

A stylized graphic of a mountain range or cloud formation in a light blue color, positioned above the main text.

DURACLOUDTM

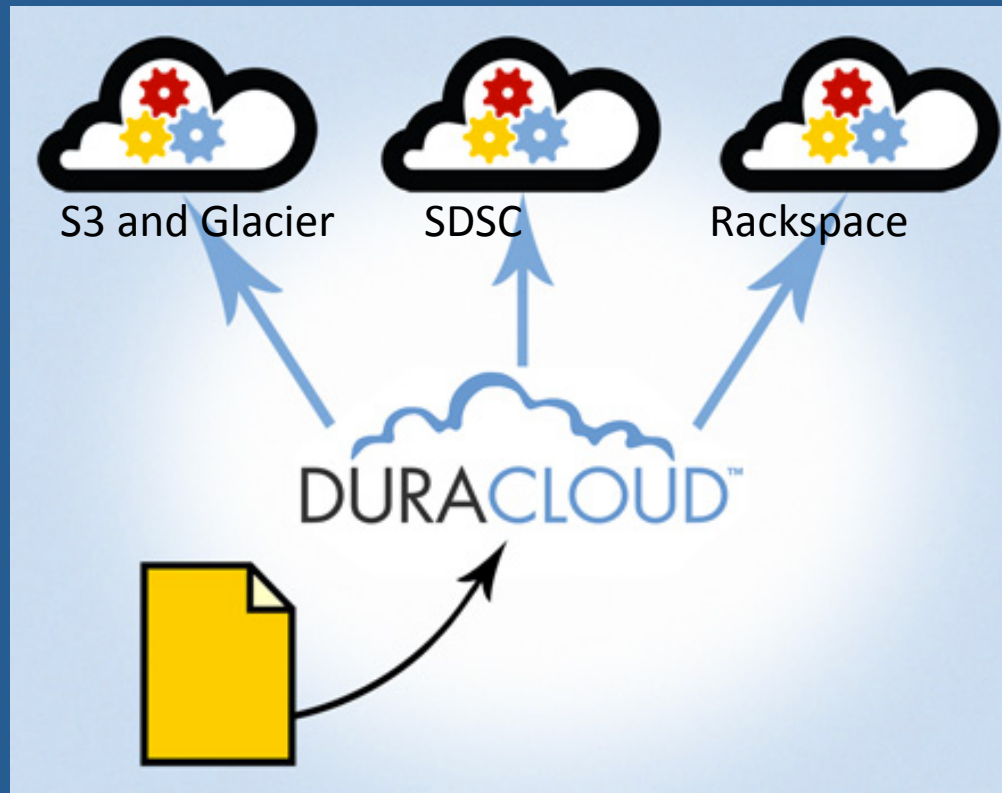
Integration with Glacier

Michele Kimpton, CEO DuraSpace
Bill Branan, Senior Tech Lead DuraCloud
Tim Harder, Business Development Manager, Amazon

What is DuraCloud?

Archiving and preservation services in the cloud

Ability to choose one or multiple cloud storage providers



What does DuraCloud do?

