A Literature Survey on Cloud Computing

P.Nagendra Babu¹, M.Chaitanya Kumari², S.Venkat Mohan³

¹M.Tech Student, Computational Engineering in CSE, RGUKT APIIIT – RK VALLEY, Andhra Pradesh, India
²M.Tech Student, Computational Engineering in CSE, RGUKT APIIIT – RK VALLEY, Andhra Pradesh, India
³Head of the Department, CSE, Govt Polytechnic College – Vempalli, Andhra Pradesh, India

Abstract—From past decade onwards Cloud Computing plays a vital role. Cloud computing is a computation, software, data access, and storage services that do not require end-user knowledge of the physical location and configuration of the system that delivers the services. Cloud computing is the fastest new paradigm for delivering on demand services over internet and can be described as internet centric software. Cloud computing describes a new supplement, consumption, and delivery model for IT services based on Internet protocols, and it typically involves provisioning of dynamically scalable and often virtualized resources. It is a byproduct and consequence of the ease-of-access to remote computing sites provided by the Internet. This frequently takes the form of web-based tools or applications that users can access and use through a web browser as if they were programs installed locally on their own computers. In this paper we discussed all the basic concepts of Cloud Computing. It is very useful to do research in this area.

Keywords— Cloud Computing, QoS, SaaS, PaaS, IaaS, PSO, Cloud Scheduling

I. INTRODUCTION

Cloud computing is a model for enabling 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. Typical cloud computing providers deliver common business applications online that are accessed from another web service or software like a web browser, while the software and data are stored on servers. Most cloud computing infrastructures consist of services delivered through common centers and built-on services. Clouds often appear as single points of access for consumer’s computing needs. Commercial offerings are generally expected to meet quality of service (QoS) requirements of customers, and typically include service level agreements (SLAs) [1], [16].

II. HISTORY OF COMPUTING

When the era of personal computers began in the 1970s, it was generally called micro-computers. It later became an affordable consumer goods when the microprocessors where developed. The workstation computers started in the 1990s. With various developments of hardware and processor components, the computers became much more easily available for cheaper costs. The applications that were running on the personal computers were platform dependent and there was no direct connectivity to share data from one to another. Security was a major concern during this time. System administration and responsibility of data, system crash and virus infections became widespread[1], [2].

It was Tim Berners Lee who developed the World Wide Web in the early 1990s. With the introduction of web browser, the World Wide Web found a turning point. The dot com bubble came into picture. Many companies found success developing various business models [2].

Cloud computing has evolved through a number of phases that includes grid and utility computing and Software as a Service (SaaS). Though the idea and concept of computer network in global terms was thought in the 1960s, cloud computing came into light when internet started offering its service.

Cloud computing reached the public in the year 2007. The model of Cloud computing changed how computers could be used and how information could be spread. Google provided the search engine that could be accessed for free from any internet connected computer.

Companies started offering cloud applications that targeted both individuals and businesses in both free and paid versions. The benefit was to reduce company’s budget as well as energy by ‘outsourcing’ the business applications to the cloud. Features like security, data back, development of new features etc. was the responsibility of the solution provider rather than the internal IT department or staff.

The different types of cloud computing services are offered today, that has become essential for everyday business process. For businesses, it is now very easy to find and get any kind of application or feature that they desire [3].

III. CHARACTERISTICS OF COMPUTING

In cloud computing, users access the data, applications or any other services with the help of a browser regardless of the device used and the user’s location. The infrastructure which is generally provided by a third-party is accessed with the help of internet. Cost is reduced to a significant level as the infrastructure is provided by a third-party and need not be acquired for occasional intensive computing tasks. Less IT skills are required for implementation [10], [12].

1. Agility improves with user’s ability to rapidly and inexpensively re-provision technological infrastructure resources.
2. Application Programming Interface (API) accessibility to software that enables machines to interact with cloud
software in the same way the users interface facilities interaction between humans and computers. Cloud computing systems typically use REST-based APIs.

3. **Costs** claimed to be greatly reduced and in a public cloud delivery model capital expenditure is converted to operational expenditure. This ostensibly lower barriers to entry, as infrastructure is typically provided by a third-party and does not need to be purchased for one-time or infrequent intensive computing tasks. Pricing on a utility computing basis is fine-grained with usage-based options and fewer IT skills are required for implementation (in-house).

4. **Device and location independence** enables users to access systems using a web browser regardless of their location or device they are using (e.g., PC, mobile phone). As infrastructure is off-site (typically provided by a third-party) and accessed via the Internet, users can connect from anywhere.

