Chapter 20
Cloud Economics: Principles, Costs, and Benefits

Asoke K. Talukder, Lawrence Zimmerman, and Prahalad H.A

Abstract Cloud computing is an important next step in the trend toward inexpensive and universal access to information and sophisticated computing resources that help close the digital divide between the computer haves and have-nots. In cloud computing, the end-users can access fully functional software and services online at little or no cost using inexpensive computers or mobile communication devices that connect them via the Internet. Innovative service providers no longer need to own and maintain development or production infrastructures and can automatically scale their production operations to meet growing demand much more easily and economically than possible with internal data centers, traditional hosting, or managed services arrangements. The cloud's inherent ability to dynamically scale up or scale down the infrastructure commitment as demand changes on a pay-as-you-go basis has a positive impact on the service provider's overhead costs, energy costs, and in reducing its carbon footprint.

Cloud economics as presented in this chapter refers to the economic forces, business drivers, and structural issues affecting the broad costs and benefits of adopting the cloud technologies or the creation of private or public utility clouds. Here, cloud economics also deal with the economy inside the cloud, which includes monetization, charging, billing, and taxation of products and services inside the cloud.

20.1 Cloud Computing Reference Model

The cloud can be divided into three major verticals, namely, the *cloud user*, the *cloud vendor*, and the *original cloud provider* (OCP) as shown in Fig. 20.1. The cloud vendor is an organization that has a local tax registration and offers the cloud services to the *cloud user* with guaranteed *quality of experience* (QoE) and *quality of service* (QoS)

A.K. Talukder (✉)
Geschickten Solutions, Bangalore, India
e-mail: asoke.talukder@gmail.com

N. Antonopoulos and L. Gillam (eds.), *Cloud Computing: Principles, Systems and Applications*, Computer Communications and Networks,
DOI 10.1007/978-1-84996-241-4_20,

within the framework of a *service level agreement* (SLA). A cloud vendor can be a compute, storage and application brokerage and clearing house that provides prenegotiated access to the cloud services such as *infrastructure as a service* (IaaS) provider, *platform as a service* (PaaS) provider, and *software as a service* (SaaS) provider [1]. The IaaS service provider offers the physical computing hardware that includes the processing power through a set of *central processing units* (CPUs) in a cluster. IaaS will also provide the online memory (Random Access Memory – RAM) and the disk storage. The PaaS provider is responsible for supplying and managing all the middleware platforms necessary to enable the software to run over the cloud. Finally, the SaaS provider will offer the software applications that will be used by the end-user.

The cloud vendor offers the QoE and the QoS that the end-user requires; the cloud vendor will provide the data security and meet the regulatory and legal requirements as required by the user or the regulators. The cloud vendor ensures that SaaS, PaaS, and IaaS are available to the end-user as services that are elastic and can scale up or scale down on demand. The cloud vendor also guarantees that the cloud service is fault-tolerant and is available on a continuous basis with proper security that includes *confidentiality, integrity, availability, authentication, authorization, accounting, and anonymity* (CI5A) [2]. The cloud vendor will charge the end-user for the consumed cloud resources based on the QoE.

At one time, energy companies used to manufacture power transmit it from generating stations to the distribution center, and finally deliver it to a household for a fee. The same was also true with telecommunications vendors who used to own the entire infrastructure starting from customer premises equipment to the transmission line. However, today many telecom and energy resellers are virtual operators who do not own any infrastructure. They use energy and telecom infrastructures from different providers to offer better QoS, cheaper tariffs, or value-added services.

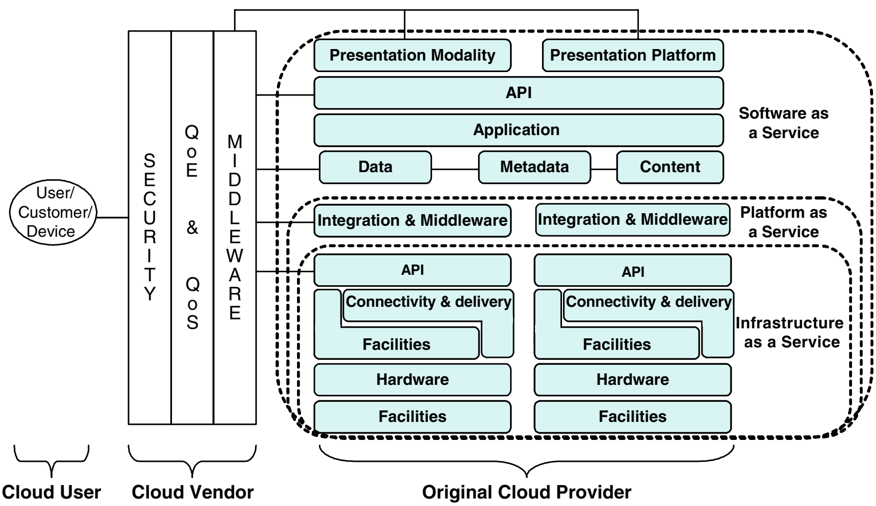

Fig. 20.1 The cloud reference model

This helped competition and improvement in service quality with new job creation and economic opportunities. Similarly, in the cloud, the cloud vendor is a virtual organization that offers the last-mile services to the end-user. It may not even own any cloud infrastructure – it will source cloud resources from various original cloud providers such as SaaS, PaaS, or IaaS from different parts of the world to offer the cloud services that meet certain SLA. The original cloud provider such as an IaaS, PaaS, or SaaS can also become a cloud vendor by offering guaranteed services quality and meeting local tax requirements.