Online backups



File health checking



File synchronization and healing



Online sharing



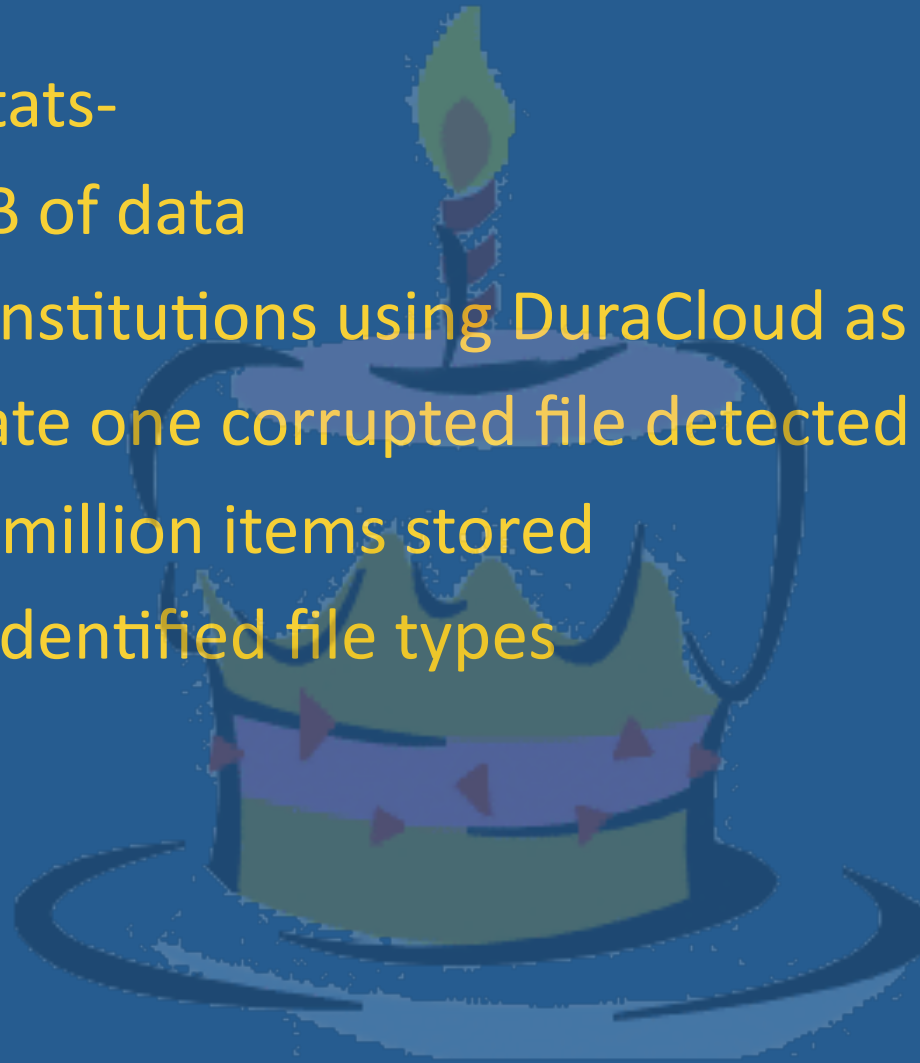
Media streaming



Content viewing

In production for one year

- Some stats-
 - 35 TB of data
 - 25+ institutions using DuraCloud as a service
 - To date one corrupted file detected and restored
 - 5.46 million items stored
 - 60+ identified file types



How can DuraCloud help me with my archiving and preservation strategy?

- ✓ Ability to easily move online copies offsite, distribute geographically
- ✓ Automatically checks the health of all copies, compared to manifest
- ✓ Easily repair or replace corrupted files
- ✓ Easily move or download content from multiple providers
- ✓ Scale storage up or down as needed
- ✓ Full Web access with ability to view/edit/delete and upload/download your content

What users wanted

- Very low cost second copy, comparable to tape
- Accessible for audit purposes
- Can be retrieved reliably and easily in disaster recovery scenario
- Can archive large amounts of data (50-100+TB)
- Simple and easy to use, and ingest content

Unique Benefits of DuraCloud/Glacier

- Content stored in DuraCloud can be automatically backed up to Glacier and synchronized
- Manifest created upon upload
- DuraCloud provides one easy to use interface across all storage providers for viewing content
- DuraSpace provides one annual invoice and service agreement
- DuraCloud plans to provide periodic health check audits on Glacier content

DuraCloud Glacier Cost

- Glacier storage +10% premium, pay once per year
- Glacier is 1/10 of the cost of S3 (\$120 per TB/yr)
- Retrieval cost calculated based on request to protect the user from unpredictable charges
- Pay upfront for storage only, pay for retrieval when you need it upon consultation with DuraSpace
- Minimum purchase is 1Tb increments, annual

