

# Winds of change: libraries and cloud computing



**Matt Goldner, product and technology advocate for OCLC, examines cloud computing, a new technology model for IT services. It allows organisations to avoid locally hosting multiple servers and equipment and having to constantly deal with hardware failure, software installs, upgrades and compatibility issues. Cloud computing can simplify processes**

**and save time and money. This article defines cloud computing and shows how it is different from other types of computing. It also discusses how cloud computing solutions could be beneficial to libraries in three basic areas: technology, data and community**

Anyone connected to the internet is probably using some type of cloud computing on a regular basis. Whether they are using Google's Gmail, organising photos on Flickr or searching the web with Bing, they are engaged in cloud computing. As Geoffrey Moore points out, the interesting thing about cloud computing is that it did not start as a technology for the businesses, but was driven by the public with services like Facebook and Flickr<sup>i</sup>. However, over the last few years businesses have started to see the value of cloud computing, driving it to become a major technology solution for organisations around the world, including libraries.

## What is cloud computing?

The Gartner Group defines cloud computing as "a style of computing in which massively scalable and elastic IT-enabled capabilities are delivered as a service to external customers using Internet technologies"<sup>ii</sup>. In various presentations KPMG breaks this into four different types of cloud computing: infrastructure, platform, applications and services:

The above table illustrates varying definitions of cloud

Type	What is it?	Examples
Infrastructure	Buying time/space on external servers	Amazon A3 Bungee
Platform	An existing software platform to build your own applications on	Facebook
Applications	Software applications accessed with a web browser	Google Docs Salesforce.com
Services	Ready to use services accessed with a web browser	ADP

computing. Many cloud services actually incorporate two or more of these types: Google Docs provide infrastructure as well as applications. In addition, many cloud applications and services are actually using another provider's cloud infrastructure to run their service.

## How is cloud computing different?

For much of the past 25 years, software development and system engineering has centred primarily on the personal computer. The PC era was characterised by monolithic, proprietary operating systems and programs that had long development times and release cycles. The design of software was isolated and all attention focused on a single application.

With cloud computing, hardware and functionality traditionally installed and run in a local environment is now performed on the network, in the internet cloud. In essence, the internet cloud becomes the development platform and the operating system to which programmers write reusable, constantly updated software components that are delivered over the network and that can be embedded or loosely coupled with other web applications.

Libraries have been using some cloud computing services for over a decade: online databases are accessed as cloud applications and large union catalogues can also be defined

as cloud applications. However, looking outside libraries will allow a better understanding of the value proposition of cloud computing.

## Why are organisations adopting cloud computing solutions?

Jeff Bezos of Amazon has repeatedly spoken of the 70/30 rule whereby businesses which run applications spend 70 percent of their time and money supporting the infrastructure required to keep their business going<sup>iii</sup>. This only leaves them 30 percent of time and money to work on innovation and ways to improve and grow their business. He demonstrates that when a business moves their core applications to a cloud-based solution, they can invert this ratio thus giving them 70 percent of their time and money to improve and grow their business. The director of a large academic research library commented recently that she wished they

## cloud computing can help libraries to create innovative new services & collections for users

were only expending 70 percent of the time and money on infrastructure.

Minnesota Online High School supports all its courses over the internet for its students. They were collapsing under the weight of supporting so many different computers, associated CDs and licensed software as well as applications... now the school work, applications and data are all stored in the cloud. This enabled the school to switch their efforts from managing technology to focusing on education.

## What can cloud computing solutions do for libraries?

Cloud computing can solve some real problems for libraries such as amplifying the power of cooperation and building a significant, unified presence on the web whilst saving time, money and simplifying workflows.

