CHAPTER 1

Introduction

The catchword *Cloud Computing* may sound new to the arena of computing as well as information technology; although the concept behind it is not completely new. In a broad sense, it is a step (forward) to access the web-based services through Internet; or in general terms, cloud computing represents an advance in the field of computing technology where management of computing services is outsourced to greater extents and in multiple dimensions. To be specific, computing facilities are being made available as services offered by reputed vendors.

Computing technology has been evolved over the years. There have been steady developments in the field of computing hardware, software architecture, web technology and network communications over the last decade. The speed of internetworks has increased day by day and it has also become cheaper. All these developments contributed in settinging the stage for the initiation of the revolutionary concept of 'cloud computing'.

Cloud computing has changed the way computation happens. It provides the means for smarter ways to do business and accordingly it makes life simpler. This chapter tries to highlight the shortcomings in the *traditional computing* approaches and it attempts to iterate how the stage was ready for the arrival of this new model of computing.

> *Cloud computing has evolved over the years and has changed the way people live and do business over the World Wide Web.*

1.1 WHAT IS THE BUZZ ABOUT

Cloud computing provides the means for users to easily avail computing facilities whenever and wherever required. They need not worry about setting up infrastructure, purchasing new equipment or investing in the procurement of licensed software. Rather they can access any volume, large or small, of computing facilities in exchange for some nominal payment. It is a new model of computing which has become possible through integration of advanced computing models, sophisticated web technologies and modern network communication technologies (especially high-speed Internet).

Cloud computing became a hot topic for large computing vendors (companies like Amazon, Google, Microsoft etc.) only from mid-2008, but the concept itself is quite old. People have different perceptions regarding this new computing model. To some, it is a new and radical innovation. Others think that it is nothing new: only the pace of modernization has been explored to obtain new uses of technologies which already existed. The fact is, like every new technology, cloud computing too has emerged from an ancestral domain of existing technology and its context.

Cloud computing is the delivery of computing services over the Internet.

1.2 LIMITATIONS OF THE TRADITIONAL COMPUTING APPROACHES

Every new technology emerges with the promise to resolve the shortcomings of the existing ones. *Traditional computing* has played a pivotal role in the field of computing and communication over the past few decades. Since cloud computing is being considered as the successor of the traditional computing system, therefore, it would be wise to recognize the limitations of the traditional computing approaches before studying the contents of cloud computing.

Computing and information technology (IT) has changed the nature and scope of the human civilization in last few decades. There was a time, nearly half-of-the-decade back, when enterprises used to execute their businesses merely by the aid of the pen, paper, telephone and fax machine. Gradually, computer systems intruded manual processes and started automating them. The pen and paper were replaced by digital communication, and even the phone and faxing services started being managed by computers.

At present, businesses from local to global, are dependent on the computing systems for almost everything they do. Even individuals depend heavily on computing systems for their day-to-day activities. IT and computing are critical factors now, and life cannot be imagined without easy and all-time access to computing systems.

Easy and cheap access to computing facilities has become essential for everyone. But, a little investigation raises concerns about several issues regarding the conventional uses of computing technology. Following section focuses on problems associated with traditional computing approaches.

1.2.1 Enterprise Perspective

Enterprises have been valuable consumers of computing since inception. They have always been among the front-runners for adopting computing-based process automation for running day-to-day business activities. Table 1.1 points out the difficulties faced by enterprises with *traditional computing* approach.

Table 1.1 Difficulties faced by enterprises in traditional computing

Traditional Computing Scenario	*Problematic facts and related questions*
Business without help of computing services is beyond imagination, and the customized software packages manage business activities. Most organizations use ERP packages (implemented by some IT enterprise) to get maximum benefits from regular business operations.	To run enterprise resource planning applications, business organizations need to invest huge volumes of capital to setup the required IT infrastructure. Servers, client terminals, network infrastructure are required, and they to be put together in a proper manner. Moreover, arranging adequate power supply, cooling system and provisioning space also consume a major part of the IT budget. Are there ways to avoid this huge initial investment for computing infrastructural setup?

