ISSN 2319-524X

Infrastructure as a Service: an Overview and its Impact on Digital Libraries

Dr. Ahmad Raza Khan

Assistant Professor

College of Computer and Information Science, Majmaah University

Al Majmaah

ar.khan@mu.edu.sa

***Abstract*— Cloud computing is a new technique of computing that is extensively used in today's' industry as well as society. It is a contemporary model and one of the latest computer industry buzzwords. Cloud computing brings the revolutionary changes in the world of Information Communication Technology because of its potential benefits such as reduced cost, accessible anywhere anytime as well as its elasticity and flexibility. This paper wills discoursing the cloud computing, definitions, historical background, characteristics, models and application of new generation libraries**.

***Keywords*— Cloud Computing, Libraries. Information Communication Technology, Cloud Computing Models,**

AN OVERVIEW

Cloud computing is a computing paradigm, where a large pool of systems are connected in private or public networks, to provide dynamically scalable infrastructure for application, data and file storage. With the advent of this technology, the cost of computation, application hosting, content storage and delivery is reduced significantly. Cloud computing is a practical approach to experience direct cost benefits and it has the potential to transform a data center from a capital-intensive set up to a variable priced environment. The idea of cloud computing is based on a very fundamental principal of reusability of IT capabilities. The difference that cloud computing brings compared to traditional concepts of "grid computing", "distributed computing", "utility computing", or "autonomic computing" is to broaden horizons across organizational boundaries.

What is Cloud Computing?

As we explained above cloud computing here are some more simple words and the paragraphs to understand efficiently cloud computing.

A. We see Cloud Computing as a computing model, not a technology. In this model "customers" plug into the "cloud" to access IT resources which are priced and provided "on-demand". Essentially, IT resources are rented and shared among multiple tenants much as office space, apartments, or storage spaces are used by tenants. Delivered over an Internet connection, the "cloud" replaces the company data center or server providing the same service. Thus, Cloud Computing is simply IT services sold and delivered over the Internet.

B. Cloud computing is the use of computing resources (hardware and software) that are delivered as a service over a network (typically the Internet). The name comes from the use of a cloud-shaped symbol as an abstraction for the complex infrastructure it contains in system diagrams. Cloud computing entrusts remote services with a user's data, software and computation.

C. The Internet is often represented as a cloud and the term "cloud computing" arises from that analogy. Accenture defines cloud computing as the dynamic provisioning of IT capabilities (hardware, software, or services) from third parties over a network. McKinsey says that clouds are hardware-based services Offering compute, network and storage capacity where: hardware management is highly abstracted From the buyer; buyers incur infrastructure costs as variable OPEX [operating expenditures]; and Infrastructure capacity is highly elastic (up or down).The cloud model differs from traditional outsourcing in that customers do not hand over their own IT resources to be managed. Instead they plug into the cloud, treating it as they would an internal data center or computer providing the same functions.

Historical Background of Cloud Computing:-

The origin of the term ***cloud computing*** is obscure, but it appears to derive from the practice of using drawings of stylized clouds to denote networks in diagrams of computing and communications systems. The word *cloud* is used as a metaphor for the Internet, based

ISSN 2319-524X

on the standardized use of a cloud-like shape to denote a network on telephony schematics and later to depict the Internet in computer network diagrams as an abstraction of the underlying infrastructure it represents. The cloud symbol was used to represent the Internet as early as 1994. In the 1990s, telecommunications companies who previously offered primarily dedicated point-to-point data circuits began offering virtual private network (VPN) services with comparable quality of service but at a much lower cost. By switching traffic to balance utilization as they saw fit, they were able to utilize their overall network bandwidth more effectively. The cloud symbol was used to denote the demarcation point between that which was the responsibility of the provider and that which was the responsibility of the users. Cloud computing extends this boundary to cover servers as well as the network infrastructure.