- Potential areas for improvement could include:
  1. Most library computer systems are built on pre-web technology
  2. Systems distributed across the net using pre-web technology are harder and more costly to integrate
  3. Libraries store and maintain much of the same data hundreds and thousands of times
  4. Library data is scattered across distributed systems, weakening the library's web presence
  5. Collaboration between libraries is difficult and expensive with independent systems
  6. Information seekers work in common web environments and distributed systems make it difficult to get the library into their workflow
  7. Many systems are only used to 10 percent of their capacity. Combining systems in a cloud environment reduces the carbon footprints, making libraries greener

## Technology improvements

The mainstay of libraries is the library management system (LMS) or integrated library system (ILS). LMSs were developed before the internet and web existed, and are generally closed proprietary systems.

Over time libraries have had to add more systems to manage their changing collections which now comprise a combination of physical, licensed and digital collections. Each of these systems typically stands alone, making integrating them difficult and relying on their vendors to do any such integration. and, at times, making integration impossible. What can change in a cloud environment for managing core library services?

First is the possibility of open service oriented architecture (SOA). Many cloud solutions offer this type of openness with published application program interfaces (APIs) that any programmer can take advantage of. This means that if a new service or technology emerges libraries will not always be dependent on a vendor or other third party to start taking advantage of these services and technologies. Existing library systems have used APIs to connect to external services but they have remained closed proprietary systems, making it hard to integrate them into external services. When library



Image created using [www.wordle.net](http://www.wordle.net)

systems are deployed as open cloud solutions then the library community itself can create extensions to their core services and, more importantly, share them throughout the community using cloud solutions. This makes it possible to integrate two services once and reuse it across the community.

Secondly, libraries can focus on collection building, patron services and innovation rather than technology. Servers can be decommissioned and no longer require replacement every five years (or less). Staff no longer have to maintain the complex software stack necessary to run local systems and worry about compatibility of the stack during upgrades. Instead technical skills can be re-deployed for extending cloud services into their environment and vice versa.

## Data efficiencies

When data is stored in the cloud it offers several advantages. Common data can now be easily shared among services

and users. The need for local storage, maintenance and backups is removed. Agreements can be forged to share data that normally would be considered private to a single

cloud computing to open up collections so that so the public can tag, edit, collect and review them

organisation. Finally, libraries can achieve web scale when they massively aggregate data and users through the cloud.

A key advantage of data storage in the cloud is removing the need to store the same data hundreds and thousands of times across libraries. Consider how many copies of the cataloguing data there are for serials such as the *Economist*. And if a change is needed to the data to keep it current, each library must perform that change. When this data is maintained in the cloud, maintenance and backup of this data is now done once and if a change is needed, one library



performs the change and all share it.

Another great benefit of data stored in the cloud is the opportunity for collaboration and cooperative intelligence. Libraries can agree to share pools of data for cooperative collection building, cooperative preservation or digitisation, cooperative sharing of materials, etc. With massively aggregated data, new services can be created such as recommender services based on a broad base of usage data.

However, when library data is widely distributed across systems it makes library web presence weak. Search engines can harvest from large data stores but collectives must work on search engine optimisation and ensuring that library collections appear more relevant to search engines thus are displayed higher in search results. This is a complex and ever-changing task that would be prohibitive for individual libraries to accomplish. Further, aggregated data can attract a much larger aggregation of users who interact with the data, add to it and re-use it. The result is every user adds benefit for every other user.

### Community power

Cloud computing offers libraries a unique opportunity to create an online information community network comprising

the internal community of libraries collaborating within a single institution and across institutions and the external community of libraries and information seekers. The value to libraries is the “network effect” that coming together in the cloud provides. The cooperative efforts of libraries creates scale savings and efficiencies, brings wider recognition for libraries, provides cooperative intelligence for better decision-making, and provides the platform on which libraries can innovate.

The first community cloud computing offers is external: social media. Organisations can build social communities around their services and participate in existing online communities such as Facebook or Twitter.

The internal community formed through the cloud offers new possibilities and efficiencies for current workflows. Starting with a single organisation, the simple task of collaboratively working on documents and maintaining version control either requires extensive manual processes between colleagues or a locally installed system to assist in collaboration and version control. Many librarians have discovered the power of services like Google Docs which allow them to easily share ongoing work whenever they want and wherever they are, to reduce the effort of working jointly.

