



Cold Data Archiving in the Cloud

Library of Congress
Designing Storage Architectures (DSA), April 2024

Vikranth Etikyala
Senior Staff Software Engineer, SoFi

Cold Data

- Infrequently accessed
- Not required real time
- Lower priority
- Older or historical data

Why so Cold?

- Regulatory Compliance - Banks
- Audit and Legal - Governments
- Historical Records - Genomic Data
- Disaster Recovery - Backup and Restore
- AI/ML Models - ChatGPT

Missing or Deleted Data will most likely be the data you need

Cold Data - Archival and Benefits

- Cost
- Scalability
- Performance

Archival Solutions on Cloud

- AWS S3 - Glacier
- Azure Archive Storage
- Google Cloud Storage
- Oracle Archive Storage
- IBM Cloud Object Storage



Cloud Archival Benefits - First Order

- Serverless
- Scalable and Elastic
- Native Cloud Integrations
- Lifecycle - Auto Transition to Archives - Date Based
- Security
- Service Level Agreements (SLA's)

Cloud Archival Benefits - Second Order

- Data Center Management - Host, Racks, Temperature
- Energy - Carbon Footprint
- Capacity Management
- Operations Team
- Distributed System Problems
- Storage Media Advancements

Example: S3 Scale

Capacity and throughput

Amazon S3 holds more than **280 trillion objects** and averages over **100 million requests per second**

Events

Every day, Amazon S3 sends over **125 billion event notifications** to serverless applications

Replication

Customers use Amazon S3 Replication to **move more than 100 PB** of data per week

Cold Storage Retrieval

Every day, customers **restore more than 1PB** from the S3 Glacier Flexible Retrieval and S3 Glacier Deep Archive storage classes

Data Integrity Checks

Amazon S3 performs over **4 billion checksum computations per second**

ns advanced metrics and
greater than the Storage Lens cost

Cost Optimization

On average, customers using Amazon S3 Storage Le
recommendations have obtained **cost savings 6x g**
in the first six months of using it.

n Amazon S3

Flexibility

Hundreds of thousands of data lakes are built o

Thank you

Questions: vik.e@live.com