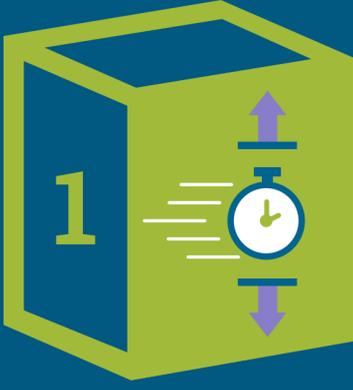


TOP 10 TIPS FOR Using Amazon S3

Amazon S3 gives modern organizations and enterprises a powerful and durable way to shape, predict, and scale storage dynamically while only paying for the space needed at any given time. S3 is easy to provision and the following tips will help your team extract the maximum value from S3 at a lower cost at scale.



GET LOG DATA INTO AND OUT OF S3 FASTER

Cutting down the time you spend uploading and downloading files can be valuable
DirectConnect ports offer significantly higher bandwidth, S3 Transfer Acceleration gets data into AWS faster by changing API endpoints, and s4cmd and aws-cli make it easy and fast to transfer large files with concurrent connections along with multi-part upload.

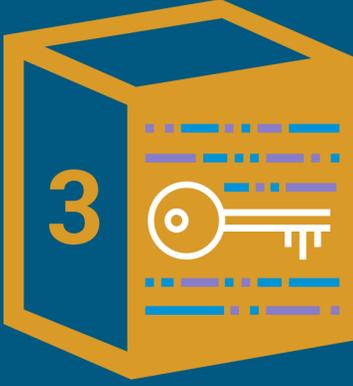
THINK THROUGH DATA LIFECYCLES UP FRONT

Before putting an item in S3, ask yourself when and how it should be deleted
Managed lifecycles and S3 object tagging make it easier to apply lifecycle policies by allowing you to delete or archive data based on object tags.



PRIORITIZE ACCESS CONTROL, ENCRYPTION, AND COMPLIANCE

Every business has sensitive data – it's just a matter of which data and how sensitive it is
Minimally, you'll want to store data with different needs in separate S3 buckets, regions, and/or AWS accounts, and set up documented processes around encryption and access control for that data.



NESTED S3 FOLDER ORGANIZATION IS GREAT – EXCEPT WHEN IT ISN'T

Newcomers are surprised to learn that latency on S3 operations depends on key names
If you need high volumes of operations, it is essential to consider naming schemes with more variability at the beginning of the key names to avoid internal "hot spots."



SAVE MONEY WITH REDUCED REDUNDANCY, INFREQUENT ACCESS, OR GLACIER

S3's standard class offers high durability, high availability, low latency access, and relatively cheap access cost
Reduced Redundancy Storage costs less but can cause data loss, while Infrequent Access and Glacier offer cheaper storage with more expensive access.

ORGANIZE S3 DATA ALONG THE RIGHT AXES

Avoid organizing data in a way that causes business risks or costs down the road
It's often best to organize your data into different buckets and paths at the highest level – not based on what the data itself is, but rather by considering sensitivity, compliance, lifecycle, realm, and visibility.

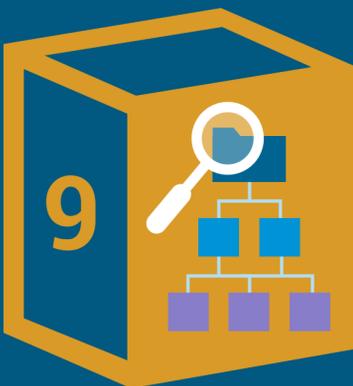
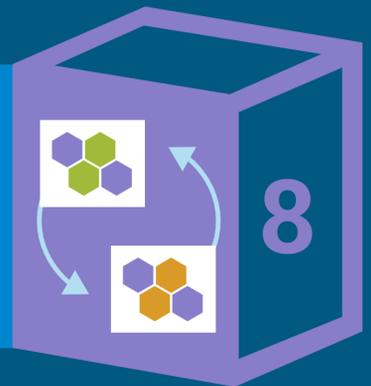


DON'T BAKE S3 LOCATIONS INTO YOUR CODE

Tying your code to deployment details is almost guaranteed to cause problems later
Instead of hard-coding S3 locations, de-couple code and S3 locations by setting up a configuration file or service and reading S3 locations like buckets and prefixes from there.

DEPLOY YOUR OWN TESTING OR PRODUCTION ALTERNATIVES TO S3

Services that are compatible with S3 APIs are helpful for testing and migration to local storage
S3Proxy (Java), FakeS3 (Ruby), and other services make it faster and easier for users to test S3-dependent code in isolation.

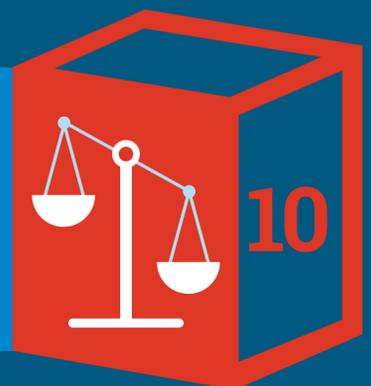


CHECK OUT NEWER TOOLS FOR MAPPING FILESYSTEM AND S3 DATA

S3 doesn't offer all the right features to make it a robust filesystem
Riofs (C) and Goofys (Go) offer improved solutions that use a different object format and allow filesystem-like access.

DON'T USE S3 IF ANOTHER SOLUTION IS BETTER

Consider that S3 may not be the optimal choice for your use case
Glacier and other S3 variants offer cheaper pricing, while EBS and EFS can be much more suitable for random-access data.



Sumo Logic can add real-time visibility to your global Amazon environment – including S3 Storage – by allowing you to examine, analyze, and validate critical data. Our technology offers better monitoring, root cause analysis, and customizable reports and auditing to help you get more out of your data, reduce MTTI and MTTR, and consolidate your organization's observability stack.

Sign up for Sumo Logic Free to try it for yourself at www.sumologic.com/lp/log-monitoring/full-stack-log-analytics/