The potential for collaboration between libraries is truly revolutionary in a cloud environment. When data and functions are shared in the cloud, libraries can make joint decisions on collection development, preservation, and digitisation, in real time. As demonstrated by OCLC's QuestionPoint virtual reference service and its 24/7 cooperative, a single library's ability to assist patrons is expanded beyond the constraint of its own walls and hours of operation to become a true cloud service (last year QuestionPoint logged its five millionth answer to a reference question).

### Real world examples of current library cloud solutions

To date, the main focus of libraries moving into the cloud has been discovery services, a mechanism to disclose their vast collections on the web. Though library OPACs attract existing patrons they are not integrated with most information seekers' common workflows. A first step for libraries has been

to start massively aggregating data about their collections into common pools. OCLC's WorldCat is now forty years old and pre-dates both the web and cloud computing. Other similar union catalogues have existed throughout the world most commonly supported by national libraries and large union catalogues, such as the National Library of Australia, the Bayerische Staatsbibliothek in Germany, and Bibsys in Norway. The advent of the web has allowed libraries to extend this original vision in new ways.

### the potential for collaboration between libraries is truly revolutionary in a cloud environment

Extending these services beyond traditional library collections is well-illustrated by the National Library of Australia's (NLA) Trove. It has used the web to combine the collections of Australian libraries with other important Australian and international collections and information sources such as Wikipedia and to open up much of this content so the public can tag, edit, collect and review it.

The explosion of digitisation projects in the last decade has driven this gathering of information in new directions with examples in addition to NLA's Trove: the Hathi trust is building a repository of digitised books and journals from major research libraries in the United States; OAISTER is a service started by the University of Michigan and now managed by OCLC which seeks to harvest all the major digital repositories around the world; and Europeana is gathering the digitised collections from Europe's galleries, libraries, archives and museums. What makes these aggregations and others like them important is their intent to allow their content to be mashed up into other services and reused.

Other benefits growing from massively aggregated data about collections is the ability to aggregate user opinion and use. LibraryThing is a good example of being able to build recommender services based on the aggregation of what thousands of people hold in their personal libraries.

However, there is no reason to extend cloud-based services only to libraries' end users. As Marshall Breeding points out,



- “We can’t let the current focus on front-end interfaces make us complacent about the software systems that we use to automate routine library functions.”<sup>iv</sup>

### Beyond library discovery services

It is here that libraries can look to gain new efficiencies both internally and among the entire library community. When library software suppliers create the user personas that will use their software, the focus is generally on external users but there are also many internal personas that need to take advantage of new technologies and web capabilities. Reference librarians are now able to better assist their patrons online and to build a large network of librarians globally who can answer specific questions and be available 24/7. Other beneficiaries could be:

- Acquisitions librarians managing increasingly diverse collections
- Cataloguing librarians seeking to describe an ever-increasing body of information and information sources the library is managing
- Serials librarians working to maintain control and access to collections spidered across the web
- Electronic resource librarians managing burgeoning collections, and ever-changing lists of vendors

The dramatic change in library collections often blurs the lines between traditional job roles in libraries. An acquisitions librarian probably also needs to manage licences for electronic materials as well as manage purchasing for multiple formats, often for the same item. They need to access information from suppliers, reviewers, local constituency and other staff in a unified manner. This begs for an open system deployed where it can easily be accessed by external systems and that pulls in data and services in from those systems.

Cloud computing solutions can create the new workflows needed by librarians through a cooperative platform for libraries to build on using four key principles:

- Openness: services and data are made available to support greater interoperability, not only within and between cloud services, but also with library-developed and third-party applications



Image: FreeDigitalPhotos.net

- Extensibility: the platform can easily accommodate the addition of new services and applications, developed either by the service provider or by members of the community;
- Data richness: a library can interact with and expose a wide variety of information about purchased, licensed, and digital content through this platform
- Collaboration: libraries can harness the collective power of the community of libraries to innovate and share solutions.