5. **Multi-tenancy** enables sharing of resources and costs across a large pool of users thus allowing for:
   - **Centralization** of infrastructure in locations with lower costs (such as real estate, electricity, etc.)
   - **Peak-load capacity** increases (users need not engineer for highest possible load-levels)
   - **Utilization and efficiency** improvements for systems that are often only 10-20% utilized.

6. **Reliability** improved if multiple redundant sites are used, which makes well designed cloud computing suitable for business continuity and disaster recovery.

7. **Scalability** via dynamic (“on-demand”) provisioning of resources on a fine-grained, self-service basis near real-time, without users having to engineer for peak loads, performance is monitored and consistent and loosely coupled architectures are constructed using web services as the system interface.

8. **Security** cloud improves due to centralization of data, increased security-focused resources, etc., but concerns can persist about loss of control over certain sensitive data, and the lack of security for stored kernels. Security is often as good as or better than under traditional systems, in part because providers are able to devote resources to solving security issues that many customers cannot afford. However, the complexity of security is greatly increased when data is distributed over a wider area or greater number of devices and in multi-tenant systems which are being shared by unrelated users. In addition, user access to security audit logs may be difficult or impossible. Private cloud installations are in part motivated by users desire to retain control over the infrastructure and avoid losing control of information security.

9. **Maintenance** of cloud computing applications is easier, because they do not need to be installed on each user’s computer. They are easier to support and to improve, as the changes reach the clients instantly.

10. **The computing is “in the cloud”** that is, the processing (and the related data) is not in a specified, known or static place(s). This is in contrast to a model in which the processing takes place in one or more specific servers that are known. All the other concepts mentioned are supplementary or complementary to this concept.

IV. ADVANTAGES OF CLOUD COMPUTING

Cloud Computing has different advantages over past techniques. It has following Advantages [3-5], [14].

a. **Easy Management**

   The maintenance of the infrastructure, be it hardware or software is simplified, thus, less headaches for the IT team. Also applications that are quite storage extensive are easier to use in the cloud environment compared to the same when used by the organization by its own. Also at the user level, what you mostly need is a simple web browser with internet connectivity.

b. **Cost Reduction**

   The main advantage for SMEs lies here. Cloud computing drastically reduces the IT spending for SMEs. Costly systems need not be required for occasional use of intensive computing resources. Also the man power required for such systems is not required. Even simple applications like email can be set up and mostly free through applications like Google Apps. Also as most of the time such providers are quite reliable in terms of availability, it is clear winner.

c. **Uninterrupted Services**

   Lower outages are provided by cloud computing services, thus providing uninterrupted services to the user. However, some occurrences of outages have occurred in the past, like the Gmail outage in 2009. Also other cloud vendors like EC2 have failed at some point of time, but however, they are much more dependable compared to the infrastructure installed on the organization.

d. **Disaster Management**

   In case of disasters, an offsite backup is always helpful. Keeping crucial data backed up using cloud storage services is the need of the hour for most of the organizations. Also cloud storage services not only keep your data off site, but they also ensure that they have systems in place for disaster recovery.
e. Green Computing -

Harmful emissions due to extensive use of systems in organizations, electronic waste generated as the time passes and energy consumption is the main disadvantage of the present day computing systems. This can be reduced to some extent by using cloud computing services. This leads to environment preserving. Also the e-waste is generated to minimum extent.

V. CLOUD ARCHITECTURE

Cloud architecture [10, 11], the systems architecture of the software systems involved in the delivery of cloud computing, typically involves multiple cloud components communicating with each other application programming interfaces. Usually web services and 3-tier architecture. This resembles the UNIX philosophy of having multiple programs each doing one thing well and working together over universal interfaces. Complexity is controlled and the resulting systems are more manageable than their monolithic counterparts.

The two most significant components of cloud computing architecture are known as the front end and the back end. The front end is the part seen by the client, i.e. the computer user. This includes the client’s network (or computer) and the applications used to access the cloud via a user interface such a web browser. The back end of the cloud computing architecture is the “cloud” itself, comprising various computers, servers and data storage devices[7].

A. Layers of Cloud

In this section we discussed how the layers of cloud will work and what the services we used to develop a cloud are [7], [8], [9], and [10].

1. Client- Cloud client consists of computer hardware and/or computer software that relies on cloud computing for application delivery, or that is specifically designed for delivery of cloud services and that, in either case, is essentially useless without it. Examples include some computers, phones and other devices, operating systems and browsers [7], [8], [10].