John McCarthy opined in the 1960s that "computation may someday be organized as a public utility." Almost all the modern-day characteristics of cloud computing (elastic provision, provided as a utility, online, illusion of infinite supply), the comparison to the electricity industry and the use of public, private, government, and community forms, were thoroughly explored in Douglas Parkhill's 1966 book, *The Challenge of the Computer Utility*. Other scholars have shown that cloud computing roots go all the way back to the 1950s when scientist Herb Grosch (the author of Grosch's law) postulated that the entire world would operate on dumb terminals powered by about 15 large data centers.[10] Due to the expense of these powerful computers, many corporations and other entities could avail themselves of computing capability through time sharing and several organizations, such as GE's GEISCO, IBM subsidiary The Service Bureau Corporation, Tymshare (founded in 1966), National CSS (founded in 1967 and bought by Dun & Bradstreet in 1979), Dial Data (bought by Tymshare in 1968), and Bolt, Beranek and Newman marketed time sharing as a commercial venture.

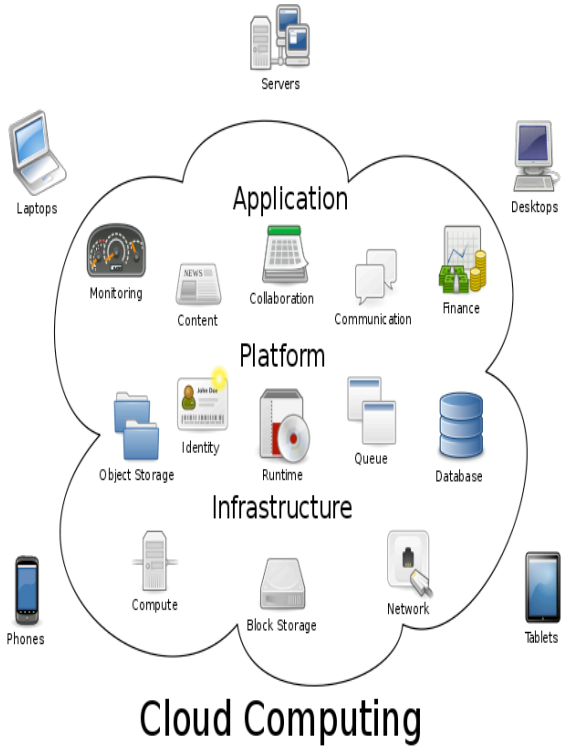
Definitions:-

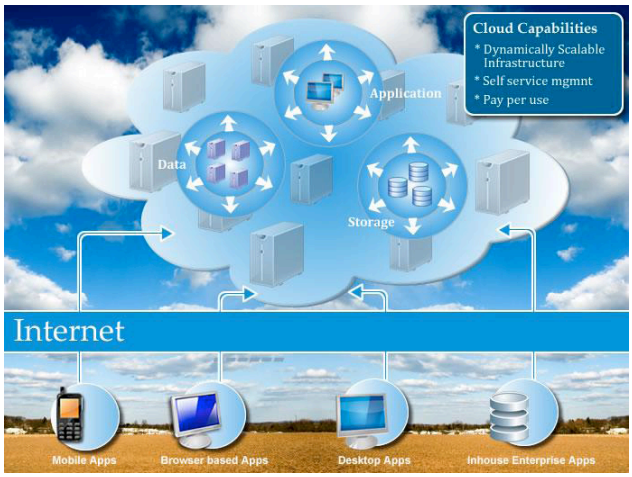
Forrester defines cloud computing as:

1. "A pool of abstracted, highly scalable, and managed compute infrastructure capable of hosting end customer applications and billed by consumption."
2. Berkely says "Cloud Computing refers to both the applications delivered as services over the Internet and the hardware and systems software in the datacenters that provide those services."

