academicJournals

Vol. 5(10), pp. 394-400, November, 2013
DOI: 10.5897/IJLIS2013.0367
ISSN 2141–2537 ©2013 Academic Journals
http://www.academicjournals.org/IJLIS

International Journal of Library and Information Science

*Full Length Research Paper*

# Development of cloud computing in integrated library management and retrieval system

**DAS Amit Kumar* and Sukumar Mandal**

Department of Library and Information of Science, Golapbag, The University of Burdwan, Burdwan, W.B., India.

Accepted 5 July, 2013

**Cloud computing comes in several different forms and this research paper discusses how service, platform and infrastructure forms of cloud computing have been used to serve library needs. Following an overview of these uses the experience of one library in migrating IT infrastructure to a cloud environment and concludes with a model for assessing cloud computing. Development of cloud computing for integrated library management and retrieval system on the basis of global recommendation include Service Oriented Architecture (SOA), Open Library Environment Project (OLE) and the recommendation of Integrated Library System for Discovery Interface (ILS-DI); and in this regard, the most comprehensive open source software for designing and developing the Internet based services is selected.**

**Key words:** Cloud computing, information mash-up, Web 2.0 and open source software.

## INTRODUCTION

The information retrieval environment has changed significantly over the past few years due to the convergence of computer, telecommunication and broadcasting technologies. The emergence of low cost electronic networks has paved the way for users to communicate daily with others around the world fast and inexpensively. In simple terms, the definition of an electronic network is two or more computers, connected in some way or the other, that are able to communicate and share information. These computers may be connected through cables, telephone lines via modems, dedicated data lines, wireless radio or satellite links. All kinds of information can be sent through electronic networks, such as text, databases, pictures, video, sound and so on. Electronic network is classified into categories according to the physical locations of its computers and terminals. The traditional approach to access and use remote information is changing rapidly, and traditional online information service providers are quickly adapting to the Web technology. User studies have always remained a critical area, and user-oriented approach to design, development and evaluation of information systems is becoming more and more important in the digital library environment. Intelligent search agents are being designed to perform a number of operations in the digital information environment and in order to solve the existing problems information professionals through open source software with their rich experience in organizing and searching information as well as in identifying user needs, should work hand in hand with software professionals in organizing the so-called 'anarchy' of the Internet.

### Cloud computing

Cloud computing is the use of computing resources (hardware and software) that are delivered as a service over a searching information as well as in identifying users' needs should work hand in hand with software professionals in organizing the so-called 'anarchy' of the Internet network (typically the Internet). The name comes from the use of a cloud- shaped symbol as an abstraction

*Corresponding author. E-mail: zahidrais@gmail.com, zahid@kashmiruniversity.ac.in.