In many cases, the *small and medium enterprises* (SMEs) or a household will interact only with the cloud vendor at the last-mile and may not even know the original cloud providers. The service model for the cloud vendor will mainly be driven by the end-to-end services they provide based on QoS, SLA, and QoE. These services may be the entire computing environment starting from software application to data storage and management or even simple resources such as four processors for 1 hour. The service can be private or public and accessible through any network whether wire-line or wireless. The pricing of the resources by the original cloud provider will be driven by some fixed price derived from raw computing power and the storage (memory and disk), whereas the pricing of the cloud service to the end-user will depend on the SLA and QoE the user perceives.

Currently, many large cloud providers are actively recruiting vendors and resellers of their services. Like any value-added reseller sales channel, the providers are looking to leverage the vendor's sales efforts, client relationship management expertise, and value-added services such as cloud application development or customization, legacy system integration, etc., to increase the provider's revenues and to maximize the utilization of their facilities. Cloud vendors will create new economic opportunities that promise to increase innovation and entrepreneurship in the delivery of the cloud services that will directly impact the QoS and QoE of the end-users.

20.2 Cloud Economics

Cloud economics as presented in this section refer to the economic forces, business drivers, and structural issues affecting the costs and benefits of adopting the cloud technologies.

20.2.1 Economic Context

As this chapter is being written, every enterprise in the world is facing a global economic recession that has profoundly affected all developed countries as well as those developing countries that develop products sold in those markets. Uncertain times also bring opportunities, but taking advantage of strategic opportunities typically must now be done quickly without additional capital funds or corporate resources.

In addition, for information technology (IT) managers, energy cost management is not a small issue[1]. The challenge today is to increase computing power utilisation with lower energy consumption. In addition, the maintenance of legacy enterprise data centers absorbs the majority of IT budgets and IT managers are looking for ways to create increased capacity and flexibility within their current computing facility and hardware footprint, thereby lowering costs and increasing their return on assets (ROA). There has been increasing attention paid to alternatives that provide the pay-as-you-go options, unlimited scalability, quick deployment, and the minimal maintenance requirements. Cloud computing is a paradigm that promises to meet all these requirements.

20.2.2 Economic Benefits

Occasionally used to refer to the economics of cloud computing, the term "Cloudonomics" was coined by Joe Weinman in a seminal article entitled "The 10 Laws of Cloudonomics" [4]. While far from being a comprehensive or exhaustive list of economic factors, his "10 Laws" serve as a useful starting point in our discussion. He examined the strategic advantages provided by public utility cloud services over private clouds and traditional data centers. He posits that public utility clouds are fundamentally different from traditional data center environments and private clouds. For individual enterprises, cloud services provide benefits that broadly fall into the categories of lowering overall costs for equivalent services (you pay only for what you use), increased strategic flexibility to meet market opportunities without having to forecast and maintain on-site capacity, and access to the advantages of the cloud provider's massive capacity: instant scalability, parallel processing capability, which reduces task processing time and response latency, system redundancy, which improves reliability, and better capability to repel botnet attacks. Further, public cloud vendors can achieve unparalleled efficiencies when compared with data centers and private clouds because they are able to scale their capacity to address the aggregated demand of many enterprises, each having different peak demand periods. This allows for much higher server utilization rates, lower unit costs, and easier capacity planning netting a much higher return on assets than is possible for individual enterprises. Finally, because the location of the public cloud vendor's facilities are not tied to the parochial interests of the individual clients, they are able to locate, scale, and manage their operations to take optimum advantage of reduced energy costs, skilled labor pools, bandwidth, or inexpensive real estate.

These are not the only benefits that have been identified. Matzke [5] suggests that the levels of required skills or specialized expertise along with the required

[1] IBM cites a study [3] that reports that US data center managers are anticipating a 35% increase in energy expenses over the next 4 years.

economies of scale drive the optimum choice for resourcing IT initiatives. For him, the availability of scalable skills combined with other economies of scale are among the compelling benefits of cloud computing[2]. This is especially true for enterprises that are located in labor markets that have very few or only very expensive IT staff resources available with the requisite skills.

20.2.3 Economic Costs

The costs associated with cloud computing facing early adopters include the potential costs of service disruptions; data security concerns; potential regulatory compliance issues arising out of sensitive data being transferred, processed or stored beyond defined borders; limitations in the variety and capabilities of the development and deployment platforms currently available; difficulties in moving proprietary data and software from one cloud services provider to another; integration of the cloud services with legacy systems; cost and availability of programming skills needed to modify legacy application to function in the cloud environment; legacy software CPU-based licensing costs increasing when moved to a cloud platform, etc.