   Even though IBM's Irving Wladawsky Berger reports a leading analyst as having said recently that 'There is a clear consensus that there is no real consensus on what cloud computing is,' here are no fewer than twenty attempts at a definition of the infrastructural paradigm shift that is sweeping across the Enterprise IT world — some of them really quite good. From the article: 'Cloud computing is…the user-friendly version of grid computing.' (Trevor Doerksen) and 'Cloud computing really is accessing resources and services needed to perform functions with dynamically changing needs. An application or service developer requests access from the cloud rather than a specific endpoint or named resource.' (Kevin Hartig)"
3. According to Gartner's Hype Cycle Special Report for 2009, "technologies at the 'Peak of Inflated Expectations' during 2009 include cloud computing, e-books… and Internet TV, while social software and Micro blogging sites…have tipped over the peak and will soon experience disillusionment among Enterprise users". Is cloud computing also heading for the trough of disillusionment?

Fig: Conceptual view of cloud Computing





Characteristics of Cloud Computing:-

ISSN 2319-524X

Cloud computing exhibits the following key characteristics:

- Agility improves with users' ability to re-provision technological infrastructure resources.
- Application programming interface (API) accessibility to software that enables machines to interact with cloud software in the same way the user interface facilitates interaction between humans and computers. Cloud computing systems typically use REST-based APIs.
- Cost is claimed to be reduced and in a public cloud delivery model capital expenditure is converted to operational expenditure. This is purported to 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). The e-FISCAL project's state of the art repository contains several articles looking into cost aspects in more detail, most of them concluding that costs savings depend on the type of activities supported and the type of infrastructure available in-house.
- Device and location independence enable users to access systems using a web browser regardless of their location or what 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.
- Virtualization technology allows servers and storage devices to be shared and utilization be increased. Applications can be easily migrated from one physical server to another.
- 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.
- Reliability is improved if multiple redundant sites are used, which makes well-designed cloud computing suitable for business continuity and disaster recovery.
- Scalability and elasticity 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.
- Security could improve 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 other 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 that 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.
- Maintenance of cloud computing applications is easier, because they do not need to be installed on each user's computer and can be accessed from different places.

Cloud Computing Models:-

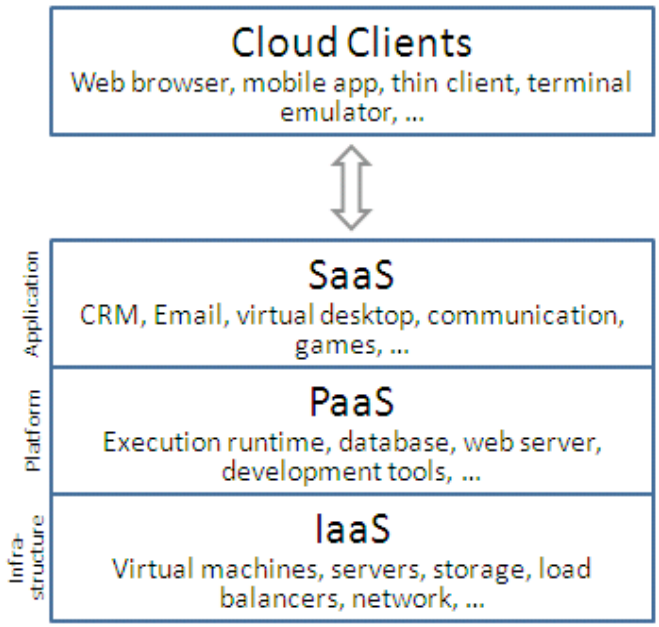
Cloud Providers offer services that can be grouped into three categories:-

1**. Software as a Service (SaaS):** In this model, a complete application is offered to the customer, as a service on demand. A single instance of the service runs on the cloud & multiple end users are serviced. On the customers side, there is no need for upfront investment in servers or software licenses, while for the provider, the costs are lowered, since only a single application needs to be hosted & maintained. Today SaaS is offered by companies such as Google, Sales force, Microsoft, Zoho, etc.

2. **Platform as a Service (Paas):** Here, a layer of software, or development environment is encapsulated & offered as a service, upon which other higher levels of service can be built. The customer has the freedom to build his own applications, which run on the provider˝s infrastructure. To meet manageability and scalability requirements of the applications, PaaS providers offer a predefined combination of OS and application servers, such as LAMP platform (Linux, Apache, MySql and PHP), restricted J2EE, Ruby etc. Google s App Engine, Force.com, etc are some of the popular PaaS examples.