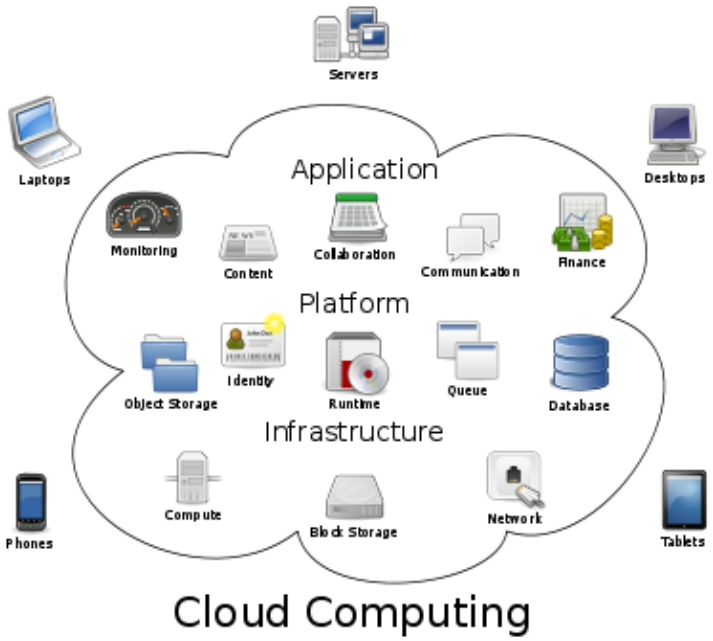


**Figure 1.** Cloud computing logical diagram.

for the complex infrastructure it contains in system diagrams (Figure 1). Cloud computing entrusts remote services with a user's data, software and computation (NIST, 2012). 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 (Lewis, 2010). The word *cloud* is used as a metaphor for the Internet, based 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 (Lewis, 2009). The cloud symbol was used to represent the Internet as early as 1994. In the 1990s, telecommunications companies that 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. There are many types of public cloud computing (Monaco, 2012) like Infrastructure as a service (IaaS), Platform as a service (PaaS), Software as a service (SaaS), Storage as a service (StaaS), Security as a service (SECaaS), Data as a service (DaaS), Test environment as a service (TeaaS), Desktop as a service (DaaS) and API as a service (APIaaS). Cloud computing is a technology that uses the internet and central remote servers to maintain data and applications. Cloud computing allows consumers and businesses to use applications without installation and access their personal files at any computer with internet access. This technology allows for much more efficient computing by centralizing data storage, processing and bandwidth. A simple example of cloud computing is Yahoo email, Gmail, or Hotmail etc. All you need is just an internet connection and you can start sending emails. The server and email management software is all on the cloud (internet) and is totally managed by the cloud service provider Yahoo, Google etc.

Essential characteristics of cloud computing are discussed as follows.

### On-demand self-service

A consumer can unilaterally provide computing capabilities, such as server time and network storage, as needed automatically without requiring human interaction with each service provider.

### Broad network access

Capabilities are available over the network and accessed through standard mechanisms that promote use by heterogeneous thin or thick client platforms (examples, mobile phones, tablets, laptops and workstations).

### Resource pooling

The provider’s computing resources are pooled to serve

multiple consumers using a multi-tenant model, with different physical and virtual resources dynamically assigned and reassigned according to consumer demand. There is a sense of location independence in that the customer generally has no control or knowledge over the exact location of the provided resources but may be able to specify location at a higher level of abstraction (examples, country, state, or datacenter). Examples of resources include storage, processing, memory and network bandwidth.

### Rapid elasticity

Capabilities can be elastically provided and released; in some cases automatically, to scale rapidly outward and inward commensurate with demand. To the consumer, the capabilities available for provision often appear to be unlimited and can be appropriated in any quantity at any time.

### Measured service

Cloud systems automatically control and optimize resource use by leveraging a metering capability1 at some level of abstraction appropriate to the type of service (examples, storage, processing, bandwidth and active user accounts). Resource usage can be monitored, controlled and reported, providing transparency for both the provider and consumer of the utilized service.

## METHODOLOGY OF DEVELOPMENT OF CLOUD COMPUTING IN INTEGRATED LIBRARY SYSTEM (ILS)

Development of cloud computing is an important task towards next level integrated library system. Now, cloud computing is of two parts: front end and back end. The front end includes the client's computer (or computer network) and the application required to access the cloud computing system. Not all cloud computing systems have the same user interface. Services like Web-based e-mail programs leverage existing Web browsers like Internet Explorer or Firefox (Dormann and Rafail, 2006). Other systems have unique applications that provide network access to clients. On the back end of the system are the various computers, servers and data storage systems that create the "cloud" of computing services.

In theory, a cloud computing system could include practically any computer program you can imagine, from data processing to video games. Usually, each application will have its own dedicated server. A central server administers the system, monitoring traffic and client demands to ensure everything runs smoothly. It follows a set of rules called protocols and uses a special kind of software called middleware. Middleware allows networked computers to communicate with each other (Jansen and Grance, 2011). Most of the time, servers do not run at full capacity. That means there is unused processing power going to waste. It is possible to fool a physical server into thinking it is actually multiple servers, each running with its own independent operating system. The technique is called server virtualization (Strowd and Lewis, 2010). By maximizing the output of individual servers, server virtualization reduces the need for more physical machines. There are many open source software in integrated library system like Koha, NewGenLib, Emilda, OpenBiblio and PhpMyLibrary and etc. Among these web enabled software, Koha and NewGenLib are tested by twelve checklists on cloud computing and compare with these two software which support cloud computing for identification, collection, organization and dissemination of digital resources for the users as well as library professionals.

## CLOUD COMPUTING SERVICES

Cloud computing providers offer their services according to three fundamental models: Infrastructure as a service (IaaS), platform as a service (PaaS), and software as a service (SaaS) where IaaS is the most basic and each higher model abstracts from the details of the lower models (Voorsluys et al., 2011).