It is precisely this that the business world and social media have demonstrated can be done with cloud computing solutions. Libraries can have the same possibilities through cooperative and community building.

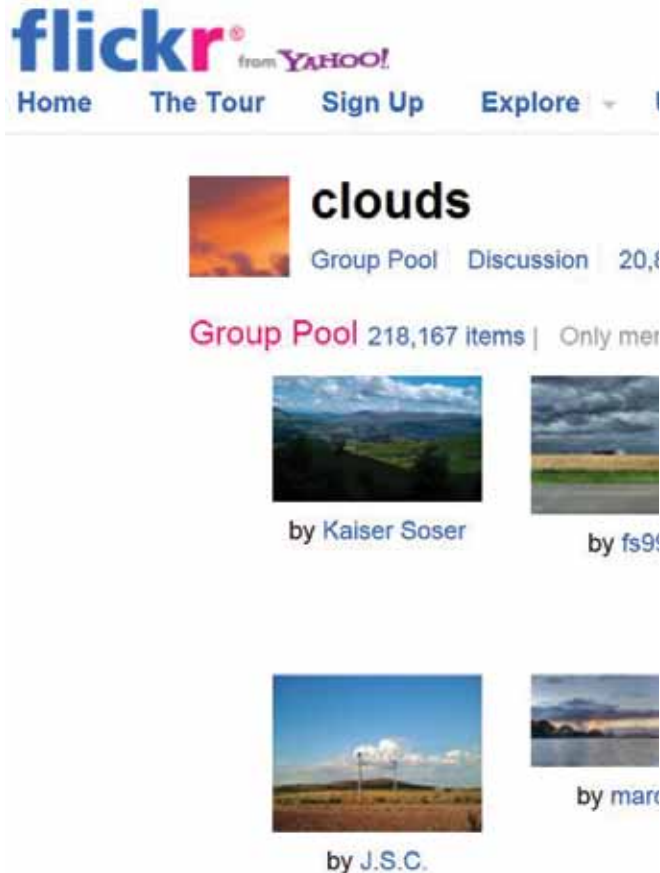
### Caveat emptor

If libraries are to consider moving more of their services into the cloud there are certain questions that must be addressed. Foremost is whether this service will make the library more

efficient and help it offer better service to its constituency. Once it is determined that a cloud solution does accomplish this for the library then considerations are:

- Does the service have built-in scalability, reliability and security?
- Is it multi-tenancy?
- Who owns the data stored in the system and what rights does the library have to extract their data for other uses or even to leave the service entirely?
- Is it an open system so that external data and services can be economically integrated into this service and its services can be economically integrated to external services?

In discussions of cloud computing, security and privacy are raised as serious concerns throughout the literature and especially by librarians. When considering a cloud application two aspects of security and especially privacy must be examined: technical and legal. Does the provider demonstrate the necessary technical expertise and



explanation of their environment to ensure there will be no unauthorised access to a library's data stored in the cloud? Have they considered the legal requirements of the government bodies the library is answerable to? This means it is also important to know exactly where the data is going to be stored since different countries have much different privacy requirements and standards. And since many cloud solutions are actually running on another supplier's cloud infrastructure due diligence is required: first to be certain where data is stored and secondly to establish what the service level agreements for access to the data and preservation of the data are for the actual infrastructure

supplier. Naturally data privacy and security are not mutually exclusive to cloud-based solutions.

Multi-tenancy is critical for scalability of any cloud service. "Multi-tenancy refers to a principle in software architecture where a single instance of the software runs on a server, serving multiple client organisations (tenants). Multi-tenancy is contrasted with a multi-instance architecture where separate software instances (or hardware systems) are set up for different organisations. With a multi-tenant architecture, a software application is designed to virtually partition its data and configuration, thus each client organisation works with a customised virtual application instance." The architecture that makes cloud solutions highly scalable thus must be considered when adopting a cloud service.