3. **Infrastructure as a Service (Iaas):** IaaS provides basic storage and computing capabilities as standardized services over the network. Servers, storage systems, networking equipment, data centre space etc. are pooled and made available to handle workloads. The customer would typically deploy his own software on the infrastructure. Some common examples are Amazon, GoGrid, 3 Tera, etc.

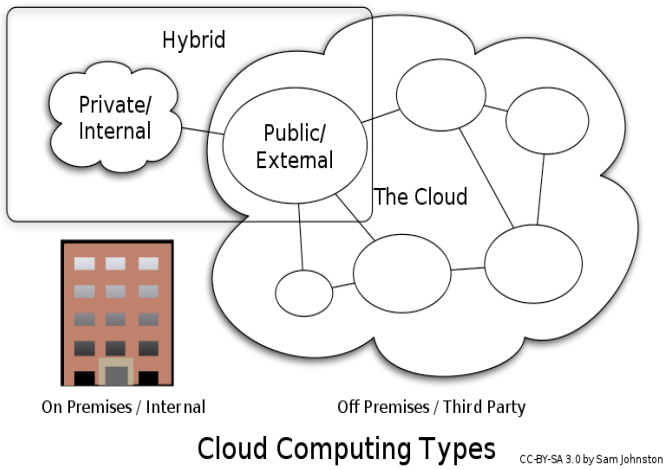
ISSN 2319-524X



Deployment Models: -

1) Public cloud

Public cloud applications, storage, and other resources are made available to the general public by a service provider. These services are free or offered on a pay-per-use model. Generally, public cloud service providers like Amazon AWS, Microsoft and Google own and operate the infrastructure and offer access only via Internet (direct connectivity is not offered).



2) Community cloud

Community cloud shares infrastructure between several organizations from a specific community with common concerns (security, compliance, jurisdiction, etc.), whether managed internally or by a third-party and hosted internally or externally. The costs are spread over fewer users than a public cloud (but more than a private cloud), so only some of the cost savings potential of cloud computing are realized.

3) Hybrid cloud

Hybrid cloud is a composition of two or more clouds (private, community or public) that remain unique entities but are bound together, offering the benefits of multiple deployment models.

By utilizing "hybrid cloud" architecture, companies and individuals are able to obtain degrees of fault tolerance combined with locally immediate usability without dependency on internet connectivity. Hybrid cloud architecture requires both on-premises resources and off-site (remote) server-based cloud infrastructure.

Hybrid clouds lack the flexibility, security and certainty of in-house applications. Hybrid cloud provides the flexibility of in house applications with the fault tolerance and scalability of cloud based services.

Private cloud

Private cloud is cloud infrastructure operated solely for a single organization, whether managed internally or by a third-party and hosted internally or externally. Undertaking a private cloud project requires a significant level and degree of engagement to virtualize the business environment, and it will require the organization to reevaluate decisions about existing resources. When it is done right, it can have a positive impact on a business, but every one of the steps in the project raises security issues that must be addressed in order to avoid serious vulnerabilities.

They have attracted criticism because users "still have to buy, build, and manage them" and thus do not benefit from less hands-on management, essentially "[lacking] the economic model that makes cloud computing such an intriguing concept".

Document is a template. An electronic copy can be downloaded from the conference website. For questions on paper guidelines, please contact the conference publications committee as indicated on the conference website. Information about final paper submission is available from the conference website.

Cloud Computing Benefits:-

Enterprises would need to align their applications, so as to exploit the architecture models
That Cloud Computing offers. Some of the typical benefits are listed below:

1. **Reduced Cost**

There are a number of reasons to attribute Cloud technology with lower costs. The billing model is pay as per usage; the infrastructure is not purchased thus lowering maintenance. Initial expense and recurring expenses are much lower than traditional computing.

2. **Increased Storage**

With the massive Infrastructure that is offered by Cloud providers today, storage &maintenance of large volumes of data is a reality. Sudden workload spikes are also managed effectively & efficiently, since the cloud can scale dynamically.