Pricing

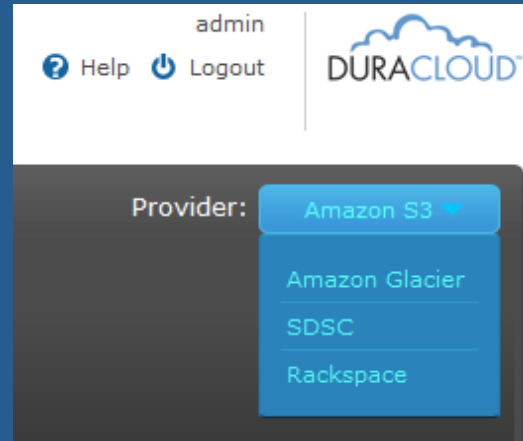
<http://www.duracloud.org/content/pricing>

Subscription Plan	Price
<p>DuraCloud Preservation</p> <p>The DuraCloud Preservation Basic plan is ideal for institutions that wish to store one copy of their content in the cloud. This subscription plan includes the following features: storage of your content in Amazon S3, online access through the DuraCloud web interface, automatic content health checking services, and storage reports.</p>	<p>Annual price (storage in Amazon S3):</p> <ul style="list-style-type: none">• \$1,500 for 1st TB• \$1,000 for additional TBs
<p>DuraCloud Preservation Plus</p> <p>The DuraCloud Preservation Plus plan is best suited for organizations that wish to store two copies of their content in the cloud. This subscription plan includes all of the features of the Preservation Basic plan as well as the following features: automatic synchronization of content between primary and secondary storage providers, your choice of secondary cloud storage providers (Amazon Glacier, San Diego Supercomputer Cloud Store, and Rackspace Cloudfiles), and automatic file recovery between copies.</p>	<p>Annual price (storage in Amazon S3 + Amazon Glacier):</p> <ul style="list-style-type: none">• \$1,635 for 1st TB• \$1,135 for additional TBs <p>Annual price (storage in Amazon S3 + SDSC):</p> <ul style="list-style-type: none">• \$2,500 for 1st TB• \$1,700 for additional TBs

Glacier Integration (Current)

- DuraCloud secondary storage providers

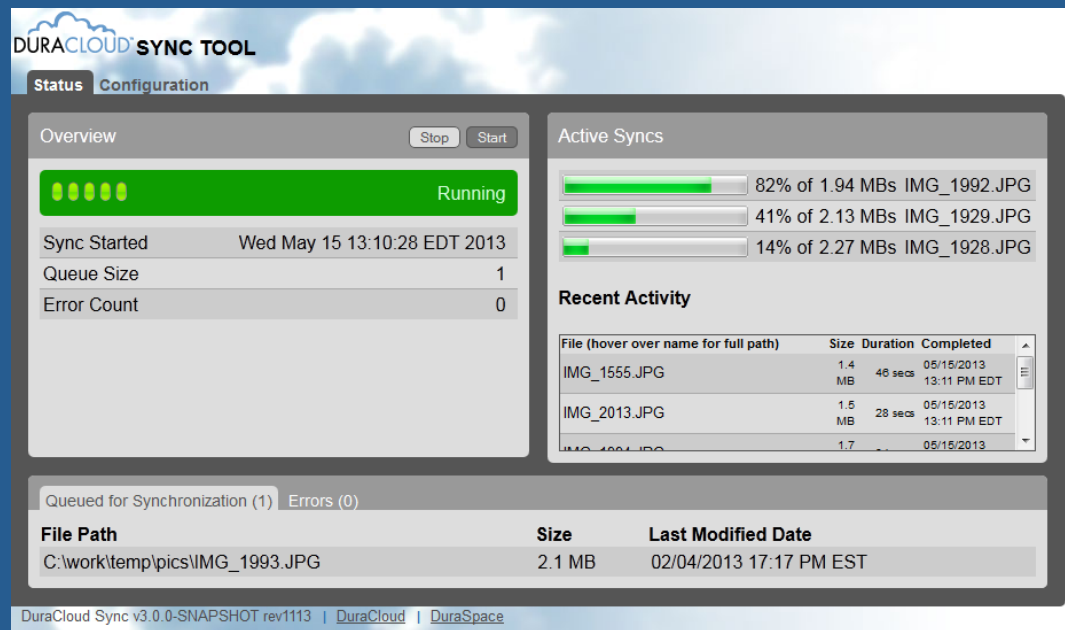
- Rackspace
- SDSC
- Glacier



- Primary storage provider remains Amazon S3
- Glacier provider the same as Rackspace and SDSC
 - **EXCEPT** that stored content is not immediately accessible
 - Can always view listing of files
 - Can always view file properties