2. Cloud application services or “software as a service (SaaS) [13]” deliver software as a service over the internet, eliminating the need to install and run the application on the customer’s own computers and simplifying maintenance and support. People tend to use the terms “SaaS” and “cloud” interchangeably, when in fact they are two different things. Network-based access to, and management of, commercially available software. Activities that are managed from central locations rather than at each customer’s site, enabling customers to access applications remotely via the web. Application delivery that typically is closer to a one-to-many model (single instance, multi-tenant architecture) than to a one-to-one model, including architecture, pricing, partnering, and management characteristics. Centralized feature updating, which obviates the need for downloadable patches and upgrades. SaaS delivers special purpose software that is remotely accessible. Figure3 [9] explains Google Maps, Live Mesh from Microsoft etc.

3. Platform - Cloud platform services [8] or “Platform as a Service (PaaS)” deliver a computing platform and/or solution stack as a service, often consuming cloud
infrastructure and sustaining cloud applications. It facilitates deployment of applications without the cost and complexity of buying and managing the underlying hardware and software layers. PaaS offers high level integrated environment to build, test, and deploy custom apps. Restrictions on software used to develop apps in exchange for built-in scalability. Figure 4 [9] explains Google App Engine.  

4. **Infrastructure**- Cloud infrastructure services, also known as “Infrastructure as a Service (IaaS)”, delivers computer infrastructure—typically a platform virtualization environment—as a service [7], [10]. Rather than purchasing servers, software, data-center space or network equipment, client instead buy those resources as a fully outsourced services. Suppliers typically bill such services on a utility computing basis and amount of resources consumed (and therefore the cost) will typically reflect the level of activity. IaaS evolved from virtual private server offerings. Cloud infrastructure often takes the form of a tier 3 data center with many 4 attributes, assembled from hundreds of virtual machines. Hardware, software, equipments can scale up and down dynamically(elastic). Figure 5 [9] explains

- Amazon Elastic Compute Cloud (EC2) and Simple Storage Service (S3)
- Eucalyptus: open source Cloud implementation compatible with EC2 (allows to set up local cloud infra prior to buying services)

5. **Server** - The server’s layer consists of computer hardware and/or computer software products that are specifically designed for the delivery of cloud services, including multi-core processors, cloud-specific operating systems and combined offerings[7], [16]

<table>
<thead>
<tr>
<th>TABLE 1 VARIOUS CLOUD COMPUTING SERVICES</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>SAAS</strong></td>
</tr>
<tr>
<td>Software As A Service</td>
</tr>
<tr>
<td>✓ Government Applications</td>
</tr>
<tr>
<td>✓ Communications</td>
</tr>
<tr>
<td>✓ Productivity tools</td>
</tr>
<tr>
<td>Examples:</td>
</tr>
<tr>
<td>✓ Oracle</td>
</tr>
<tr>
<td>✓ SalesForce.com</td>
</tr>
<tr>
<td>✓ LinkedIn</td>
</tr>
</tbody>
</table>

VI. DEPLOYMENT MODELS

There are different ways in which a cloud environment can be deployed. Figure 6 [15] shows the different types of cloud computing can be deployed are

1. **Public Cloud**—

Public cloud or external cloud describes cloud computing in the traditional mainstream sense, whereby resources are dynamically provisioned on a fine-grained, self-service basis
over the Internet, via web applications/web services, from an off-site third-party provider who bills on a fine-grained utility computing basis. It is a standard computing model [15, 16].

2. Community Cloud–
A community cloud may be established where several organizations have similar requirements and seek to share infrastructure so as to realize some of the benefits of cloud computing. With the costs spread over fewer users than a public cloud (but more than a single tenant) this option is more expensive but may offer a higher level of privacy, security and/or policy compliance. Examples of community clouds include Google’s “Gov. Cloud”[16].

3. Hybrid cloud and hybrid IT delivery–
The main responsibility of the IT department is to deliver services to the business. With the proliferation of cloud computing (both private and public) and the fact that IT departments must also deliver services via traditional, in-house methods, the newest catch-phrase has become “hybrid cloud computing”. Hybrid cloud is also called hybrid delivery by the major vendors including HP, IBM, Oracle and VMware who offer technology to manage the complexity in managing the performance, security and privacy concerns that results from the mixed delivery methods of IT services.

A hybrid storage cloud uses a combination of public and private storage clouds. Hybrid storage clouds are often useful for archiving and backup functions, allowing local data to be replicated to a public cloud [15, 16].

Another perspective on deploying a web application in the cloud is using Hybrid Web Hosting, where the hosting infrastructure is a mix between cloud hosting and managed dedicated servers- this is most commonly achieved as a part of a web cluster in which some of the nodes are running on real physical hardware and some are running on cloud server instances.