### Infrastructure as a service (IaaS)

In this most basic cloud service model, cloud providers offer computers, as physical or more often as virtual machines, and other resources. The virtual machines are run as guests by a hypervisor, such as Xen or KVM. Management of pools of hypervisors by the cloud operational support system leads to the ability to scale to support a large number of virtual machines (Amies et al., 2012). Other resources in IaaS clouds include images in a virtual machine image library, raw (block) and file-based storage, firewalls, load balancers, IP addresses, virtual local area networks (VLANs), and software bundles.

### Platform as a service (PaaS)

In the PaaS model, cloud providers deliver a computing platform typically including operating system, programming language execution environment, database, and web server. Application developers can develop and run their software solutions on a cloud platform without the cost and complexity of buying and managing the underlying hardware and software layers. With some PaaS offers, the underlying computer and storage resources scale automatically to match application demand such that cloud user does not have to allocate resources manually.

### Software as a service (SaaS)

In this model, cloud providers install and operate application software in the cloud and cloud users access the software from cloud clients. The cloud users do not manage the cloud infrastructure and platform on which the application is running. This eliminates the need to install and run the application on the cloud user's own computers simplifying maintenance and support (Hamdaqa, 2012). What makes a cloud application different from other applications is its elasticity. This can be achieved by cloning tasks onto multiple virtual machines at run-time to meet the changing work demand.

These services are used in integrated library management and retrieval system for cloud computing. The open source software and open standards are also used for Internet based services towards next automated library system. The most important cloud computing based services are provided for accessing journals, e-books in library OPAC as well as librarian interface and these can be achieved in different way such as information mashup, import bibliographic and authority data, reference management and web 2.0 by using twelve check list. These important services are described in the following way.

