even trickier by the fact that the law that exists, created in a different era, states that data handed over to a third party in the normal course of business can be subpoenaed without notice. What customers are doing with cloud service providers is passing data on to third parties but for storage, not for sharing, as was the norm when the laws were first couched (Gruenspecht 2010).

2.7.2 The Current Legal Framework for Cloud

The uncertainties outlined in the above section may be one reason for an organisation being wary of investing in the cloud. However, elsewhere in this book, we have seen its many advantages, and as with all business decisions, organisations will just weigh benefits against risk. Other players, such as governmental institutions, will also provide input to the decision-making. In the EU, for example, the Commission President indicated that he foresaw that digital commerce would be a significant area of growth for Europe:

> Half of European productivity growth over the last 15 years was driven by information and communication technologies. This trend is set to intensify. Our European Digital Agenda will deliver a single digital market worth 4% of EU GDP by 2020 (Barroso 2010).

Many companies have already committed to cloud. They will therefore need to work within the existing legal framework. Uncertainty is not an excuse to ignore the laws that do exist.

Remember that one of the building blocks of cloud, particularly the public aspects thereof, is the idea of pooling resources and charging them out on a pay-for-use basis. The service provider will typically offer certain guaranteed services, and the service contract will usually include service-level agreements (SLAs). The guarantees are usually expressed in measurable terms, some examples of which include:

- Availability of the service
- Minimum performance benchmarks
- Minimum help-desk response time

These SLAs are part of normal contract law. The jurisdiction in which any legal disputes will be settled is often stipulated within the SLA itself but if it isn't determining the appropriate jurisdiction can be a lengthy (and expensive) precursor to any actual legal action. The question, in short, is the following: Which national, or subnational, laws apply? Those of the providing company's head office? Those of the customer? Those of the location of the data centre? The safest advice to give, therefore, is to ensure that jurisdiction is explicitly agreed in the SLA.

2.7.3 Privacy and Security

As we cover in the Security and Governance chapter of this book (Chap. 10), there is much for potential cloud adopters to worry about in terms of privacy and security. This section only covers the legal aspect of these concerns.

Until legislation specific to cloud computing is forthcoming, both service providers and their customers need to rely heavily on their SLAs to effectively deal with security risks, a process that requires an element of trust from the customer perspective. Further to the comments above about the EU putting cloud high on their economic policy agenda, the EU has created a body called the European Network and Information Security Agency (ENISA) to review and respond to cybersecurity issues within the European Union. Its website says it is

> … the 'pace-setter' for Information Security in Europe, and a centre of expertise. The objective is to make ENISA's web site the European 'hub' for exchange of information, best practices and knowledge in the field of Information Security.

ENISA's cloud computing risk assessment report (http://www.enisa.europa.eu/activities/risk-management/files/deliverables/cloud-computing-risk-assessment) states 'loss of governance' as one of the biggest single risks for cloud adopters. The customer passes responsibility for security to the service provider, who may not provide adequate guarantees in their SLAs. Any adopter therefore needs to carry out a risk assessment, perhaps as discussed in the ENISA report, and must ensure that their privacy protection is built into the SLA.

Suppliers of cloud infrastructure and services are not going to allow a perceived lack of security to prevent them from maximising profits. If you Google 'cloud security IBM' and then repeat for the major cloud players, you will see many pages on each site dedicated to explaining the supplier's security. And current security specialists, too, have noticed how cloud is becoming important. McAfee recently released its Cloud Security Platform, for example, and Symantec's have their Symantec.Cloud.

But these are still all sales pitches, and some caution needs to be taken. With the best will in the world businesses do not, and should not, blindly believe suppliers' claims. Again, until legislation catches up, it is the customers' task to ensure that they have contracts which ensure their data is secure and that services are delivered as promised.

2.8 Summary

In this chapter, we explored the different methods by which cloud computing can be adopted by organisations and by individuals. The adoption types we examined were public, private, hybrid and community. These terms will be used throughout this book and are in wide usage in the computing arena and have become the de facto way of describing the differing approaches. The ways that these are implemented technically are explored in the next part of the book, whilst the business aspects are explored in Part III.

We also analysed the way that these adoption types may be used by different types of business, from small to enterprise sized. We have a chapter in Part III which discusses large-scale enterprise cloud in more detail.

One of the major difficulties for organisations trying to decide whether to adopt cloud computing is which model to adopt. We began to explore tools to assist in analysis of the major factors and looked at the Jericho Cloud Cube Model. A more detailed review of the financial and investment appraisals issues is to be found in Chap. 8.

2.9 Review Questions

The answers to these questions can be found in the text of this chapter.

1. List the types of service that are available from cloud providers today, being clear that you understand the differences between them.
2. How might cloud be an easy solution for smaller businesses looking for business continuity and disaster recovery?
3. What is meant by hybrid cloud?
4. Is a community cloud a public or private cloud solution? Or both? Or is it something else?
5. Why is the policing of cloud seen as problematic for many law makers?