20.2.4 Company Size, Economic Costs, and Benefits of Cloud Computing

The economic costs or benefits of implementing cloud services vary depending on the size of the enterprise and its existing IT resources/overheads including legacy data center infrastructure, computer hardware, legacy software, maturity of internal processes, IT staffing, and technical skill base. These factors determine the strategic costs and benefits that accrue to individuals and corporations depending on their relative size.

In the past, large corporations have had an advantage over small corporations in their access to capital and their ability to leverage their existing human, software, and hardware resources to support new marketing and strategic initiatives. However, since the advent of cloud computing, the barriers to entry for a particular market or market segment for a startup company have been dramatically reduced and cloud computing may have tipped the balance of strategic advantage away from the large

[2] Those with low requirements for economy of scale and skills can be addressed with on-site resources. Initiatives with low scalability requirements but higher skill requirements can be handled through traditional outsourcing arrangements. Projects with high scalability requirements but low skill scalability requirements can be addressed through collocation or traditional hosting arrangements. Finally, projects that require both economies of scale as well as scalable skills are best addressed by cloud computing all other things being equal [4].

established corporations towards much smaller or startup companies. A small, dedicated, and talented team of individuals can now pool their individual talents to address a perceived market need without an immediate need for venture capital funds to provide the necessary IT infrastructure. There are a number of cloud providers who provide software development environments that include the requisite software development tools, code repositories, test environments, and access to a highly scalable production environment on pay-as-you-go basis.

Also contributing to this trend is the open-source movement. While licensing issues, support, and feature considerations may dissuade larger enterprises from using open-source software in the development and deployment of their proprietary products, the availability of open-source software in nearly every software category has been a boon to SMEs, the self-employed, and startups.

As these small companies grow into midsize and large companies, they face changing cost equations that modify the relative costs and benefits of cloud computing. For instance, at certain data traffic volumes, the marginal costs of operating with a cloud provider's infrastructure may become more expensive than providing the necessary IT infrastructure in-house. At that point, there may be advantages of a mixed-use strategy in which some of the applications and services are brought in-house and others continue to be hosted in the cloud. The following tables will identify the differences that SMEs and large enterprises face in both the benefits and costs of cloud services (Tables 20.1 and 20.2).

20.2.5 The Economics of Green Clouds

The development of green data centers and green clouds is shaped by two important factors. The first is a global awareness of the devastating potential of climate change due to human activity primarily through carbon emissions. The second is the rising costs of energy. These two factors have focused IT infrastructure planning and decision-making on energy cost reduction, dynamic resource allocation strategies, and have moved green issues from the category of nice-to-do to strategically important for all midsize and large corporations. In 2008, IBM did more than 30 energy assessments around the world and found that 60–70% of the energy used in the data centers was used for indirect purposes such as cooling and lighting the facilities with only 30–40% of the energy being used directly by the computing hardware [3].

Public cloud providers locate their data centers where bandwidth, cheap energy, abundant water for cooling, and proximity to markets are optimal. Google [6] and other cloud providers have focused on creative approaches to efficient resource usage including not only electricity usage but also water recycling and equipment recycling upon disposal. Through purchasing servers and other equipment designed to minimize energy usage, these cloud providers minimize the non-computing energy overhead and maximize their utilization rates through the dynamic allocation of

Table 20.1 Economic benefits of cloud adoption

Economic benefits	Small and medium enterprises (SMEs)	Large enterprises
Strategic flexibility	Critical in getting quickly to market. Cloud services allow startups to rapidly develop and deploy their products as long as they can use the open source or proprietary development platforms of the cloud providers. As the cloud market offerings mature, there will be many more platform options available.	Cloud services can provide large enterprises the same strategic benefits as startups for new initiatives as long as legacy software integration and data issues are not significant. With appropriate software development talent, operating units can rapidly develop and market test new innovations without putting additional strain on IT budgets, staff, or hardware. Longstanding internal IT management policies and standards may have to be re-examined and modified to allow this to happen.
Cost reduction	Pay-as-you-go pricing may be critical if operating capital or venture capital funding is not available. With cloud services, growth can more easily be funded through operating revenues and there may be tax advantages to converting what would have been longer-term depreciation expenses to fully loaded current expenses.	Cloud services provide the same cost benefits for isolated and exploratory initiatives. Instant availability and low setup costs for new development and deployment environments allow operating units to explore new initiatives quickly at low cost without increasing internal IT hardware or staff overheads. For high data traffic volumes, it may become more economical to bring the operations in-house. Because maintaining legacy hardware and software absorb the majority of IT costs, large corporations may see significant costs savings by selectively moving noncritical applications and processes to external clouds.
Software availability	Software as a Service (SaaS) and Platform as a Service (PaaS) provide necessary software and infrastructure at low entry cost. Limited online version functionality may be more than offset by dramatic cost savings.	Existing volume licensing of legacy desktop and process-integrated enterprise software may make the status quo more attractive if end-user retraining, process modifications, and other change costs are high. Legacy desktop software may have more features and functionality than is currently available in SaaS versions. But the legacy software licensing costs may dramatically increase if it is hosted in a private cloud environment.