Getting Content into Glacier

- Transfer to DuraCloud primary storage
 - Sync Tool
 - GUI
 - Command line
 - REST API
 - Java client



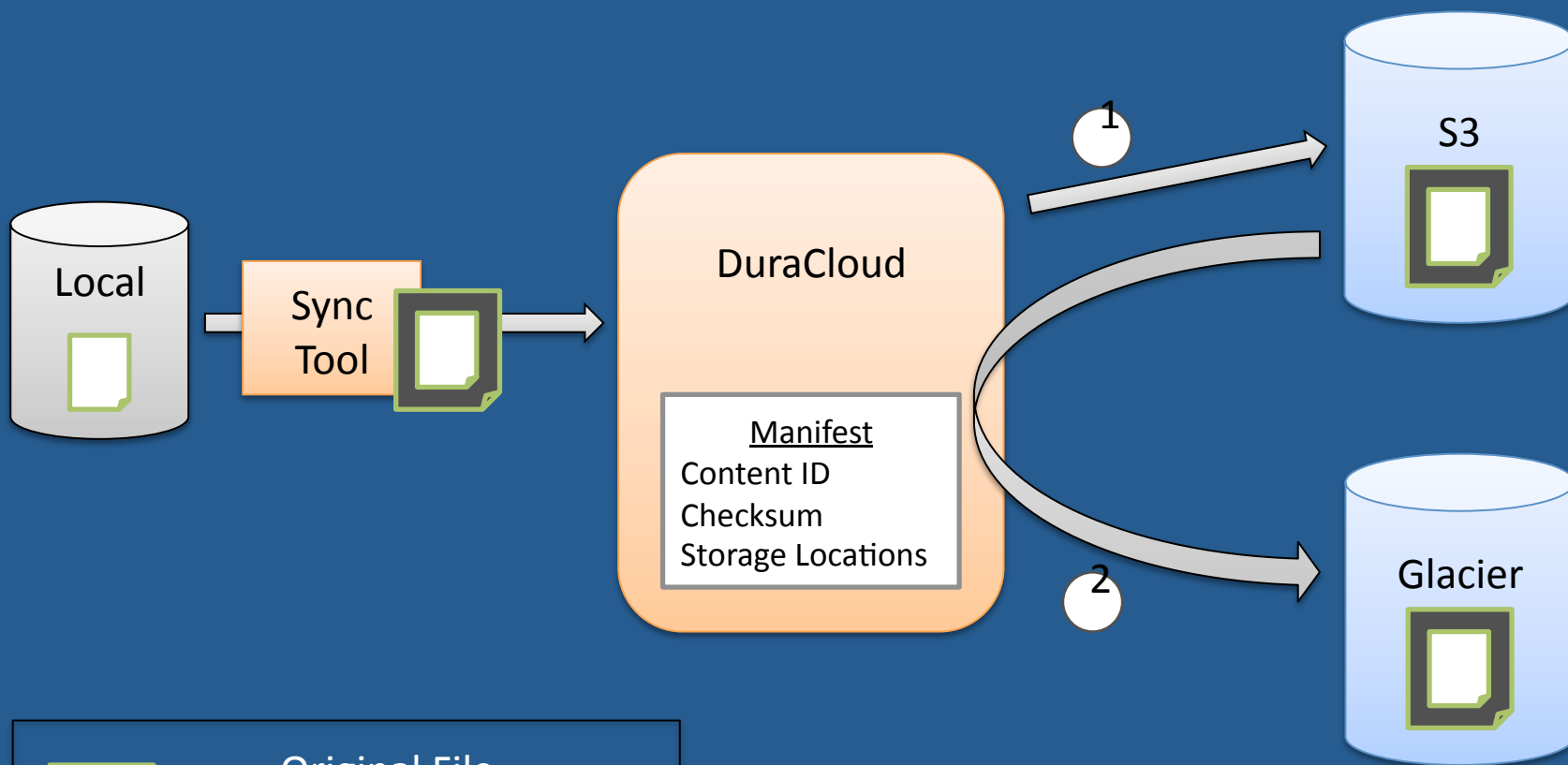
The screenshot displays the DuraCloud Sync Tool interface. The 'Overview' section shows a green progress bar and a 'Running' status. The 'Active Syncs' section lists three files with their respective progress: IMG_1992.JPG (82% of 1.94 MBs), IMG_1929.JPG (41% of 2.13 MBs), and IMG_1928.JPG (14% of 2.27 MBs). The 'Recent Activity' table shows completed syncs for IMG_1555.JPG, IMG_2013.JPG, and IMG_1993.JPG. The 'Queued for Synchronization' section shows one file, IMG_1993.JPG, with a size of 2.1 MB and a last modified date of 02/04/2013 17:17 PM EST.