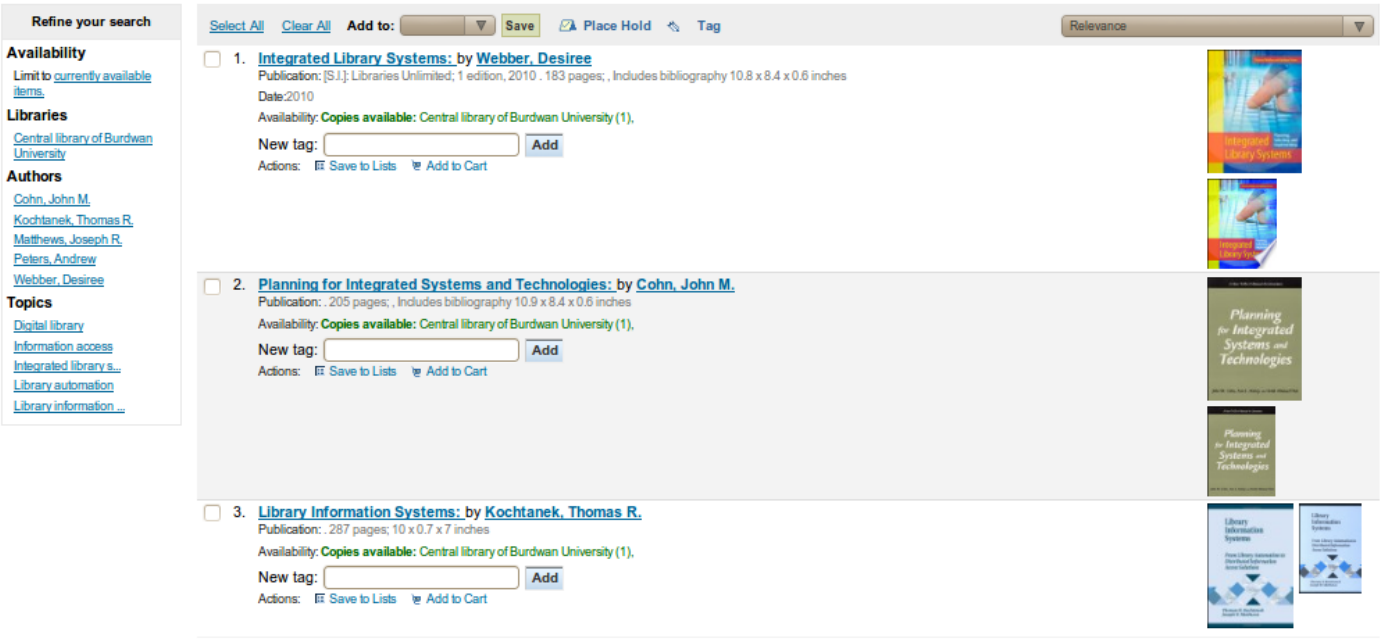


**Figure 2.** Information Mashup.

### Information Mashup

A mashup in web development is a web page or web application that uses and combines data, presentation or functionality from two or more sources to create new services. The term implies easy, fast integration, frequently using open application programming interfaces (API) and data sources to produce enriched results that were not necessarily the original reason for producing the raw source data. Information mashup is a collection of structured and unstructured contents combined in a single, web-based location so they can be browsed and searched as a single body of information. The book covers images and links that are snatched through open source software like koha in OPAC interfaces. This will attract the user to choose the necessary document which is available on-line. Koha is a web enabled architecture (Figure 2) and this is represented in user interfaces.

### Import Bibliographic and Authority Data

In integrated library system, bibliographic and authority data are an essential element for staff as well as user interfaces. There are many tools and servers available for importing the authority and bibliographic records from other library OPAC. It saves time and can easily manage the MARC tag (fields and subfields) except in tag 9. Authority data are snatched through authority control tools like Name Authority Cooperative Cataloguing (NACO), Subject Authority Cooperative Cataloguing (SACO), Virtual International Authority File (VIAF), Library of Congress (LoC) etc; among these tools library of congress is the most comprehensive tool compared to other authority control tools. Library of congress also manages the bibliographic records and these records import though client server tools like Z39.50 server (Figure 3). Integration of bibliographic and authority control tools is easily possible in koha both for admin as well as user interfaces. Bibliographic and authority data import is possible if computer is connected to Internet.

## REFERENCE MANAGEMENT

Reference management entails the collection, annotation and citation of published information. Zotero greatly simplifies these tasks and it can perform Internet searches to retrieve references and associated pdf format or web pages, or immediately find and import references for which you already have the pdf format. Versatile groups (static, smart, and virtual), Term Lists (keywords, authors, etc.), reference cross-linking (links), multiple note cards per reference, tag clouds, and near-instantaneous live searches help organize and find information of user personal collection. Highly con-figurable displays let you view the reference information and attachments the way you want. To create publication-ready documents complete with bibliographies and footnotes, scan your word processor manuscripts and this can be exported in open-office word for list of references like APA. Figure 4 represents the Zotero for managing the references both for web and local files (like bibtex) and this helps the researcher or advanced user for better management of references.

### Twelve check list

This research work compiled a list of commonly acknowledged features for next-generation catalogs found in the library literature and summarized in Marshall Breeding's Introduction in Library Technology Reports (Breeding, 2007) and Peter Murray's Power Point presentation on OPAC discovery layer tools (Murray, 2008).

### Single point of entry for all library information

The library catalog should be a single search or federated search for all library materials, including pointers to the articles in electronic databases as well as records of books and digital collections. One search should retrieve all relevant materials. Presently, patrons have to search the catalog for books and videos, databases for