Traditional Computing Scenario	*Problematic facts and related questions*
Business application package implementation also over-burdens the IT enterprises with many other costs. Setting up infrastructure, installation of OS, device drivers, management of routers, firewalls, proxy servers etc. are all responsibilities of the enterprise in traditional computing approach.	Enterprises (or IT service firms) need to maintain a team of experts (system maintenance team) in order to manage the whole thing. This is a burden for HR management and incurs recurring capital investment (for salaries). Can enterprises get relief from these responsibilities and difficulties? It would help them concentrate fully on the functioning of business applications.
Even those IT enterprises whose sole business interest is developing applications are bound to setup computing infrastructure before they start any development work.	This is an extra burden for enterprises who are only interested in application development. They can outsource the management of infrastructure to some third party, but the cost and quality of such services varies quite a bit. Can IT enterprises avert such difficulties?
Computing infrastructure requires adequate hardware procurement. This procurement is costly, but it is not a one-time investment. After every few years, existing devices become outdated as more powerful devices appear.	It becomes difficult to compete in the market with outdated hardware infrastructure. Advanced software applications also require upgraded hardware in order to maximize business output. Can this process of upgrading hardware on a regular basis be eliminated from an enterprise's responsibility?
It is not unusual to find an updated version of application with new releases that is more advanced and apt to keep up with changing business scenario.	Adopting an updated version of an application requires necessary efforts from subscriber's end. Fresh installation and integration of components need to be done. Can subscribers be relieved of this difficulty of periodically upgrading the applications?
Capacity planning of computing resources is a critical task for any organization. Appropriate planning needs time, expertise and budgetary allocation since low resource volume hampers the pace of the performance of applications.	Enterprises generally plan and procure to support the maximum business load that they have anticipated. But average resource demand remains far less, most of the time. This causes resource wastage and increases the recurring cost of business. If this capacity planning task could be made less critical and resource procurement strategy more cost effective?
Resource requirements of a system may increase or decrease from time to time.	Individual enterprises cannot manage system contraction in a way that unutilized resources of a system can be utilized in some other system so that the cost of the business could be reduced. If this were somehow possible?
Many enterprise computing systems run forever without stopping. Such systems host applications which require round-the-clock availability to fulfill business demand.	When resource capacity expansion of such system becomes an absolute requirement for the respective business, an system shutdown (hence service disruption) becomes unavoidable which may cause loss in the business. If a system could be expanded without shutting it down?

Hence, an enterprise can be overburdened when handling multiple issues in the traditional way of computing. Issues like huge initial investment for setting up computing infrastructure,

difficulty of managing the complex systems, difficulty in regularly replacing outdated things with latest available technology etc. have always been the matters of concern.

1.2.2 Individual User's Perspective

Individual users have also been consumers of computing since a long time. They use computing for various purposes like editing documents, developing program, playing games, watching videos, accessing Internet etc. Table 1.2 focuses on the troubles individual users of computing usually face with the *traditional computing* approach.

Table 1.2 Difficulties faced by individual users in traditional computing approach

Traditional Computing Scenario	*Problematic facts and related questions*
To work with software applications (like text editor, image editor, programming, games etc.), users first need to procure computing system where these software applications run.	For general users (who don't want to experiment with computer hardware devices), this initial capital investment for setting up computing infrastructure is often more than the software applications they use! If this huge unnecessary investment for procuring hardware components could be avoided?
Requirement analysis and procurement of hardware infrastructure are responsibility of the users. But, actual utilization of these resources depends on frequency of user access and the kind of software applications they run over it.	General users are usually not experts of computing systems. They are often misguided and procure unnecessary volume/capacity of hardware resources, most portions of which remain unutilized. This reduces the return on investment (ROI). If this approach could be changed; if users did not have to procure a fixed volume of hardware resource prior to its actual use or demand?
A hardware component may fail for many reasons. Maintenance of the hardware infrastructure is the users' responsibility.	Time, cost and the uncertainty are involved in the process. If users could get relief from these responsibilities and difficulties?
Computing systems (desktop, laptop etc.) procured by most users are hardly used for few hours daily on an average.	Non-utilization of procured systems results in wastage of resource with regard to total investment. If the hardware resources would be available on payment of usage basis?
Software licensing cost needs separate budgetary allocation. Licenses are sold for fixed period of time (usually for one year-duration).	If software is used for 2–5 hour per day on an average during licensing period, it depicts 8 %–20 % utilization of entire investment. If this cost could be reduced? If the licensing fee would be paid on hourly usage basis?
Users are burdened with installation and critical customization of software. They also troubleshoot in case the software crashes.	Professional help can be obtained against payment, or users can troubleshoot themselves, thereby investing more time. If users could get relief from these responsibilities and difficulties?