Data ownership cannot be overstressed. The library must know it has complete access to all their data while using the service so they can take it and re-use it as necessary, whether in another service or simply for reporting purposes. Just as important, they must be certain they can extract all their data at any point in the future should they decide to leave the service. It is also important to know what provisions are in place for data access should the supplier go out of business.

### Conclusion

Libraries have the opportunity to improve their services and relevance in today's information society. Cloud computing is one avenue for this move into the future. It can bring several benefits to libraries and give them a different future.

The cooperative effect of libraries using the same, shared hardware, services and data — rather than hosting hardware and software on behalf of individual libraries — can result in lowering the total costs of managing library collections and enhancing the both library user's experience and library staff workflows.

While local library systems served an important purpose earlier in library automation they now represent a tremendous duplication of effort. Each library builds and maintains a database, buys equipment and installs and updates the software. Some libraries can get stuck in perpetual upgrade mode, which involves lots of testing, retesting and time-consuming customisation.

With cloud computing, all this is taken care of transparently for the library and user. Among the benefits of a cloud computing approach:

- Take advantage of current and rapidly emerging technology to fully participate in the web's information landscape
- Increased visibility and accessibility of collections
- Reduced duplication of effort from networked technical services and collection management
- Streamlined workflows, optimised to fully benefit from network participation
- Cooperative intelligence and improved service levels enabled by the large-scale aggregation of usage data
- Make libraries greener by sharing computing power thus reducing carbon footprints

The vision is to use cloud computing to deliver library resources, services and expertise at the point of need, within user workflows and in a manner that users want and understand. It should free libraries from managing technology so they can focus on collection building, improved services and innovation. The cloud computing model will encourage libraries and their users to participate in a network and community of libraries by enabling them to reuse information and socialise around information. It can also create a powerful, unified presence for libraries on the web and give users a local, group and global reach. ■

Matt can be contacted on: [matt.goldner@oclc.org](mailto:matt.goldner@oclc.org)

Visit OCLC at: [www.oclc.org](http://www.oclc.org)

This article is based on Matt's whitepaper: 'Winds of Change: libraries and cloud computing'

References:

<sup>i</sup> Moore, Geoffrey: *Core Content and the Cloud*.

[www.youtube.com/watch?v=0swJCYLH2C](http://www.youtube.com/watch?v=0swJCYLH2C)

<sup>ii</sup> Gartner Group: [www.gartner.com/it/page.jsp?id=1035013](http://www.gartner.com/it/page.jsp?id=1035013)

<sup>iii</sup> Bezos, Jeff: [www.slideshare.net/goodfriday/amazon-web-services-building-a-webscale-computing-architecture](http://www.slideshare.net/goodfriday/amazon-web-services-building-a-webscale-computing-architecture)

<sup>iv</sup> Breeding, Marshall: *It's time to break the mold of the original ILS*. In: *Computers in Libraries* 27:10 (2007) S. 39-41

<sup>v</sup> Wikipedia: <http://en.wikipedia.org/wiki/Multi-tenancy>

## COPYRIGHT INFORMATION



**Author:** Goldner, Matt

**Title:** Winds of change: libraries and cloud computing

**Source:** Multimedia Inf Technol 37 no3 Ag 2011 p. 24-28

**ISSN:** 1466-190X

**Publisher:** Library Assoc. Multimedia & Information Technology Group and  
Multimedia Group of Aslib

Coach House Frongog, Llanbadarn Fawr, Aberystwyth SY23 3HN, Wales

The magazine publisher is the copyright holder of this article and it is reproduced with permission. Further reproduction of this article in violation of the copyright is prohibited.

This article may be used for research, teaching and private study purposes. Any substantial or systematic reproduction, re-distribution, re-selling, loan or sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden. The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.