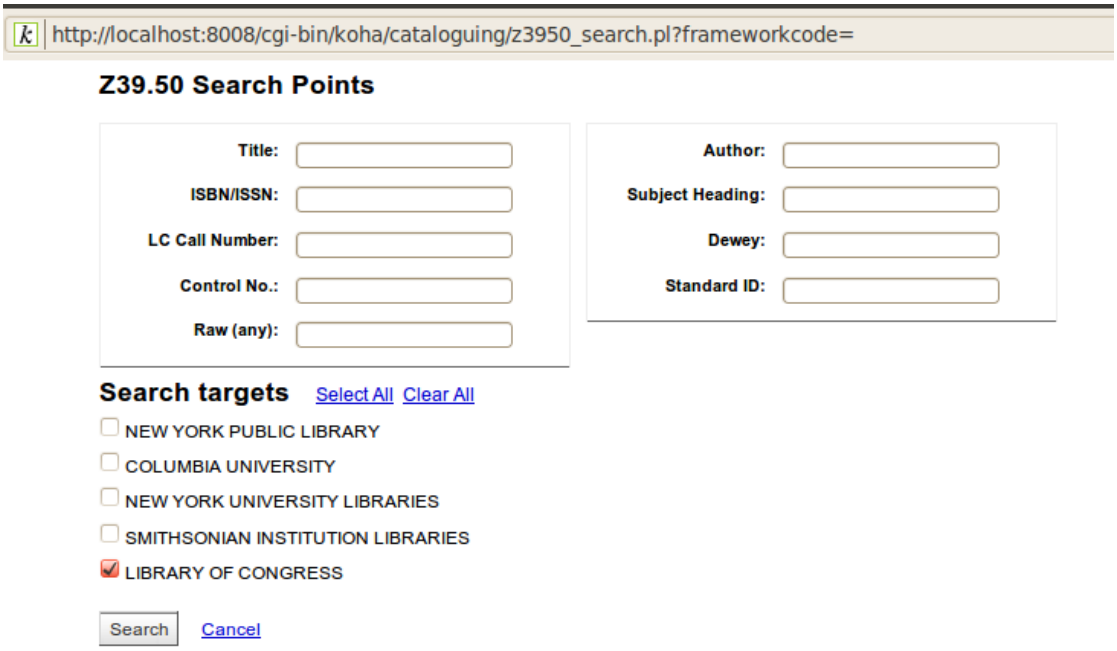


**Figure 3.** Z39.50 search in Koha.

journal articles, and digital collections and archives for local images and materials.

### State-of-the-art web interface

Library catalogs should have a modern design similar to commercial, e-business sites. This criterion is highly subjective and as such is difficult to quantify. A next-generation catalog should look and feel like popular sites such as Google, Netflix and Amazon.

### Enriched content

Library catalogs should include book cover images, user driven input such as comments, descriptions, ratings, and tag clouds. Traditionally, only professionally trained cataloging librarians have the ability to create or add content to bibliographical records.

### Faceted navigation

Library catalogs should be able to display the search results as sets of categories based on some criteria such as dates, languages, availability, formats, locations, etc. Users can conduct a very simple, initial search by their preferred keyword method and then refine their results by clicking on the various results facets.

### Simple keyword search box on every page

The next generation catalog starts with a simple keyword search box that looks like that of Google or Amazon. A link to advanced search should be provided. The simple search box should appear on every page of the interface as users navigate and conduct searches.

### Relevancy

Librarians complain that OPAC relevancy results are problematic or that they do not understand how relevance is determined. The next-generation catalog does better in relevancy ranking with increased precision. In addition circulation statistics should influence the relevancy results. More frequently circulated books indicate popularity and usefulness. They should be ranked higher in the display. Items deemed important enough to have multiple copies should also receive higher relevancy ranking.

### Did you mean...?

A spell-checking mechanism should be present in a next-generation catalog. When an error appears in the search, there should be a pop-up with the correct spelling or suggestions from a dictionary.

Clicking on any of these runs a search terms.

### Recommendations/related materials

Commonplace in e-commerce sites, the customer is shown additional items with a suggestion like “Customers who bought this item also bought...” Likewise, a next-generation catalog should recommend books for readers on transaction logs. This should take the form of “Readers who borrowed this book also borrowed the following... ” or a link to “Recommended Readings”.

### Users’ contribution

The next-generation catalog allows users to add data to records. The user input includes descriptions, summaries, reviews, criticism, comments, rating and ranking, and tagging or folksonomies.