(continued)

Table 20.1 (continued)

Economic benefits	Small and medium enterprises (SMEs)	Large enterprises
Scalability	One of the most dramatic benefits for SMEs and startups. If successful, applications designed to autoscale can scale endlessly in a cloud environment to meet the growing demand.	Large enterprises with significant hardware, legacy software, and staff resources can benefit from cloud scalability by identifying CPU-intensive processes such as image processing, PDF conversion, and video encoding that would benefit from the massively scalable parallel processing available in clouds. While this may require modifying legacy applications, the speed benefits and reduced local hardware requirements may far outweigh the software modification costs.
Skills and staffing	While the proper design of cloud applications requires high-level software development skills, their maintenance and support is vastly simplified in the cloud environment. Cloud providers handle all maintenance and support issues for both hardware and platform software at costs that are either bundled into the usage fees or available in various configurations as premium services. This allows significant cost savings through reduced staff overheads.	Because the majority of enterprise IT costs goes to support legacy applications and hardware, the greatest staffing benefits will be seen in new cloud initiatives that do not add to the staffing burden. Longer term, as the enterprise begins to analyze cloud technology potential for its legacy operations, retraining of existing staff or bringing in new staff with cloud technology skills will be necessary to take advantage of the new paradigm. Thus, some investment will have to be made before large-scale or long-term benefits will be seen. The staffing investment may be significant if the enterprise is attempting to create a private cloud to handle dynamic resource allocation and scalability across its operating units. In this case, it may face significant staff investment as well as the required hardware, software, and network investment to implement and maintain their private cloud.
Energy efficiency	Because SMEs can dramatically reduce or eliminate local servers, cloud computing provides direct utility cost savings as well as environmental benefits.	Even very large enterprise IT data centers cannot achieve the energy efficiencies found in the massive facilities of public cloud providers even with aggressive high-density server and virtualization strategies. In periods of economic downturns, green initiatives typically cannot compete for scarce capital funds. By employing a mixed strategy that off-loads applications and processing to external clouds when feasible, IT managers are able to minimize their energy costs and carbon footprint.

(continued)

Table 20.1 (continued)

Economic benefits	Small and medium enterprises (SMEs)	Large enterprises
System redundancy and data backup	This is a large benefit for SMEs, the majority of which are poorly prepared for hardware failures and disaster recovery [3]. Cloud storage can reduce downside risks at low cost.	Because cloud technologies distribute both data storage and data processing across potentially large number of servers, the likelihood of data loss due to hardware failure is much lower than in most large private data centers. The cloud data storage can provide a cost effective supplemental back-up strategy.

computing resources. This combination of lower energy overhead amortized over a much higher server utilization rate allows cloud suppliers to provide computing services far more efficiently with a much smaller energy and carbon footprint.

Because of the scale of operations of large cloud providers, they are able to achieve efficiency rates and server utilization rates that are unachievable in even large corporate data center operations. Thus, cloud computing holds the promise of not only providing attractive cost savings at the enterprise level but also may contribute to the larger societal objectives of energy efficiency and environmental protection and sustainable development.

20.3 Quality of Experience in the Cloud

To retain and recruit customers in the cloud, the experience of the customer has to be managed in a very sensitive fashion. In the cloud, experience will be measured in terms of experience in a *virtual environment* (VE) [8] where challenges will relate to user-agents and devices, the virtualized environments used, the presence attributes, and the tasks to be performed. *Experience assurance* (AE) in the cloud will deal with a community of vendors, providers, and partners; where the cloud vendor will empower the customer – the customer will be able to choose and measure the perceived value of a service. In addition, the cloud vendor must be proactive – communicating a problem before the customer discovers it; also, a remediation must be in place before customer asks for it. Experience happens through *moment of truth* (MoT), when people meet people; therefore, the cloud vendor must be in constant touch with the customer and also must improve based on the feedback from the customer.

To ensure security and service quality in the cloud, a cloud vendor has to go beyond its own domain of control. This becomes even more complex when the cloud vendor is a virtual organization and does not own service infrastructures. For example, to provide a secure and fault-tolerant service, the cloud vendor must ensure that all the original cloud providers in the value-chain agree on some level of security