ISSN 2319-524X

3. Flexibility

This is an extremely important characteristic. With enterprises having to adapt, even more rapidly, to changing business conditions, speed to deliver is critical. Cloud computing stresses on getting applications to market very quickly, by using the most appropriate building blocks necessary for deployment.

Cloud Computing Challenges:-

Despite its growing influence, concerns regarding cloud computing still remain. In our opinion, the benefits outweigh the drawbacks and the model is worth exploring. Some common challenges are:

1. **Data Protection:** Data Security is a crucial element that warrants scrutiny. Enterprises are reluctant to buy an assurance of business data security from vendors. They fear losing data to competition and the data confidentiality of consumers. In many instances, the actual storage location is not disclosed, adding onto the security concerns of enterprises. In the existing models, firewalls across data centers (owned by enterprises) protect this sensitive information. In the cloud model, Service providers are responsible for maintaining data security and enterprises would have to rely on them.

2. Data Recovery and Availability: All business applications have Service level agreements that are stringently followed. Operational teams play a key role in management of service level agreements and runtime governance of applications. In production environments, operational teamsSupport Appropriate clustering and Fail over Data Replication System monitoring (Transactions monitoring, logs monitoring and others)
Maintenance (Runtime Governance) Disaster recovery Capacity and performance management If, any of the above mentioned services is under-served by a cloud provider, the damage & impact could be severe.

3. Management Capabilities: Despite there being multiple cloud providers, the management of platform and infrastructure is still in its infancy. Features like „Auto-scaling" for example, is a crucial requirement for many enterprises. There is huge potential to improve on the scalability and load balancing features provided today.

4. **Regulatory and Compliance Restrictions:** In some of the European countries, Government regulations do not allow customers personal information and other sensitive information to be physically located outside the state or country. In order to meet such requirements, cloud providers need to setup a data center or a storage site exclusively within the country to comply with regulations. Having such an infrastructure may not always be feasible and is a big challenge for cloud providers. With cloud computing, the action moves to the interface — that is, to the interface between service suppliers and multiple groups of service consumers. Cloud services will demand expertise in distributed services, procurement, risk assessment and service negotiation areas that many enterprises are only modestly equipped to handle.

5. Services are available from any location.

6. Cloud computing can be ordered online without detailed formal contracts.

Libraries and the clouds:-

Libraries are also not left blank with the emerging technology i.e. cloud computing. Many enterprises are implementing cloud computing as a new technology model for ICT infrastructures and services. cloud computing allows libraries to avoid locally hosting multiple servers and equipment and constantly deals with hardware failure, software installs, upgrades and compatibility issues and this is possible because processes get simplified and libraries save time and money. ICT related headaches of libraries get reduced by the introduction of this technology. Furthermore it can concurrently make workflows simpler and permit the libraries to make available improved end-user customer services with highly developed library find means and assist them online through a large network of collaborating librarians globally

Some live examples where Libraries are adopting Cloud Computing:-

I. OCLC: - OCLC Online Computer Library Center is a nonprofit, membership, computer library service and research organization dedicated to the public purposes of furthering access to the world's information and reducing the rate of rise of library costs. In a sense OCLC has been functioning as a cloud computing vendor. They provide cataloguing tools over the Internet and allow member institution to draw on their centralized data store. This centralized database allows for the sharing of catalog records between libraries and greatly reduces the time spent in cataloging incoming material.
World Cat is another example of cloud computing architecture drawing on the union catalog infrastructure they have built up over the years.

Library Thing

It is one of the sites that combine aspects of social networking and Cloud computing is Library Thing, originator of which is Tim Spalding. Library Thing offers services which are just like social networking site, authorizes people to contribute information and suggestion about books and allows them to interconnect globally to share interests. This site also contributes web services for libraries after paying a nominal fee it allows libraries to draw on the vast database of recommendation and other users available in Library Thing.