Today’s users increasingly look for what other users have to say

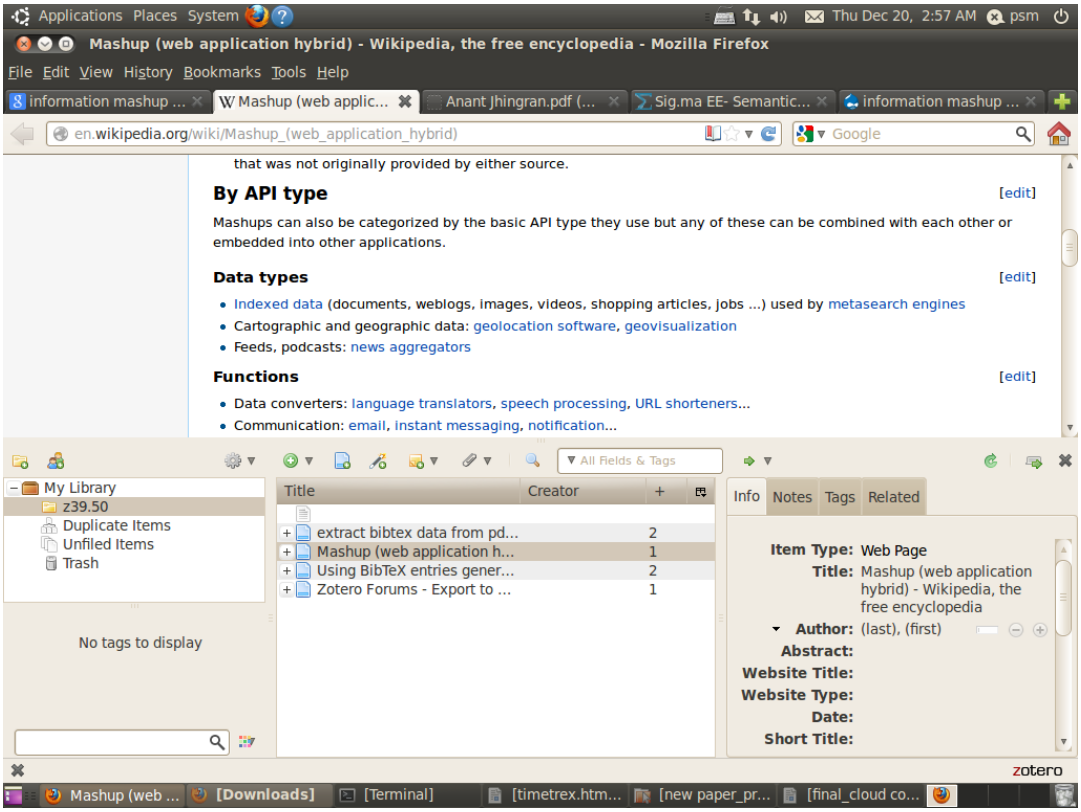


**Figure 4.** Reference management on Zotero.

**Table 1.** Twelve check list.

| No. | Check list | Koha | | NewGenLib | |
|---|---|---|---|---|---|
| | | Support | Score | Support | Score |
| 1 | Single point of entry for all library information | Yes | 1 | Yes | 1 |
| 2 | State-of-the-art web interface | Yes | 1 | Yes | 1 |
| 3 | Enriched content | Yes | 1 | Yes | 1 |
| 4 | Faceted navigation | Yes | 1 | Partial | 0.5 |
| 5 | Simple keyword search box on every page | Yes | 1 | Yes | 1 |
| 6 | Relevancy | Yes | 1 | Yes | 1 |
| 7 | Did you mean . . . ? | Yes | 1 | No | 0 |
| 8 | Recommendations/related materials | Yes | 1 | No | 0 |
| 9 | User contribution | Yes | 1 | Yes | 1 |
| 10 | RSS feeds | Yes | 1 | Yes | 1 |
| 11 | Integration with social network sites | Yes | 1 | Partial | 0.5 |
| 12 | Persistent links | Yes | 1 | No | 0 |

Total Score (Out of 12); Koha Score: 12; NewGenLib Score: 8.

about items found online, and value what they feel to be their peers' review of items. Tagging clouds can serve as access points and descriptive keywords leading to frequently used items.

### RSS feeds

Really Simple Syndication allows users to connect themselves to content that is often updated. Next-generation interfaces include RSS feeds so that users can have new book lists, top-circulating book lists, canned searches, and "watch this topic" connections to the catalog on their own blog or feed reader page.

### Integration with social network sites

When a library's catalog is integrated with social network sites,

patrons can share links to library items with their friends on social networks like Twitter, Facebook and Delicious.

#### Persistent links

Next-generation catalog records contain a stable URL capable of being copied and pasted and serving as a permanent link to that record.

This check list represented in Table 1 supports online public access catalog in library automation of two open source software like Koha and NewGenLib.

From Table, 1 represents the twelve check lists in library OPAC which is essential and inevitable in colleges, universities and special libraries. Twelve check lists in library online public access catalog like Single point of entry for all library information, State-of-the-art web interface, Enriched content, Faceted navigation, Simple keyword search box on every page, Relevancy, Did you mean...?, Recommendations/related materials, User contribution, RSS feeds, integration with social network sites and Persistent links. These are the new innovative features in library automation or next generation cataloguing in library OPAC. It helps the user and librarian with web 2.0 features to identify their necessary materials and resources with minimum time. So it can save the time of the reader by using or applying the twelve check lists in library OPAC interface. All features are available in koha software with OPAC interface while NewGenLib does not fulfill all the features in library OPAC. The score of koha is 12 out of 12 where NewGenLib scored 8 out of 12. Obviously, it can be concluded that Koha is the most efficient and comprehensive software in library OPAC compared to NewGenLib on the basis of cloud computing checklist.