Table 20.2 Economic costs of cloud adoption

Economic costs	Small and medium enterprises (SMEs)	Large enterprises
Data security	SMEs are better able to use third-party services such as payment processing to handle secure transactions.	Data is an enterprise's most important IT and operating asset. Current uncertainty regarding the security of the data assets stored in public clouds is one of the most significant barriers in cloud adoption. Large enterprises may not want their data stored in countries where intellectual property piracy is prevalent. Some companies may not want their data stored on equipment used by their competitors.
Data confidentiality	SMEs face the same data confidentiality issues as large enterprises.	One of the advantages of cloud computing and storage for confidentiality is that the data transfer and storage algorithms encrypt the data into units that are difficult to reconstruct without the specialized algorithms/keys if the data are intercepted in transfer or the cloud security is compromised.
Data regulations	SMEs face the same regulatory data location issues as large enterprises.	Depending on the company's industry, there may be significant regulatory issues regarding data location. Data that identifies the individual in certain health and financial contexts are subject to US regulations. Similarly, the EU has laws that restrict the transfer of certain data outside of its borders.
Data integrity	The data integrity and reliability of cloud suppliers may be higher than that provided by the existing internal systems.	Cloud technologies are relatively new and storage and data transfer algorithms slice the data into small units, which are stored and transferred dynamically within the storage region. Estimating and factoring the risks of potential data corruption of mission critical data at this early stage of cloud implementation may be difficult leading to nonadoption, especially if the existing internal systems, processes, and protocols are working.
Data transfer costs	For new initiatives that do not require the transfer of legacy data to the clouds, transfer costs are minimal. Getting locked into a particular cloud service provider is currently a market concern due to the lack of open standards among the providers.	Moving the existing data sets to clouds requires data integrity check to ensure that all of the data has been transferred fully and that it has not been corrupted. For very large data sets, this may represent significant staff costs. Cloud vendors typically charge data transfer costs. If the data set is large and there is significant data churn due to transaction processing, it may be more cost-effective to look at more traditional hosting options.

(continued)

Table 20.2 (continued)

Economic costs	Small and medium enterprises (SMEs)	Large enterprises
Integration costs and legacy application reengineering	In startups and small companies, potentially little or no integration is required between cloud applications and legacy applications.	Potentially significant costs to have new cloud applications interact with legacy applications or to modify legacy applications to offload processing to cloud-based components. Conversely, there may be advantages to reengineering legacy applications and hosting them in a public cloud when integrating Web 2.0 functionality with legacy applications.
Software licensing	Cloud services (SaaS, PaaS) provide significant software licensing cost savings for startups and small companies.	Migrating large enterprises to cloud based SaaS may not be cost-effective relative to the existing enterprise licensing agreements. Depending on the licensing agreements for third-party software, especially if licensing fees are based on the number of CPUs using the software, hosting legacy applications in a cloud environment may involve significantly increased licensing costs or noncompliance with the agreements if the software is installed on a machine image used for autoscaling as the user demand increases.
Cloud availability – "rolling brownouts"	Unavailability of the cloud services or slow performance due to heavy traffic is a serious concern when choosing a cloud vendor.	Same as with SMEs. Currently, even large vendors have experienced slow performance or suspended service due to overwhelming utilization.

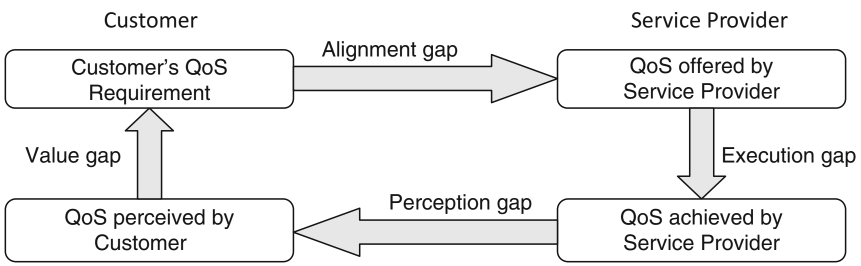

Fig. 20.2 Viewpoints of quality of service

policy and QoS guarantee. The cloud vendor can learn from telecom industry and implement the ITU Recommendation G1000 [9] where QoS is expressed on a service-by-service basis. For QoS to be truly useful across the industry, it must be meaningful from four viewpoints, which are illustrated in Fig. 20.2. These viewpoints are:

- Customer's QoS requirements
- Service provider's offerings of QoS (or planned/targeted QoS)
- QoS achieved or delivered
- Customers' survey ratings of QoS (perceived QoS)

To address these viewpoints in a timely manner, the cloud vendor can consider the *grade of service* (GoS) used in telecom traffic engineering [10]. The GoS deals with resource instead of services – it deals with number of controls to provide a measure of adequacy of a group of resources under specified conditions. The key point to conform to GoS standards is to apportion individual values to each network element in a fashion that the target end-to-end QoS is obtained.

The challenge here is that the GoS and QoS have different interpretations. While the QoS views the situation from the customer's point of view, the GoS takes the provider's point of view. To resolve the ambiguity, it is necessary to introduce *service level agreement* (SLA) in this context. An SLA is a contract between a customer and the cloud vendor to define QoE. The purpose of SLA will be to ensure that QoE is understood in the same manner by the customer and the cloud vendor. Also, it can be implemented in the cloud using definitions and rules [11]. Furthermore, the SLA defines what is to happen in case the terms of the contract are violated [12].