2.10 Extended Study Activities

These activities require you to research beyond the contents of the book and can be tackled individually or as a discussion group.

2.10.1 Discussion Topic 1

What factors are suitable for inclusion in an SLA between cloud provider and customer? You should not only review the factors themselves but also decide on their relative importance and how they might be measured and monitored. You should also consider what the likely impact of requiring extremely demanding levels would be on cost.

We saw that SLAs are key for organisations in terms of ensuring satisfactory levels of service from providers. Some of the more obvious factors are around performance and availability. Five 9 s are industry-speak for as available as possible and mean that a system is up and running 99.999% of the time. However, availability levels set so high are extremely expensive to enable, as the provider will need many layers of redundancy built into their offering.

Measurement too can be a problem. The organisation may have in mind that performance can be measured in terms of user-click-to-returned dataset times. But for cloud applications, the timings can be out of the provider's hands since much will depend upon local Internet speeds and connections.

2.10.2 Discussion Topic 2

Many commentators see hybrid as the likely model for cloud adoption in the long term, allowing companies to use the best of both public and private platforms. In an era when many applications are built with data sharing built in, you should explore the significant challenges that will be faced by organisations with mixed public–private application portfolios.

When attempting this question, you should look to see what standards are in place for cloud computing. If you advise your organisation to use Salesforce CRM, for example, what pressure does that put on other organisational systems in terms of preventing needless data duplication? Is there a threat that cloud could actually result in more siloed data and less sharing?

References

Barroso, J.M.D.: State of the Union 2010 Strasbourg, 7 Sept 2010. http://europa.eu/rapid/pressReleasesAction.do?reference=SPEECH/10/411 (2010). Last accessed 22 May 2012

Best, J.: G-Cloud will lead to shorter contracts and IT 'bought like stationery'. Guardian Professional, Thursday 26 Jan 2012. http://www.guardian.co.uk/government-computing-network/2012/jan/26/gcloud-contracts-liam-maxwell-procurement (2012)

Claybrook, W.: Cloud interoperability: problems and best practices. ComputerWorld, June 2011. http://www.computerworld.com/s/article/9217158/Cloud_interoperability_Problems_and_best_practices (2011)

DDP Website.: http://digitaldueprocess.orgspecific page; http://digitaldueprocess.org/index.cfm?objectid=99629E40-2551-11DF-8E02000C296BA163 (2011)

Gartner, Inc.: Gartner Says Worldwide Cloud Services Market to Surpass $68 Billion in 2010. Gartner press release, Stamford, 22 June, 2010. http://www.gartner.com/it/page.jsp?id=1389313 (2010)

Gruenspecht, J.: "Reasonable" grand jury subpoenas: asking for information in the age of big data. Harv. J. Law Technol. **24**(2), 543–562 (2010). http://jolt.law.harvard.edu/articles/pdf/v24/24HarvJLTech543.pdf

ITIF Press Release: ITIF Calls for Updates to Privacy Laws, 30 Mar, 2010. http://www.itif.org/pressrelease/itif-calls-updates-privacy-laws (2010). Last accessed 22 May 2012

Johnson, R.: Cloud computing Is a trap, warns GNU founder Richard Stallman, guardian.co.uk, Monday 29 Sept 2008. http://www.guardian.co.uk/technology/2008/sep/29/cloud.computing.richard.stallman (2008)

Kaufman, L.M.: Can public-cloud security meet its unique challenges? IEEE J. Security Priv. **8**(4), 55–57 (2010). ISSN: 1540–7993

Li, A., Yang, X., Kandula, S., Zhang, M.: Comparing public cloud providers. IEEE Internet Comput. **15**(2), 50–53 (2010)

Maison, A.: How charities could save money by getting on 'the cloud'. Guardian Professional, Wednesday 1 June 2011. http://www.guardian.co.uk/voluntary-sector-network/2011/jun/01/charities-save-money-cloud (2011). Last accessed 22 May 2012

Mearian, L.: Fortune 1000 firms shun public cloud storage. ComputerWorld, May 2011. http://www.computerworld.com/s/article/356680/Survey_Big_Firms_Shunning_Public_Cloud_Storage (2011). Last accessed 22 May 2012

Mell, P., Grance, T.: The NIST Definition of Cloud Computing, NIST Special Publication 800–145 (Draft). Recommendations of the National Institute of Standards and Technology. http://csrc.nist.gov/publications/nistpubs/800-145/SP800-145.pdf (2011). Last accessed 22 May 2012

Opengroup: Cloud Cube Model - Selecting Cloud Formations for Secure Collaboration April 2009, The Jericho Forum, a Forum of The Open Group Available online from: https://collaboration.opengroup.org/jericho/index.htm (2010). Last accessed 22 May 2012

Vaquero, L.M., Rodero-Merino, L., Caceres, J.: A break in the clouds: towards a cloud definition. ACM Comput. Commun. Rev. **39**(1), 50–55 (2009). doi:10.1145/1496091.1496100. ISSN:0146–4833