File (hover over name for full path)	Size	Duration	Completed
IMG_1555.JPG	1.4 MB	46 secs	05/15/2013 13:11 PM EDT
IMG_2013.JPG	1.5 MB	28 secs	05/15/2013 13:11 PM EDT
IMG_1993.JPG	1.7 MB		05/15/2013

File Path	Size	Last Modified Date
C:\work\temp\pics\IMG_1993.JPG	2.1 MB	02/04/2013 17:17 PM EST

- DuraCloud automatically:
 - Copies all content to Glacier
 - Keeps all providers in sync

DuraCloud Ingest Process



 = Original File
= Calculated Checksum
= File Properties

Content in Glacier

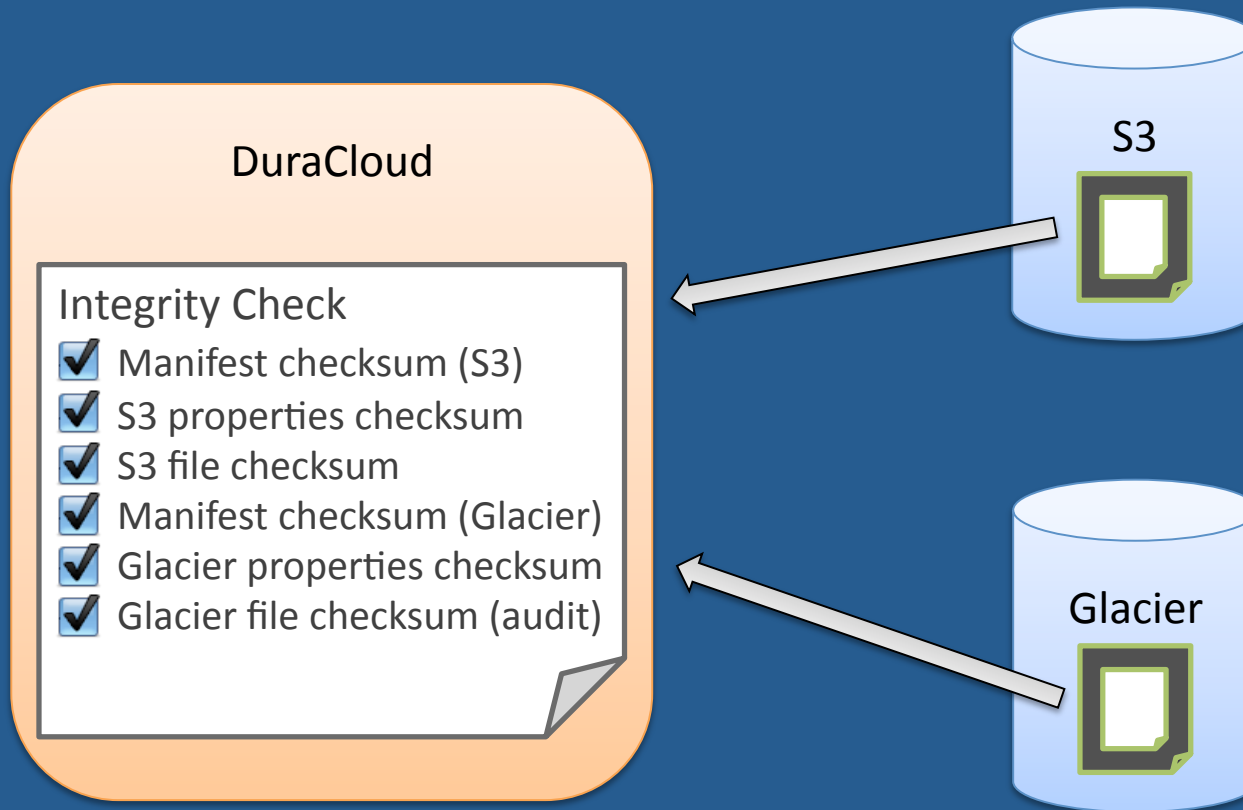
The screenshot displays the DuraCloud Administrator interface. At the top, the DURA CLOUD logo is on the left, and user information (admin) and navigation links (Help, Logout) are on the right. Below the logo, there are tabs for 'Spaces' and 'Services', and a 'powered by amazon web services' badge. The main content area is divided into three sections:

- Spaces:** A list of spaces including 'bb-glacier-test' and 'x-service-out'. A search filter is present.
- Content Items:** A list of 12 items, with 'nasa.tiff' selected. The list includes files like 'family/bill-plus-kids.jpg', 'shuttle/launch/atlantis-1.jpg', and 'test-file-1.txt'. A 'Refresh' button and a 'Get Sync Tool' button are at the top of this section.
- Content Detail:** A detailed view of the selected 'nasa.tiff' file. It shows the mime type as 'image/tiff' and provides 'Download' and 'View' buttons. Below this is a 'Details' section with a table of metadata:

Details	
Space	bb-glacier-test
Size	7.9 MB
Modified	2013-02-21
Checksum	7d5d1890e19e19c8256d9c3afd66c527
Properties	
creator	root
content-file-last-accessed	2011-06-14T22:41:53.053
content-file-path	C:\work\other\bb-pics\nasa.tiff
content-file-modified	2009-08-10T14:24:50.050
content-file-created	2011-05-23T17:00:24.024

At the bottom of the interface, there is a footer with the text: 'Duracloud Administrator Release v2.3.1 1107 | ©2013 DuraCloud | DuraSpace | Management Console | Help Center | Contact Us' and the DÜRASPACE logo on the right.

DuraCloud Verification Process



Getting Content out of Glacier

- Files not available for immediate download
- Submit a DuraCloud support ticket
- We work with you to:
 - Define the set of content to be downloaded
 - Determine the optimal speed of retrieval
 - Goal: Retrieve content within the necessary time frame for the lowest possible cost
- We execute the retrievals from Glacier
- You download files through DuraCloud
 - Same process as downloading from any other provider

Glacier Integration (Future)

- Glacier as the primary DuraCloud storage provider
 - No secondary providers
 - Lowest possible cost
 - Goal: In place by the end of 2013
 - Working with Amazon to define preferred pricing and retrieval options
- Automated retrieval options
 - For simple, well defined use cases

Amazon Glacier Overview

Timothy K. Harder / harder@amazon.com +1 508-308-4660



What we heard from our customers

**Reliable and cheap storage
of data for:**

- Data with long retention periods**
- Multi-PB, infrequently accessed data sets**

Our customers love Amazon S3 for its simplicity, security, durability, and performance.

Our customers that need to archive data wanted a highly secure, extremely durable, and extremely cost effective option for archiving data for years while retaining the ability to easily retrieve data occasionally.