The security challenges in the cloud will be higher and more complex compared with what the world has seen earlier. The major difference is that a user does not have full control of the infrastructure and the people who manage the data and the cloud infrastructure. Many security attacks that were not possible in a private network will be possible in the cloud owing to its large attack surface. Therefore, in addition to standard security offered by the cloud providers, there will be separate end-to-end security services provided by the cloud vendor.

To realize QoS, QoE, and security, we propose the *cloud service quality manager* (CSQM) architecture as shown in Fig. 20.3. There are six entities in this architecture.

1. Access requestor (AR)
2. Policy decision point (PDP)
3. Policy repository (PR)
4. Policy enforcement point (PEP)
5. Cloud decision point (CDP)
6. Service Quality Manager (SQM)

The *access requestor* (AR) is an endpoint device or user-agent seeking access to some service or resource from the service provider. The *policy decision point* (PDP) is a system where a policy decision related to security requirement or QoS requirement is made. Typically, the policies fall into two main categories: general policies that are applicable to all the users; specific policies that are applicable to an individual user, a particular service, or a group of users or services defined in an SLA. The SLA is designed to meet certain *key performance indicators* (KPIs) based on certain *key quality indicators* (KQIs). Policies are stored in *policy repository* (PR). The PR will coordinate with other databases such as inventory for services, resources, and GoS. The policy server will host the PR and evaluate the policy

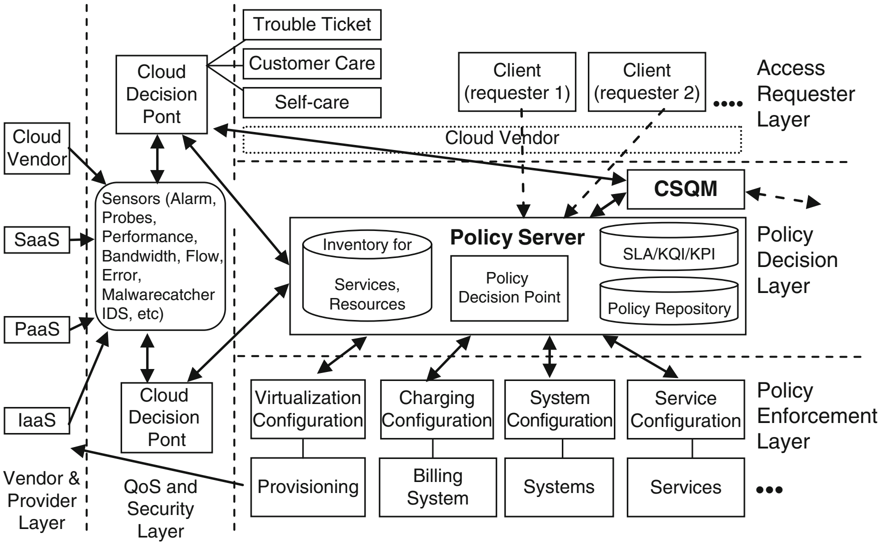

Fig. 20.3 The cloud service quality management architecture

conformance through the PDP. The *policy enforcement point* (PEP) is responsible for enforcing a policy. Because policy may not directly be understood by all equipments or applications, it is necessary to translate these policies into service-specific configuration rules and enforce through activation and control systems. The *cloud decision point* (CDP) captures, interprets, and decides about the events received from the cloud vendor and original cloud providers like SaaS, PaaS, and IaaS. These events are alarms, performance, and security data collected in proactive and reactive fashion. A CDP works like a sensor that processes various events and sends them to the PDP for review and policy enforcement. All these policy servers and CDPs will be managed by the *cloud service quality manager* (CSQM).

20.4 Monetization Models in the Cloud

In the cloud, there are four different models of monetization:

1. Each and every service is priced and charged to the consumer. IaaS and PaaS will fall in this category – IaaSs and PaaSs will monetize the services they offer. All single tenancy resources will fall into this category – in single-tenancy a resource can be used by only one user at any given point in time – here demand--supply-driven pricing will prevail. Some SaaS services will also fall in this category – this model for SaaS will evolve from the earlier model of *application service provider* (ASP). This model is quite successful in the wireless networks where network operators are in control of the network and therefore all the services that are offered through these networks are monetized. Monetization of

SaaS will be transaction-based. Even a multitenancy object will be converted into single-tenancy object through *digital rights management* (DRM).

2. The second model of monetization will be offering part of the service free and part of the service as chargeable. Here, the free part of the service will mainly be matchmaking platforms, such as job sites portals, dating sites, search engines, or the virtual travel agents. Here, the monetization will be through the match or completion of a transaction. In this model, the service provider will offer the *content* free and determine the *intent* of the user for using this content. Once the intent is known, the provider will propose a match and commit a business transaction.
3. The third model is where a service is free. The user is free to use or modify the service or content. This will follow the principle of *Bhikshu economy*. Bhikshus are Buddhist monks who offer service for free – in return, community supports their livelihood ("366: A Bhikshu who, though he receives little, does not despise what he has received") [7]. If one finds a value in it, one makes a contribution. Unlike a capitalistic economy, pricing is not dependent on demand and supply – one can pay any amount that is worth the experience. Another interesting concept of *Shramadana* from Buddhist philosophy will prevail in the cloud, wherein public pays back by joining the community and offering their intellect, time, and labor instead of cash. Wikis and GNU software are examples of this practice.
4. The fourth model is free service that might have some restriction for monetization. Many governments are following the principle that all outcomes of research projects funded by governments will be open-domain where not only the results but also the data will be available in the open domain for not-for-profit usage. Healthcare-related projects in the USA and other parts of the world fall in this category.