Traditional Computing Scenario	*Problematic facts and related questions*
Users need to have physical access to the system for using a personal computing system.	Though portable computing devices are available (like laptop, tablet etc.), it may not be possible to carry them all the time. Could there be a way of accessing personal computing systems remotely, from any location, any time?
Within few years, hardware systems become outdated. It becomes difficult to run advanced or new software on them.	Users have no other option but to throw out the whole setup and replace it with a new one. If there be a permanent solution to this wastage (from users' end)?

So, individual users also face many difficulties with traditional computing. Initial necessities of procuring expensive physical system, maintenance of the hardware, maintenance of software have always been prime concerns for general users. Moreover, within just few years of the whole investment the need for hardware up-gradation emerges.

All these limitations of traditional computing paved the way for the emergence of an alternative model of computing. Cloud computing has redefined the way computing has ever worked.

1.3 IS THERE ANY SOLUTION TO THESE WORRIES?

When an organization or an individual user needs to set up in-house IT infrastructure (small or big) for computing purpose, they have two options:

- Doing it themselves (by deploying *in-house* team, in case of an organization)
- *Outsourcing* the responsibility to some third party who are well-trained for the job

But in both of these cases, there remain concerns which bother the users, such as hardware management or software installation or application implementation etc. which are not likely to be part of the core competencies of users. Other worries have already been listed in Table 1.1 and Table 1.2. Is there a way to remove these worries from the minds of individual users or enterprises? If the answer is 'Yes' then they will be able to concentrate more on their core tasks.

IT *service outsourcing* is a popular model which has been adopted by enterprises over the last two decades, where all of the application development and implementation related activities/ responsibilities are transferred to IT service companies. But, traditional outsourcing is not the best solution to all problems mentioned above.

Traditional outsourcing may remove the worries of system or application maintenance to some extent but it cannot eliminate all of them, as different levels of computing facility subscribers have different types of requirements. For example, enterprises who only use applications to run their business activities have some kind of computing requirements, but those IT service organizations to whom they outsource the complex tasks of application development or implementation have their separate computing requirements. Application development team of an IT service organization may in turn depend on some system (physical computing system), whereas the assembling and maintenance group who need not to have knowledge of application development may have concerns that are entirely different. They will certainly have separate requirements. In brief, difficulties faced by users (of computing) depend on the layers of computing they work on.

The concerns of the users of computing vary with the layers of their computing activities.

1.4 THREE LAYERS OF COMPUTING

Computers and computing have become an integral part of our daily lives. Different people use different categories of computing facilities. These computing facilities can be segmented into three categories:

- Infrastructure
- Platform
- Application

These three categories of computing facilities form three layers in the basic architecture of computing. Figure 1.1 represents the relationships between these three entities.

1.4.1 Infrastructure

The bottom layer or the foundation is the 'computing infrastructure' facility. This includes all physical computing devices or hardware components like the processor, memory, network, storage devices and other hardware appliances. Infrastructure refers to computing resources in their bare-metal form (without any layer of software installed over them, not even the operating system). This layer needs basic amenities like electric supply, cooling system etc.

1.4.2 Platform

In computing, platform is the underlying system over which applications run. It can be said that the platform consists of the physical computing device (hardware) loaded with layer(s) of software where the program or application can run. The term 'computing platform' refers to different abstract levels. It consists of:

- Certain hardware components, only.
- Hardware loaded with an operating system (OS).
- Hardware and OS, additionally, loaded with run-time libraries.

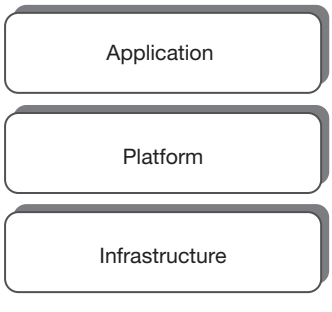

FIG 1.1: Three layers of computing facilities

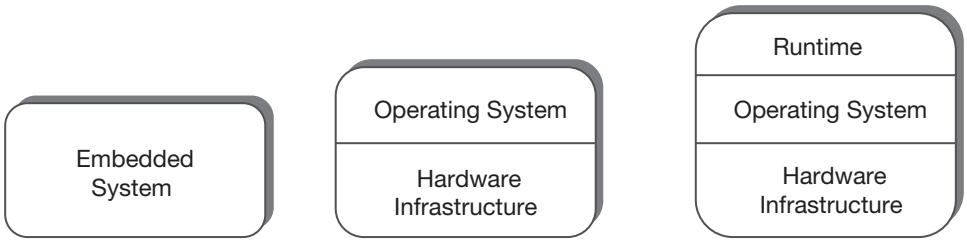

FIG 1.2: Computing platforms in different forms

Hardware alone can be considered as the platform in case of embedded systems, where physical resources can be accessed without any operating system.

