Cloud Computing

Mike Gunderloy, a contributor to the Web Worker Daily blog and a frequent author on tech topics, made an interesting analogy in a post on his blog on July 30, 2008.[1] He compared the way we handle information to the way we handle money: either we can keep it all on our personal computers, subject to risk of hard-drive failures or natural disasters, or we can "deposit" it in the bank of servers that exist in the network–sometimes called the "cloud"–and just pull it out when we need it, much like using an ATM.

> *Web Worker Daily blog*
> http://webworkerdaily.com

Defining Cloud Computing

The cloud, in this case, is the massive network of storage devices (servers) that exist somewhere "out there" on the Internet. Wikipedia defines *cloud computing* as an "Internet-based ('cloud') development and use of computer technology ('computing')."[2] For this report, we will consider the cloud to be the network of servers that run the services (Facebook, Flickr, etc.) that can be used as collaborative platforms. When we use this cloud, we are uploading documents, data, images, and other artifacts of our work to a server or server farm that is under someone else's control. A server farm is a number of servers that are linked together to provide more storage and more computing power than a single server could alone (see figure 2). Often, in the Web 2.0 world, these services are free, but just as often there are versions that offer more storage, more features, or better accessibility for a price. This chapter will discuss some of the benefits that we can realize from our use of cloud computing and the pitfalls of using other organizations' servers to hold data.

Benefits of Cloud Computing

There are many benefits for those who use cloud computing. The ability to outsource much of the day-to-day technical duties–the commodity part of any job–frees employees to concentrate on other aspects of their work that could otherwise have been neglected. If an organization doesn't have to worry about doing backups, keeping hackers out of its data, or providing more virtual storage space, it can worry about the bigger picture and the more mission-focused projects that it might be working on. It is important to note that choosing cloud computing assumes a high degree of trust between the organization and its cloud computing provider, as the provider will be trusted with sensitive information and security details.

A cloud computing provider can be any company that runs many servers that are available for use either directly or as a part of an application like Facebook, Flickr, or Ning. Direct service providers give companies access to computing resources "in the cloud" that are scalable–many companies use services such as Amazon's Simple Storage Service (S3) to provide extra bandwidth and storage space that is reliable and can be expanded on a

moment's notice. The type of cloud computing provider I will be referring to in this report, however, offers more than just computing resources. Facebook, Flickr, and Ning are all providers of applications that make use of the cloud and can be considered cloud computing providers. When you post a message to Facebook, upload a picture to Flickr, or add a document to a Ning group, you are taking advantage of a cloud computing provider's resources.

Amazon S3
http://aws.amazon.com/s3

Chris Brogan, a popular social media blogger, provided a great example of cloud computing when he wrote a post about what happened when his main personal computer died.[3] Surprisingly, the event actually turned out to be much less trouble than he anticipated. Most of his daily computing life was conducted on various websites instead of on his personal computer—his life was "in the clouds." Because he used cloud-based applications like Gmail, Google Calendar, Evernote, Flickr, Google Docs, and Delicious (formerly del.icio.us) for his data, none of it was lost when his computer refused to boot up. Without even realizing it, he had started to make the move toward cloud computing.

Even those who prefer to work on their own computers can still take advantage of the cloud by using a service called Google Gears. Gears is a browser extension that works with Firefox and Internet Explorer to provide a connection between the cloud and your computer. Using Gears with a tool like Google Docs allows users to download the document they are working on to their personal computer for use when there might not be an Internet connection. Google's Documents products, which include spreadsheets, also work with Gears, as does the Zoho Office suite and other cloud-based services that are available on the Internet. Users interested in this approach can check the feature list of any service they use to see if it is compatible with Google Gears.