Data logistics will play a significant role in the cloud monetization. Data logistics will include functions like

- Data acquisition
- Data cleaning
- Data transformation
- Data transportation
- Data storage (offline)

Data acquisition or cleaning of data will be a complex process where it might be a service provided by the SaaS provider or the cloud vendor. Though not likely, data acquisition and cleaning service might be offered by the IaaS or the PaaS through a partner. Data acquisition will deal with a first-time user where the data need to be transformed into the electronic form. Data might exist in paper form or some other nonelectronic form, which need to be converted into electronic form understandable and accessible by the software application. In data cleaning service, the data will be examined and validated to ensure that the data that has been captured is indeed correct and free from redundancy or missing components. Data transformation will be a service where the data of the end-user is transformed into a format that is understandable by the software application. Transportation of data will mainly be the role of the cloud vendor where the data is transported from the end-user's premise to the computing infrastructure in the cloud.

20.5 Charging in the Cloud

The charging for the resources and invoicing the end-user will be the responsibility of the cloud vendor. For the cloud usage, the cost to the end-user will be the combination of communication cost and the charges the user will pay to the cloud vendor. Communication will be provided by an *internet and communication service provider* (ICSP). The ICSP charges will mainly be based on the traditional spatial and temporal properties of the single-tenancy resource usage like bandwidth, and duration of usage. Cloud computing has the following three characteristics.

1. Infinite virtual computing resources available on demand, thereby eliminating the need for cloud computing users to plan far ahead for provisioning.
2. The elimination of an upfront commitment by cloud users, thereby allowing companies to start small and increase hardware resources only when there is an increase in their needs.
3. The ability to pay for use of computing resources on a short-term basis as needed (e.g., processors by the hour and storage by the day) and release them as needed, thereby rewarding conservation by letting machines and storage go when they are no longer in use.

Capital expenses (capex) versus *operational expense* (opex) is one of the advantages of using cloud computing. There have been many discussions comparing the cost of a 24 × 7 use of a cloud infrastructure from a cloud vendor like Amazon EC2 instance against the cost of hosting a server within the data center. Usually, providers take the average price of a 1U server, divide it by 36 (the number of months in the typical expected service life of a piece of equipment), and compare the savings. If operational costs like energy, cooling, manpower, etc. are included, the cloud looks very attractive from an operational costs' point of view.

20.5.1 Existing Models of Charging

The existing models of charging can be divided into charges by the IaaS, PaaS, and the SaaS. In case of a SaaS business with varying demand over time and revenue proportional to user hours in an IaaS, Armbrust et al. [13] have proposed the tradeoff charging model through the following equation:

$$UserHours_{cloud} \times (revenue - Cost_{cloud}) \geq UserHours_{datacenter} \times \left(revenue - \frac{Cost_{datacenter}}{Utilization}\right)$$

They also proposed the revenue equation for adverts-supported model in which the number of adverts served is roughly proportional to the total visit time spent by end-users on the service.

20.5.1.1 On-Demand IaaS Instances

On-Demand Instances from an IaaS allows a customer to pay for compute capacity by the hour with no long-term commitments. This frees the customer from the costs and complexities of planning, purchasing, and maintaining hardware and transforms what are commonly large fixed costs into much smaller variable costs. For example, at Amazon for an Extra Large Instance with 15 GB of memory, 8 EC2 Compute Units (four virtual cores with two EC2 Compute Units each), 1,690 GB of instance storage, 64-bit platform will cost \$0.80 per hour. However, there are hidden costs in many of these charging models; one such hidden cost worth mentioning here is the data access. Some cloud vendors offer storage at a very attractive price but charge on transactions that accesses the disk.

20.5.1.2 Reserved IaaS Instances

Reserved Instances by an IaaS gives the customer an option to make a low, one-time payment for each instance the customer wants to reserve and in turn receives a significant discount on the hourly usage charge for that instance. After the one-time payment for an instance, that instance is reserved for the customer, and the customer has no further obligation.

Simple statistics reveal that *reserved instances* though give a cloud customer the option to make a low, one-time payment for each instance, they are not suitable for a short-term usage. Hence, we envisage a new charging model of *Value Instance*. Here, the one-time payment for each instance to be reserved is calculated taking into consideration a percentage of the on-premise hardware cost.