A fully configured physical computer loaded with an operating system is considered as a platform for computing. Different platforms can be installed over the same computing infrastructure. Linux or Windows operating systems installed over the same physical computer (computing infrastructure) can provide two different computing platforms.

The platform layer is also the place where software developers work. Hence Integrated Development Environments (IDEs) and runtimes are part of this layer. Java Development Kit (JDK) or .NET are examples of popular computing frameworks. Software applications can be developed and run over these platforms.

Platform layer provides the platform to execute the applications; in addition it facilitates application development activities.

1.4.3 Application

Applications (application software) constitute the topmost layer of this layered architecture. This layer generally provides interfaces for interaction with external systems (human or machine) and is accessed by *end users* of computing. A user actually works on the application layer while he or she is going to edit a document, play a game or use the calculator in a computer. At this layer, organizations access enterprise applications using application interfaces to run their business.

Different types of people work at different layers of computing. They need to have different skill-sets and knowledge. Figure 1.3 shows the general users or subscribers of three different computing layers.

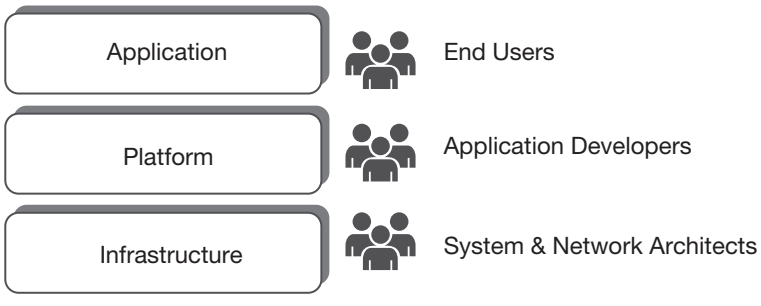

FIG 1.3: Different users/subscribers of three computing layers

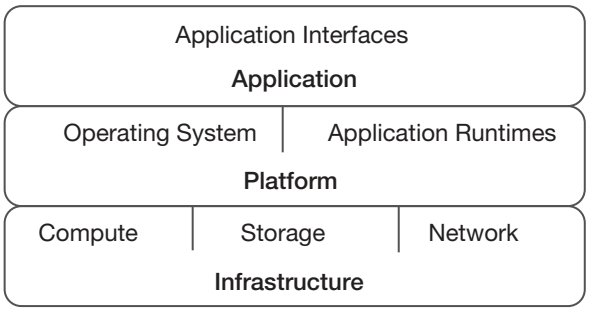

FIG 1.4: Elements of three computing layers

An upper layer in this architecture is dependent on underlying layer(s). Access to the topmost layer effectively consumes facilities from all underlying layers. Thus, a person working online or offline with some software application on a personal computing device or directly accessing the Internet basically consumes all these three facilities together.

Figure 1.4 illustrates the different constituents of the three layer of computing. In this figure, the term 'compute' refers to the set of resources required for assembling a computing system or computer, particularly processor and memory components.

1.5 THREE LAYERS IN TRADITIONAL COMPUTING

In the traditional approach, the boundaries between computing layers were not very clear to the general users. Most *end users* unknowingly had to bear the burden of being concerned about all these three computing layers, while their actual area of interest was only around the application layer. Similarly, developers had to be concerned about the infrastructure layer apart from their own layer of activities (the platform layer). This section briefly focuses on the complexities associated with these three computing layers in the traditional computing model from users' viewpoint.

1.5.1 Traditional Infrastructure Model

Traditionally enterprises arrange computing infrastructure themselves by forming an internal computing system management team or they depend upon some third-party vendors. These vendors or internal system management departments are responsible for managing the whole computing infrastructure of the enterprise, including planning, designing, assembling of physical servers, network or data storage infrastructure build ups and other services like load balancing of electricity supply.