Our goals with Amazon Glacier

- ❏ Deliver an extremely low cost storage service that provides secure and durable storage for data archiving and backup
- ❏ Replace physical media for archiving, with an easy to use storage service that is infinitely scalable, secure, designed for an annual average 99.999999999% durability per saved object, for as little as \$0.01 per gigabyte per month
- ❏ Change the game for data archiving and backup with no upfront payment, a very low price for storage, and ability to scale up and down as needed



Benefits with Amazon Glacier



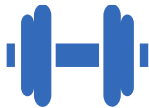
Low cost

As little as \$0.01/GB/month with no up-front capital commitments.



Secure

Secure and durable technology platform with industry-recognized certifications and audits.



Durable

Average annual durability of 99.999999999% per archive.



Simple

Eliminate hardware, software, and capacity planning.



Flexible

Add any amount of data, quickly. Easily expire and delete without handling media.



Use multiple services

Easily leverage other AWS services once your data is in the AWS cloud.



These are some customer use cases

Offsite Archive

- ❏ Glacier allows you to cost-effectively and securely store enterprise data offsite, making it simple, inexpensive and safe to retain archived data for as long as desired. Common use cases include enterprise data, media assets, and research and scientific data.

Digital Preservation

- ❏ Libraries, historical societies, non-profit organizations and governments are increasing their efforts to preserve valuable but aging digital content such as websites, software source code, video games, user-generated content and other digital artifacts.

Tape Replacement

- ❏ Tape can be one of the biggest backup pain points in the IT org. Amazon Glacier is cost competitive, even at scale, and eliminates pain points like capacity planning, capital budgeting and investments, media formats, hardware refreshes, and off-site storage costs, shipping and retrieving.



Here are some typical use cases

	Archive	Backup	DR
Amazon S3	Data accessed ~>10% / month 11 9s durability	Snapshots Shorter term data backup with rapid RTO	Rapid RTO Expiration policies
Amazon S3 RRS	Lower cost when 11 9s not required	Lower cost	Lower cost
Amazon Glacier	Long term archiving Infrequent data access (~<10% data/month)	Use policies to move cold backup data for long term retention	Retain write once read never copy in case of worst case scenario



“Every day our genome sequencers produce terabytes of data. As our company moves into the clinical space, we face a **legal requirement to archive** patient data for years that would drastically raise the cost of storage.

Thanks to Amazon Glacier’s secure and scalable solution, we will be able to provide **cost-effective, long-term storage** and thereby eliminate a barrier to providing whole genome sequencing for medical treatment of cancer and other genetic diseases.”

Keith Raffel, Senior Vice President and Chief Commercial Officer, Complete Genomics

What people are
saying





“An organization like ours thinks in centuries when it comes to content retention, and long term preservation of our Master Archives is a critical part of our mission here at NYPR.

