

DuraCloud

<http://duracloud.org/>

Reviewed by Blake E. Relle, Louisiana Secretary of State/Louisiana State Archives

DuraCloud is an open source platform used to store and manage digital content in the cloud through a single dashboard.¹ DuraSpace, a not-for-profit organization that provides services to facilitate the long-term accessibility of digital collections, hosts and manages the platform. This review assesses the free trial version of DuraCloud's Preservation Plus plan.

DuraCloud allows users to store multiple copies of content in the cloud with the option of selecting one, two, or three vendors in different geographic areas.² Storing content with multiple providers in different geographic areas can decrease the risk of losing content in the event of a disaster.³ Users can choose from Amazon Glacier, Amazon Web Services, and Rackspace. DuraCloud handles negotiations with cloud providers and problems that may arise. By having DuraCloud handle day-to-day operations, archivists can use their time on other projects. One is not limited to a provider after initial selection because users can switch storage providers without renegotiating.

DuraCloud provides storage, backup, infrastructure, and preservation management services, such as verification of file fixity, preservation event tracking, and replication across multiple locations. DuraCloud also streams video and audio files to any device linked to the Web. Users don't need to install software because DuraCloud is completely web-based. Additionally, users don't need advanced technical knowledge.

DuraCloud's automatic health-checking service verifies the integrity of the repository's files.⁴ When a file is uploaded, DuraCloud computes its checksum and periodically recalculates it to determine if any changes have occurred.⁵ If a change is detected, the next step depends on the repository's preservation plan, which determines the number of copies stored. If the repository only stores one copy, the system emails the repository that an issue has been detected with a file and that it needs to be replaced. If the repository stores multiple copies, the system finds the

¹ "Overview," DuraCloud, accessed March 21, 2016, <http://duracloud.org/overview>; "What is DuraCloud?" DuraCloud, accessed June 2, 2016, <http://duracloud.org/tour>.

² "Cloud Broker," DuraCloud, accessed March 21, 2016, http://duracloud.org/cloud_broker; "Subscription Plans," DuraCloud, accessed March 21, 2016, <http://duracloud.org/pricing>.

³ "Benefits," DuraCloud, accessed March 21, 2016, <http://duracloud.org/details/benefits>.

⁴ "Health Checkup," DuraCloud, accessed March 21, 2016, http://duracloud.org/health_checkup.

⁵ Smith, Carissa, email to the author, March 15, 2016.

file in the secondary storage location. If the second file is healthy, DuraCloud replaces the corrupted file with a healthy one and notifies the repository.⁶ If both copies are corrupt, the system emails the repository so its staff can upload a replacement copy.

DuraCloud provides two ways to upload files to the platform: individually or using DuraCloud's Sync Tool.⁷ Uploading files individually is completed either through file selection and upload or drag-and-drop using DuraCloud's interface. This method requires the user to initiate upload every time the file is updated. The second option is to download DuraCloud's free application called Sync Tool to transfer content. The Sync Tool begins with a configuration wizard and provides an interface display showing the current status of the sync process.⁸ Then users can create a "watch list," a list of files stored on the repository's servers that will be monitored for new content or changes once the initial uploading of files is completed. The Sync Tool adds new and changed files to the repository's DuraCloud account. The Sync Tool can run in single pass uploading mode or continuous uploading mode. In the single pass mode, the application does not monitor the watch list for changes or new content once the initial uploading is completed. The continuous uploading mode continues to monitor for any changes or new content to the watch list after the initial uploading is complete. One can download the Sync Tool by going to the DuraCloud Wiki page.

After uploading content users can create tags; however, searching by tags is not provided as an option. Carissa Smith, Service Product Manager at DuraCloud, said the company has identified tag-based searching as a feature they would like to add to DuraCloud in 2017.⁹ Currently, tags don't facilitate discovery. DuraCloud provides content sharing in all the plans offered to make files accessible to the public. Each file is assigned a specific storage URL and is only accessible to users who log into the platform. Archivists can make specific files accessible to the public or closed to the public. The public receives read-only copies to prevent any changes to the file stored in the platform.

DuraCloud provides five plans that cost from \$1,235 to \$5,550. All plans provide content sharing, a web-based administrative dashboard, checksum verification, storage reports and statistics, and one terabyte of storage. Additional terabytes can be purchased for costs ranging from \$625 to \$1,400, depending on one's subscription plan. DuraCloud provides quotes for more than ten terabytes of storage. Two plans store one or two copies of content in the cloud and provide

⁶ Carissa Smith, email to the author, June 29, 2016.

⁷ Bill Branan, "How to Automatically Upload Files and Directories to DuraCloud," DuraSpace Wiki, accessed June 17, 2016, <https://wiki.duraspace.org/display/DURACLOUDDOC/How+to+Automatically+Upload+Files+and+Directories+to+DuraCloud>.

⁸ "Download and Installed," DuraSpace Wiki, accessed June 27, 2016, <https://wiki.duraspace.org/display/DURACLOUDDOC/DuraCloud+Sync+Tool>.

⁹ Carissa Smith, email to the author, March 15, 2016.

access, and two are for storage without access. Some plans provide additional features, including automatic synchronization of content between copies and automatic file recovery. DuraCloud also offers a plan called DuraCloud Enterprise Chronopolis. This plan was developed to provide repositories with a Trustworthy Repositories Audit and Certification (TRAC) dark storage option. The subscription price for this plan is \$2,890 with a one-time ingest fee of \$310 per terabyte and an annual storage cost of \$500 per terabyte.

I found the free trial of DuraCloud's Preservation Plus Plan easy to use.¹⁰ The "Getting Started" feature links to videos detailing how to use DuraCloud's tools, features, and security. I could upload files individually or use the Sync Tool. However, my computer's security settings would not allow me to download the Sync Tool. After uploading, I could create subject tags for the uploaded files and choose between two cloud providers. Since I could not search using the subject tags, this feature does not offer a benefit. Once the uploading process finishes, DuraCloud provides the following file information: file size, modification date, checksum, and storage location. DuraCloud also continues to check that file health and monitor changes.¹¹

DuraCloud is a useful tool for archivists because it provides a way to preserve a repository's digital content without having to purchase and maintain equipment. However, the annual fee may be cost prohibitive for some repositories. DuraCloud helps repositories make data accessible in the future by being another set of eyes that checks to make sure the data is healthy, unchanged, and safe. The inability to conduct tag-based searching, however, may hinder users in locating content. Despite this drawback and the subscription cost, DuraCloud is a useful tool to assist archivists in managing and preserving digital content.

¹⁰ "Try DuraCloud Right Away," DuraCloud, accessed July 8, 2016, <http://duracloud.org/trial-account-request>.

¹¹ "Frequently Asked Questions," DuraCloud, accessed March 21, 2016, <http://duracloud.org/faq>.