The whole process of setting up of such an infrastructure is quite complex. For a new office, this process generally takes weeks and sometimes months. Even the purchase cycle of additional hardware to power up the infrastructure is not counted in hours or minutes, but in days or weeks. Protection and security of the infrastructure appears as an extra burden for any organization.

1.5.2 Traditional Platform Model

In the traditional model, the platform building and management tasks need special expertise. Enterprises either need to maintain an internal team or can outsource these jobs. *Computing platform* building activities involve installation of an operating system (OS), application runtimes or application development environments on computing infrastructure. For example, Java Runtime Environment (which provides a platform for the applications) is essential to run any java based application.

Traditionally the entire responsibility of platform installation, configuration, updates and other staffing at appropriate levels often fall on the shoulders of the people concerned with application development or even on users who are interested only in using applications. Thus, the actual assignments get delayed. Licensing, timebound installation of patches or maintenance of the platform cause difficulty for users (application developer or application user).

1.5.3 Traditional Application Model

At application layers, the users are the end users of computing. They need not have any knowledge about a complex computing system. Their only interests are to access or use different software applications to fulfill requirements of personal or business related works. Software applications provide solutions for various needs. In traditional model, computing infrastructure or computing platform or both often become concerns of application subscribers, which is a critical scenario. Apart from this, the traditional software applications attract fixed and prepaid licensing costs and annual support costs.

> *In traditional computing model, subscribers of one computing layer cannot fully escape the responsibilities and difficulties of arranging or managing the underlying layer(s).*

1.6 THE END OF TRADITIONAL COMPUTING

The limitations of traditional computing were evident for a long time and several futuristic ideas were proposed on computing techniques. But for various reasons, especially due to technological constraints, it was not possible to implement these ideas.

With technological advancement in several fields of computing and communication (i.e. in network communication), computing giants came up with a new model of computing during the second half of the last decade. The philosophy of this model says that all kind of computing facilities can be delivered as services. According to this model, the three main *layers of computing*, infrastructure, platform and application can be delivered to the consumers as ready-made stuff arranged and maintained by others, whenever they need it, relieving the consumers from the burden of arranging all the stuff themselves. This new model of computing is known as *cloud computing*.

> *In traditional approach, computing subscribers were always over-burdened with many additional difficulties and cost.*

1.6.1 Computing as the Utility Service

Introduction of cloud computing precipitated a significant shift in the responsibilities of managing computing resources. Computing facilities in this model are supplied in the same way as like as a civic authority supplies water or electricity in a city. Customers can use those facilities without being worried about how they are being supplied or who is managing all of these activities.

The three major aspects of computing which were represented in three-layered computing architecture in Figure 1.1 earlier, are delivered as *utility services* in cloud computing model. Hence three services corresponding to this model are:

- Infrastructure Service
- Platform Service
- Application or Software Service

Customers wanting to avail some kind of computing facilities, need not start from scratch any more. If a customer wants to work at the application level, he/she no longer needs to worry about setting up the infrastructure and platform facilities. Rather, she can directly avail the ready-made infrastructure and platform services from the cloud vendors. Again customers who want to work on platform layer of computing, get readily available infrastructure service without any difficulty.

The only thing required at customer's end to avail these readymade *computing services* (infrastructure, platform or software) are Internet connection and any basic computing device (PC, laptop, tab etc.) where software interface with cloud computing systems can run. These computing devices need not be highly configured since the local computers no longer have to do heavy tasks and cloud interface applications are fairly light-weight. If local computers are free from doing the heavy tasks that they used to do earlier, who then, is doing these on their behalf? The answer is: 'cloud' which is going to be assumed as a network of computers now onward.

Since, cloud resides at some remote location and does all the heavy computing tasks, the hardware and software demands on the local computers at users' side decrease.

1.6.2 Who Provides the Service?

Cloud computing vendors who are Independent Software Vendors (ISV) generally provide the service. These vendors are reputed computing/IT giants. They are the owners or developers of the clouds and manage everything. Users can use cloud services by payment on use basis, just like they pay monthly bill for electricity service. This becomes profitable for both parties, the customers and the vendors. Vendors can supply the service at cheaper rates because of the large size of their business (due to, economy of scale) since the number of computing subscribers is very large.

This model of computing is much talked about because it provides a lot of flexibility compared to the traditional way of computing. For instance, it drastically reduces the cost and complexity of owning and operating computers and networks. Moreover, since specialized computing vendors manage the cloud the quality of service undoubtedly gets better.