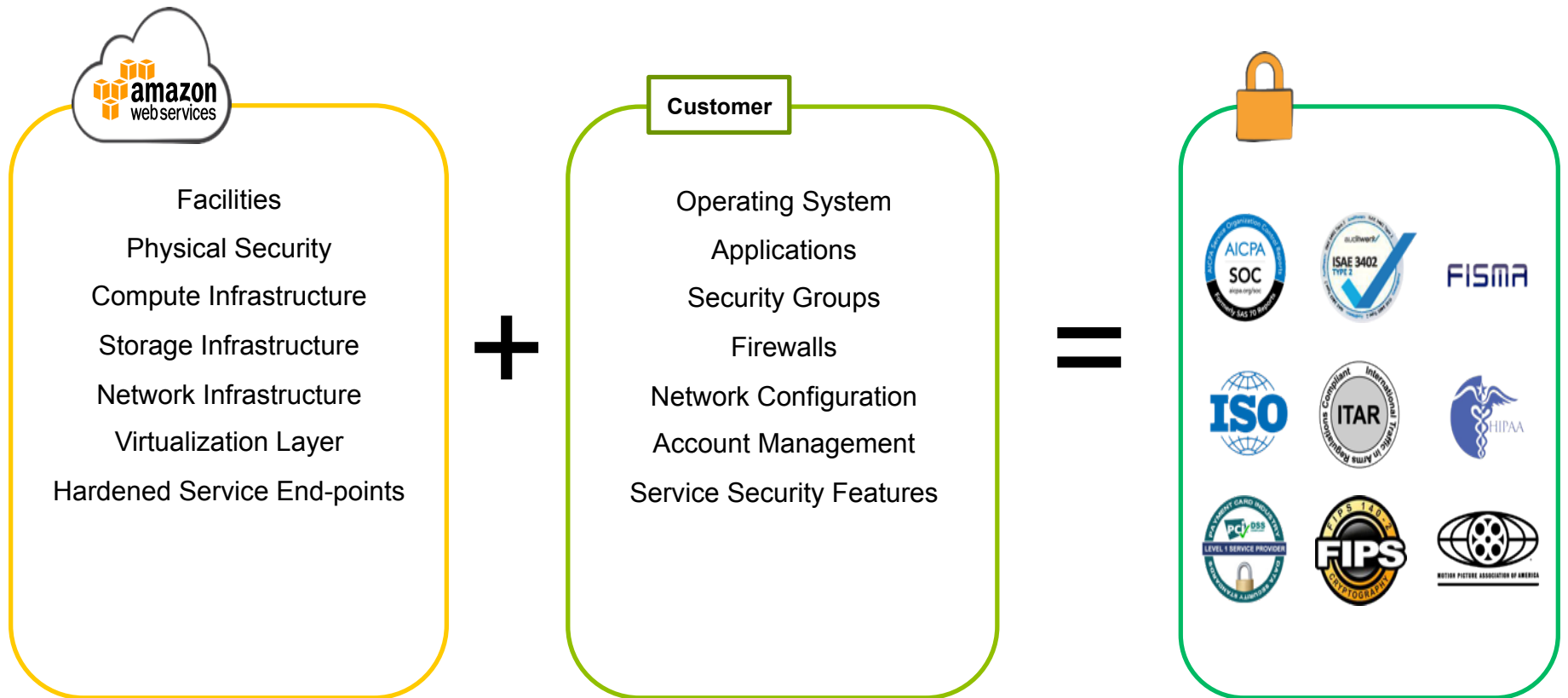
Storing these core assets on traditional media such as local disk and off-site tape exposes us to corruption and even outright-loss of data. We are excited to move our archives to Amazon Glacier, which will be a better long-term solution.”

Steve Shultis, CTO, New York Public Radio

What people are saying



Shared Responsibility Model



Available Reports & Certifications

Information at <https://aws.amazon.com/choosing-a-cloud-platform/>

- AWS SOC 1 Report (Type 2)
- AWS SOC 2 – Security Report (Type 2)
- AWS SOC 3 – Just Published
- NIST 800-53 control mappings and FISMA Moderate ATOs
- PCI DSS Level 1
- ISO 27001
- ITAR and GovCloud
- Other
 - FedRAMP
 - CSA Controls
 - MPAA



AWS Security and Compliance Center

(<http://aws.amazon.com/security/>)

- Answers to many security & privacy questions
 - Security whitepaper
 - Risk and Compliance whitepaper
- Security bulletins
- Customer penetration testing
- Security best practices
- More information on:
 - AWS Identity & Access Management (AWS IAM)
 - AWS Multi-Factor Authentication (AWS MFA)



Next steps

- Glacier available as secondary backup to all DuraCloud customers
- “Glacier only” choice available by end of year
- Working with Amazon to present use cases specific to the academic community and improve current retrieval costs for disaster recovery
- Sign up to keep informed and for the beta launch

To find out more

- Web:
 - www.duracloud.org
- Documentation:
 - <http://wiki.duraspace.org/display/duracloud>
- Webinars and Training Videos:
 - <http://www.youtube.com/user/duracloudvideos>
- Email:
 - csmith@duraspace.org or mkimpton@duraspace.org