## FINDINGS AND OBSERVATIONS

The findings of cloud computing for integrated library management and retrieval system are described in the following way:

a) There is location independence as long as there is access to the Internet and web 2.0 features by twelve check list through open source software like Koha.
b) Faceted navigation, Did you mean...? Recommendations/related materials, and Persistent links: these type of cloud computing facilities are available in Koha but in NewGenLib, improvement is needed.

i) Koha and NewGenLib both are web-enabled architecture for managing the web resources both for bibliographic and authority data.
ii) Reference management is possible through Zotero (Open source software) in both integrated library management system because it is connected to Firefox add-in.
iii) Increased flexibility and market agility as the quick deployment model of cloud computing increases the ability to re-provision rapidly as required.
iv) Increased security at a much lesser cost as compared to traditional standalone applications due to centralization of data and increased security – focused resources.

## Conclusion

Colleges and universities are expected to provide a wide and growing array of technology services in integrated library management system; some of which are highly specialized or idiosyncratic to individual campuses, whereas others simply need to be available. By offering commodity services over the Internet, cloud computing offers one way for institutions to increase operational efficiency and focus scarce resources on services that are institutional differentiators. Operating in a cloud environment requires IT leaders and staff to develop different skills, such as managing contracts, overseeing integration between in-house and outsourced services, and mastering a different model of IT budgets. Cloud services might facilitate inter-institutional collaboration because they are more easily accessed by students and faculty at disparate institutions. In addition, despite the potential security risks posed by cloud services, some would argue that cloud services offer more security than on-campus solutions, given the complexity of mounting an effective IT security effort at the institutional level.

#### REFERENCES

Amies A, Sluiman H, Tong QG, Liu GN (July 2012). "Infrastructure as a Service Cloud Concepts". Developing and Hosting Applications on the Cloud. IBM Press. ISBN 978-0-13-306684-5. http://www.ibmpressbooks.com/bookstore/product.asp?isbn=_9780133066845. (Accessed on Nov 12, 2012).

Dormann W, Rafail J (2006). Securing Your Web Browser. Retrieved from http://www. cert.org /tech_tips /securin gbrowser/.(Accessed on Nov 20, 2012).

Hamdaqa M (2012). A Reference Model for Developing Cloud Applications. Retrieved from http://en.wikipedia.org/wiki/Cloud_computing (Accessed on Sept 27, 2012).

Jansen W, Grance T (2011). Guidelines on Security and Privacy in Public Cloud Computing. National Institute of Standards and Technology.

Lewis G (2009). Cloud Computing: Finding the Silver Lining, Not the Silver Bullet. Retrieved from http://www.sei.cmu.edu/newsitems/cloudcomputing.cfm. (Accessed on Nov 25, 2012).

Lewis G (2010). Basics About Cloud Computing. Retrieved from http://www.sei.cmu.edu/library/abstracts/whitepap ers/cloudcomputingbasics.cfm. (Accessed on Nov 25, 2012).

Monaco A (7 June 2012 [last update]). "A View inside the Cloud". theinstitute.ieee.org (IEEE). Retrieved from http://theinstitute.ieee.org/technology-focus/technology-topic/a-view-inside-the-cloud (Accessed on Nov 17, 2012).

NIST (2012). "The NIST Definition of Cloud Computing". National Institute of Science and Technology. Retrieved from http://csrc.nist.gov/publications/nistpubs/800-145/SP800-145.pdf. (Accessed on Sept 25, 2012).

Strowd H, Lewis, G (2010). T-Check in System-of-Systems Technologies: Cloud Computing (CMU/SEI-2010-TN-009). Software Engineering Institute, Carnegie Mellon University. Retrieved from http://www.sei.cmu.edu/library/abstracts/reports/10tn009.cfm (Accessed on Nov 25, 2012).

Voorsluys W, Broberg J, Buyya R (February 2011). "Introduction to Cloud Computing". In R. Buyya J, Broberg A, Goscinski (2012). Cloud Computing: Principles and Paradigms. New York, USA: Wiley Press. pp.1-44. ISBN978-0-470-88799-8. http://media.johnwiley.com.au/product_data/excerpt/90/04708879/0470887990-180.pdf. (Accessed on Nov 2, 2012).