Reed Elsevier

Reed Elsevier is a service provider for scientific information working with hospitals to provide point in time information to medical technicians as they need the information. It is capitalizing on the cloud computing model. There is the possibility to place monographic and article content or even technical manuals so that technician and other medical

ISSN 2319-524X

personnel can get assistance exactly when they need it. This utilizes the cloud computing model in the way that computers and other devices used in the medical profession can be tied into the data and application provided by Elsevier from anywhere.

Amazon and Google

These are among the leading enterprises also providing solutions for libraries by having partnerships between library automation vendors. Amazon has been developing for years a large web services architecture and they now offer hosting services for data which are priced at gigabytes-month and CPU hour rates. We basically pay what we actually use.

Google for years is working for the dissemination of information also taking interest in library solutions, going to implement "APP Engine" which provides a hosted service for application within their server farms and on massive and highly redundant storage system. IBM are showing curiosity in the world and has begun developing an infrastructure known by the name "Blue Cloud".

Kindle and Mobile Me services

In the electronic book arena Amazon is providing some reading services with „Kindle". If one have wireless connection, can purchase and read a rapidly growing list of books and periodicals from the Kindle no matter for the location. With this service largest text can be downloaded in seconds

Another like service is "Mobile Me" provided by Apple computing. The concept is distributed calendaring and messaging no matter what device you are using. Modifications made via one device are instantly reflected on all of the devices and computers that are tied into "Mobile Me". This has many applications in the library world foe e.g. with the library acting as the gatekeepers, institutions could provide mobile access to say, a list of articles to their students simply by selecting them and giving them a code which would bring up the lists of articles from a vendors" cloud. The same cloud works for preprint archives, data archives and digital object repositories. (Fox 2009)

DuraSpace

A hosted service and open technology to help organizations and end users effectively utilize public cloud services. Built upon existing cloud services. The service can work on Amazon, Atmos, Sun, Rackspace, and other cloud services.LOCKSS in the cloud based on DuraCloud.

Chronopolis Project – designed primarily as a preservation storage system. Chronopolis Tools also monitors files and does auditing. (http://surferblue.wordpress.com)

TerraPod – digital video library. It allows you to outsource upload and data creation to the creators of the content. (http://surferblue.wordpress.com)

Conclusion:-

Cloud computing is an entirely new form of computing which many enterprises such as Google, Yahoo, Microsoft, Amazon, Zoho, and sales force are adopting for infrastructure solutions. Cloud computing is attracting enterprises and now libraries due to number of reasons. The concept is that it shifts the bulk of the responsibility for infrastructure support out to another vendor and basically outsources all data centers and software support to a company that specializes in web-based computing. Cloud computing is a pay-per-use model to facilitate on-demand access to configurable and trustworthy resources through its architectural layers such as Infrastructure as a service (IaaS), Platform as a service (PaaS) and Software as a service (SaaS). Cloud computing have various benefits such as the reduced cost, ease of maintenance, sharing of resources, etc. It is just renting of services instead of buying them. Instead of having many advantages, several difficulties must be overcome for cloud computing to be used on a large scale; the first one is the standardization of services offered by cloud vendors. Another obstacle is the limited support to relational database offered by current cloud solutions; the later difficulty to be overcome is the privacy of data located in a cloud. Once all of these difficulties will be surmounted, cloud computing will have the possibilities to be a massively used paradigm. (Giordanelli & Mastroianni 2010).