4. Combined Cloud –
Two clouds that have been join together are more correctly called a “combined cloud”. A combined cloud environment consisting of multiple internal and/or external providers “will be typical for most enterprises”. By integrating multiple cloud services users may be able to ease the transition to public cloud services while avoiding issues such as PCI compliance [15, 16].

5. Private Cloud–
Douglas park hill first described the concept of a “private computer utility” in his 1966 book the challenge of computer utility. The idea was based upon direct comparisons with other industries (e.g., the electricity industry) and the extensive use of hybrid supply models to balance and mitigate risks.

“Private cloud” [15 - 17] and “internal cloud” have been described as Neologisms, but the concepts themselves pre-date the term cloud by 40 years. Even with in modern utility industries, hybrid models still exist despite the formation of reasonably well-functioning market sand the ability to combine multiple providers.

Some venders have used the term to describe offerings that emulate cloud computing on private networks. These (typically virtualization automation) products offer the ability to host applications or virtual machines in a company’s own set of hosts. These provide the benefits of utility computing – shared hardware costs, the ability to recover from failure, and the ability to scale up or down depending upon demand.

Private clouds have attracted criticism because users “still have to buy, build, and manage them” and thus do not benefit from lower up-front capital costs and less hands-on management, essentially “[lacking] the economic model that makes cloud computing such an intriguing concept”. Enterprise IT organizations use their own private cloud(s) for mission critical and other operational systems to protect critical infrastructures.
VII. CLOUD SCHEDULING

The Scheduling [18] is a key concept in computer multitasking, multiprocessing operating system and real-time operating system designs. Scheduling refers to the way processes are assigned to run on the available CPUs, since there are typically many more processes running than there are available CPUs. This assignment is carried out by software known as a scheduler or dispatcher. A Scheduling algorithm is the method by which threads, processes or data flows are given access to system resources. This is usually done to load balance a system effectively or achieve a target quality of service. The need for a scheduling algorithm arises from the requirement for most modern systems to perform multitasking and multiplexing. Scheduling refer to the way process are assigned to run on the available CPU.

Job scheduling [19] is a fundamental issue in achieving high performance in cloud computing systems. However, it is a big challenge for efficient scheduling algorithm design and implementation. Unlike scheduling problems in conventional distributed systems, this problem is much more complex as new features of cloud systems such as its dynamic nature. And the high degree of heterogeneity of jobs and resources must be tackled. Job scheduling in computational cloud is a multi-objective optimization problem.

A. Purpose of the System

Cloud computing environments facilitate applications by providing virtualized resources that can be provisioned dynamically. However, users are charged on a pay-per-use basis. User applications may incur large data retrieval and execution costs when they are scheduled taking into account only the ‘execution time’.

In addition to optimizing execution time, the cost arising from data transfers between resources as well as execution costs must also be taken into account. In this project, we present a Particle swarm optimization (PSO) based heuristic to schedule applications to cloud resources that takes into account both computation cost and data transmission cost. We experiment with workflow application by varying its computation and communication costs.

B. Scope

The CLOUD Scheduler is intended to work as an add-on to the queuing, scheduling and resource managing module of the open source CLOUD Engine resource management system, previously known as SCHEDULER IN CLOUD. The scheduler will offer market-based economy driven service for managing batch jobs on clouds by scheduling CPU time according to user utility rather than system performance considerations. Hence, the main objective of CLOUD SCHEDULER is to provide Quality of Service (QoS) computational economy in CLOUD computing [16].

The PSO algorithm which is developed for cloud systems can also be used where

- Alternate solutions are too slow or overly complicated
- Need an exploratory tool to examine new approaches

✓ Problem is similar to one that has already been successfully solved by using a PSO
✓ Want to hybridize with an existing solution.
✓ Quality of Service (QoS)[18] – meeting what users actually want, thus maximizing their utility.
✓ Computational Economy - Computational economy refers to the inclusion of user-specified QoS parameters with jobs so that resource management is based on a user-centric approach rather than on a system-centric approach. This essentially means that user constraints such as deadline and budget are more important in determining the priority of a job by the scheduler, than system policies like ordering jobs according to the basis of submission time. Currently, there is no holistic scheduling mechanism in cloud computing to enable differing QoS levels for different clients.
✓ Sequential Jobs - These are jobs that must be run on a single node, and cannot be split into separate jobs.
✓ Process Migration - It is the process of moving a job from one computer system to another without restarting the job from the beginning.