Google Gears
http://gears.google.com

For our purposes, the main benefit of cloud computing is the fact that all of the data that is being used, created, and referenced is kept "in the cloud." It is available from any computer and by any person who has the proper credentials—whether that is a password or just the URL of the data. While cloud computing comes in handy for
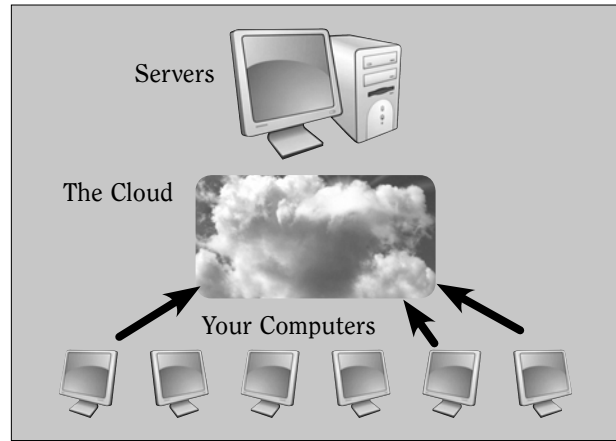
Figure 2
Basic cloud computing structure.

backing up personal data, as Chris Brogan discovered, it is also very handy for storing group data and team documents.

Problems with Cloud Computing

The previous section alluded briefly to potential security risks that come with using cloud computing. The potential for security breaches does exist, so it is extremely important that the library trust the system it is using and those administering use of that system to mitigate these risks.

While using the cloud to perform backups of important data is a good idea, it is also a good idea to use a personal computer or hard drive to perform backups of data kept in the cloud. If the data is lost or hacked and changed, a good copy will still be available. This practice eliminates the potential for damage that could be caused by either the loss of data on the service or a loss of network connection on either end of the link. In the summer of 2008, the Kevin Pipe website posted a roundup of outages or connection problems with cloud-based services.[4] The post included a number of examples of data that was not accessible, potentially at the moment when people needed it most. Pipe asked his readers this question: "Do you trust the cloud?"

When working with data or documents that an entire group can access and change, it is important to keep a close eye on versions of documents that are in the cloud, on individual computers, and elsewhere. This is where the Google Gears service comes in very handy—users can "back up" the cloud, just like you back up your personal computer, by using that browser extension. This will help to keep the panic from setting in when a service becomes unavailable or suddenly goes black—backups on a personal

computer will be available for staff, who can then work from those documents until the service comes back up or the organization decides to find a new home on the cloud for its data.

Terms of Service and Legal Issues: Who Owns Your Data?

The emergence of cloud computing has set off a debate over data ownership. Once data is uploaded to the cloud, who owns it? Does that data belong to the person who created it, or did that person give up ownership by uploading it? Data ownership is an issue any library must consider before making the decision to upload to a public server. Sensitive information like budget data, internal memos, or documents concerning major organizational decisions should never be uploaded to a public server (unless the library has a reason for wanting them to be public). On the other hand, for many collaborative projects, there isn't "ownership" in the traditional sense. For more open projects like these, the collaborative benefits of cloud computing are likely to outweigh any drawbacks.

Before uploading, it is important to check both the terms of service document (ToS) and legal statements of the service being used. Services like Google have a clear and easy-to-find ToS, a link to which can be found at the bottom of every page in the Google Docs help site. While the ToS for other services may not be this easy to locate, any reputable service should have a ToS that is accessible through its webpage or help documentation. A library administrator (or a lawyer if your library is big enough to have one) should read through the entire document and be sure that users understand how the service can use uploaded data and what steps it takes to protect that data. If the organization cannot find a public service that fits its data policy needs, downloading and installing one of the internal groupware suites that are discussed in chapter 7 would be a suitable alternative.

Google ToS
www.google.com/accounts/TOS

For information on specific terms of service for the services and applications that are profiled in this report, check the end of each service's section in chapter 6, which includes a small summary of the ToS for each tool.

Notes

1. Mike Gunderloy, "Is Your Information under the Mattress or in the ATM?" Web Worker Daily, July 30, 2008, http:// webworkerdaily.com/2008/07/30/information-under -mattress-or-in-atm (accessed March 17, 2009).
2. "Cloud Computing," Wikipedia, http://en.wikipedia.org/ wiki/Cloud_computing (accessed March 17, 2009).
3. Chris Brogan, "Life In The Clouds," Chris Brogan: Community and Social Media, July 31, 2008, www.chris brogan.com/life-in-the-clouds (accessed March 17, 2009).
4. "Do You Trust the Cloud? [Ask the Readers]," The Kevin Pipe, Aug. 12, 2008, http://thekevinpipe.com/2008/08/12/do-you -trust-the-cloud-ask-the-readers (accessed March 17, 2009).