20.5.1.3 PaaS Charging

Just getting the computing resource from the IaaS provider may not be sufficient; the charges for the PaaS need to be provisioned. PaaS cloud vendors enable an application where they charge their platforms on rental basis. These rentals are based on the number of servers or number of instances of PaaS the customer will need to use. If the application is not cloud ready, there could be additional charges for cloud enablement. There are different charging models for the PaaS user. These are sometimes charged per-resource like a piece of middleware, which might be in a range of \$100–500 a year. Some of the PaaS providers charge on a per user basis, a model similar to Google App Engine.

20.5.1.4 Cloud Vendor Pricing Model

Because QoS and SLA play a significant role in the cloud, the cloud vendor will have a back-to-back QoS and SLA with both the ICSP and the cloud providers that will need to provide QoE- and SLA-based charging as well. If there is an SLA violation, a credit to the user will have to be initiated.

20.5.1.5 Interprovider Charging

There will be many cases where the revenue collected by the cloud vendor needs to be shared with partners and other providers who are part of the value-chain. This demands an inter-provider charging agreement that will rate and calculate the charges payable to or receivable from the partner provider like the SaaS, PaaS, or the IaaS. This will be driven by the following considerations:

1. Bill & keep – this is a special type of billing agreements between the providers where the provider keeps [14] all the money they collect from the subscriber. Nobody shares any revenue with any other provider.
2. Usage of resource is measured, rated, and billed at the *point of interconnection* (POI). Rates will be determined by service combined with spatial, temporal, and instance attributes.

20.6 Taxation in the Cloud

It is easy to formulate a taxation policy for tangible movable or immovable assets; it is also possible to formulate a taxation policy when these assets cross the border of a state. Tax is levied at the point of consumption of the service; therefore, conventional taxation principles will not be able to support the complex needs of taxation in a virtual cloud environment. Cloud computing is predicated on a concept of borderless global services. Governments, for one reason or another, do not like this idea – at a basic level, governments need borders.

The taxation in the cloud will be the responsibility of the cloud vendor who will have a local tax registration and be governed by the local tax regulations. Taxation in the cloud can be managed with similar taxation model as mobile network operators or *mobile virtual network operator* (MVNO). A mobile subscriber can consume the service of the home service provider while at the home network; the subscriber can use the service of a foreign network being present at the home network. The subscriber can also be roaming in a foreign country with different taxation policies and consume services of the foreign network or the home network. Similarly, in the cloud, the end-user could be in one country and the cloud vendor could be in another country offering services from providers that originate in other countries.

Over a period of time, we believe that there will be clearing houses that will manage the interstate and intercountry taxations of the consumables. This might lead to a situation where there are dangers of double taxation. If tax is based on the location of the registered office of a cloud computing company, then there is always an option to the virtual offices to be located in a lower tax or tax-free export zone.

References

1. Security Guidance for Critical Areas of Focus in Cloud Computing (2009, April) Prepared by the Cloud security alliance
2. Talukder AK, Chaitnya M (2008) Architecting secure software systems. CRC Press, Boca Raton, FL
3. Lee S, Cooper LF (2009, August) IT managers discover the high cost of ignoring data center efficiency problems. BizTechReports.Com. Cited in an IBM WebEx presentation entitled: dynamic infrastructure in action: reducing costs while increasing value. http://researchlibrary.theserverside.net/detail/RES/1254491035_498.html
4. Weisman J (2008) GigaOM network: the 10 laws of Cloudonomics. BusinessWeek Online. http://www.businessweek.com/technology/content/sep2008/tc2008095_942690.htm. Originally posted 6 September 2008
5. Matzke P (2008, November) Cloud computing: from vision to reality. http://download.sczm.t-systems.de/t-ystems.de/en/StaticPage/55/02/30/550230_10_Presentation_Cloud-Computing-ps.pdf. Original presentation given 25 November 2008
6. Google (2009). Efficient computing: data center efficiency measurements. http://www.google.com/corporate/green/datacenters/measuring.html. Accessed September 2009
7. The Bhikshu (Mendicant) from The Dhammapada, A Collection of Verses. http://www.sacred-texts.com/bud/sbe10/sbe1027.htm
8. Gaggioli A, Bassi M, Delle Fave A (2003) Quality of experience in virtual environments. In: Riva G, Davide F, IJsselsteijn WA (eds) Being there: concepts, effects and measurement of user presence in synthetic environments. Los Press, Amsterdam, p 121
9. ITU-T Recommendation G.1000, Communications quality of service: A framework and definitions
10. ITU-T E.600, Terms and definitions of Traffic Engineering, 1993
11. RFC3644, Policy Quality of Service (QoS) Information Model
12. Oodan A et al (2002) Telecommunications quality of service management: from legacy to emerging services. Institution of Electrical Engineers
13. Armbrust M et al (2009, Feb 10) Above the clouds: a Berkeley view of cloud computing. UC Berkeley Reliable Adaptive Distributed Systems Laboratory. http://radlab.cs.berkeley.edu/
14. Berger U (2004) Bill-and-Keep vs. cost-based access pricing revisited. http://ideas.repec.org/p/wpa/wuwpio/0408002.html