References:-

1. "A network 70 is shown schematically as a cloud", US Patent 5,485,455, column 17, line 22, filed Jan 28, 1994
2. "the cloud indicated at 49 in Fig. 1.", US Patent 5,790,548, column 5 lines 56-57, filed April 18, 1996
3. "July, 1993 meeting report from the IP over ATM working group of the IETF". CH: Switch. Retrieved 2010-08-22.
4. "Jeff Bezos' Risky Bet", *Business Week*
5. "Amazon's early efforts at cloud computing partly accidental", *IT Knowledge Exchange*, Tech Target, 2010-6-17
6. B Rochwerger, J Caceres, RS Montero, D Breitgand, E Elmroth, A Galis, E Levy, IM Llorente, K Nagin, Y Wolfsthal, E Elmroth, J Caceres, M Ben-Yehuda, W Emmerich, F Galan. "The RESERVOIR Model and Architecture for Open Federated Cloud Computing", IBM Journal of Research and Development, Vol. 53, No. 4. (2009)
7. D Kyriazis, A Menychtas, G Kousiouris, K Oberle, T Voith, M Boniface, E Oliveros, T Cucinotta, S Berger, "A Real-time Service Oriented Infrastructure", International Conference on Real-Time and Embedded Systems (RTES 2010), Singapore, November 2010
8. Keep an eye on cloud computing, Amy Schurr, Network World, 2008-07-08, citing the Gartner report, "Cloud Computing Confusion Leads to Opportunity". Retrieved 2009-09-11.
9. Amies, Alex; Sluiman, Harm; Liu, Guo Ning (July 2012)."Infrastructure as a Service Cloud Concepts". *Developing and Hosting Applications on the Cloud*. IBM Press.
10. Gartner Says Worldwide IT Spending On Pace to Surpass Trillion in 2008, Gartner, 2008-08-18. Retrieved 2009-09-11.
11. "Launch of IBM Smarter Computing". Retrieved 1 March 2011.
12. "What's In A Name? Utility vs. Cloud vs Grid". Datacenterknowledge.com. Retrieved 2010-08-22.
13. "Distributed Application Architecture". Sun Microsystem. Retrieved 2009-06-16.
14. "Sun CTO: Cloud computing is like the mainframe". Itknowledgeexchange.techtarget.com. 2009-03-11. Retrieved 2010-08-22.
15. "It's probable that you've misunderstood 'Cloud Computing' until now". TechPluto. Retrieved 2010-09-14.

ISSN 2319-524X

16. Danielson, Krissi (2008-03-26). "Distinguishing Cloud Computing from Utility Computing". Ebizq.net. Retrieved 2010-08-22.
17. "Recession Is Good For Cloud Computing – Microsoft Agrees". CloudAve. Retrieved 2010-08-22.
18. "Defining "Cloud Services" and "Cloud Computing"". IDC. 2008-09-23. Retrieved 2010-08-22.
19. E-FISCAL project state of the art repository".
20. Farber, Dan (2008-06-25). "The new geek chic: Data centers". CNET News. Retrieved 2010-08-22.
21. King, Rachael (2008-08-04). "Cloud Computing: Small Companies Take Flight". Businessweek. Retrieved 2010-08-22.
22. "Defining and Measuring Cloud Elasticity". KIT Software Quality Departement. Retrieved 13 August 2011.
23. "Economies of Cloud Scale Infrastructure". Cloud Slam 2011. Retrieved 13 May 2011.
24. "Encrypted Storage and Key Management for the cloud". Cryptoclarity.com. 2009-07-30. Retrieved 2010-08-22.
25. Mills, Elinor (2009-01-27). "Cloud computing security forecast: Clear skies". CNET News. Retrieved 2010-08-22.
26. Stevens, Alan (June 29, 2011). "When hybrid clouds are a mixed blessing". *The Register*. Retrieved March 28, 2012.
27. "Is a Private Cloud Really More Secure?". Dell.com. Retrieved 07-11-12.
28. Foley, John. "Private Clouds Take Shape". InformationWeek. Retrieved 2010-08-22.
29. Haff, Gordon (2009-01-27). "Just don't call them private clouds". CNET News. Retrieved 2010-08-22.
30. "There's No Such Thing As A Private Cloud". InformationWeek. 2010-06-30. Retrieved 2010-08-22.
31. http://linuxsysadminblog.com/2009/02/best-definition-of-cloud-computing-to-date/
32. http://www.thbs.com/pdfs/Cloud-Computing-Overview.pdf
33. http://www.luitinfotech.com/kc/what-is-cloud-computing.pdf
34. http://en.wikipedia.org/wiki/Cloud_computing