C. Types of Schedulers

A scheduler is a software product that allows an enterprise to schedule and track computer batch tasks. These units of work include running a security program or updating software. Job schedulers may also manage the job queue for a computer cluster. A scheduler is one of the main components of IT infrastructure. A scheduler may also be known as a job scheduler. The following we discuss some types of Schedulers in Cloud Computing [20 - 25].

1. Optimized Resource Scheduling Algorithm:

An optimized scheduling algorithm [26] to achieve the optimization or sub-optimization for cloud scheduling. In this algorithm an Improved Genetic Algorithm (IGA) is used for the automated scheduling policy. It is used to increase the utilization rate of resources and speed.

2. Improved Cost-Based Algorithm for Task Scheduling:

An improved cost-based scheduling algorithm[27] for making efficient mapping of tasks to available resources in cloud. This scheduling algorithm measures both resource cost and computation performance, it also Improves the computation/communication ratio.

3. Innovative transaction intensive cost-constraint scheduling algorithm:

A scheduling algorithm which takes cost and time [28]. The simulation has demonstrated that this algorithm can achieve lower cost than others while meeting the user designated deadline.

4. Compromised-Time-Cost Scheduling Algorithm:

Compromised-time-cost scheduling algorithm [29] which considers the characteristics of cloud computing to accommodate instance-intensive cost-constrained workflows by compromising execution time and cost with user input enabled on the fly.
5. Particle Swarm Optimization-based Heuristic for Scheduling Workflow Applications:

A particle swarm optimization (PSO) based heuristic to schedule applications [30] to cloud resources that takes into account both computation cost and data transmission cost. It is used for workflow application by varying its computation and communication costs. The experimental results show that PSO can achieve cost savings and good distribution of workload onto resources.


SHEFT workflow scheduling algorithm to schedule a workflow elastically on a Cloud computing environment [31]. The experimental results show that SHEFT not only outperforms several representative workflow scheduling algorithms in optimizing workflow execution time, but also enables resources to scale elastically at runtime

7. Market-Oriented-Hierarchical Scheduling:

A market-oriented hierarchical scheduling strategy [32] which consists of a service-level scheduling and a task-level scheduling. The service-level scheduling deals with the Task-to-Service assignment and the task-level scheduling deals with the optimization of the Task-to-VM assignment in local cloud data centers.

8. Multiple QoS Constrained Scheduling Strategy of Multi-Workflows (MQMW):

This algorithm [33] has a strategy implemented for multiple workflow management system with multiple QoS. The scheduling access rate is increased by using this strategy. This strategy minimizes the make span and cost of workflows for cloud computing platform.

9. Optimal Workflow based Scheduling (OWS)algorithm:

OWS algorithm for scheduling workflows in a cloud [34] environment. The scheduling algorithm finds a solution that meets all user preferred QoS constraints. With this algorithm, a significant improvement in CPU utilization is achieved.

10. Resource-Aware Scheduling Algorithm (RASA):

A new task scheduling algorithm RASA. [35] It is composed of two traditional scheduling algorithms; Max-min and Min-min. RASA uses the advantages of Max-min and Min-min algorithms and covers their disadvantages. The experimental results show that RASA outperforms the existing scheduling algorithms in large scale distributed systems.


This algorithm [36] first calculates average execution time for each task and average communication time between resources of two successive tasks. Then tasks in the workflow are ordered (non-increasing) on a rank function. The task with higher rank value is given higher priority. In the resource selection phase tasks are scheduled in the order of their priorities and each task is assigned to the resource that can complete the task at the earliest time.

VIII. CONCLUSIONS

In this paper we discussed the basics of Cloud Computing, History of Computing, Characteristics, Advantages and Disadvantages of Computing. These are very useful to understand the cloud computing. Besides of above topics we discussed some more topics Cloud Architecture, Deployment models of Cloud, Cloud Scheduling, Cloud Architecture, Deployment models of Cloud is used to understand the working procedure of cloud, how it is used for organizations and how the cloud is used to store, retrieve and modify the data without physical equipment. In Cloud Scheduling we discussed maximum number of scheduling algorithms. We hope this paper will give efficient information what the readers want about Cloud Computing and very useful to researchers.

REFERENCES

[20] Er. Shimpny, Mr. Jagandeep Sidhu, “Different Scheduling Algorithms In Different Cloud Environment,” ISSN 2278-1021


[34] P.Varalakshmi, Aravindh Ramaavamy, Arwath Balasubramaniyan and Palanippan VijayaKumar, “An Optimal Workflow Based Scheduling and Resource allocation in Cloud”
