50 S3 Features that every Cloud Engineer should know

1. **Data Transfer Acceleration:**

   Amazon S3 Transfer Acceleration enables fast, easy, and secure transfers of files over long distances between your client and an S3 bucket.

2. **S3 Select and Glacier Select:**

   These allow you to retrieve only a subset of data from the object by using simple SQL expressions.

3. **Object Lock:**

   S3 Object Lock is used to prevent an object from being deleted or overwritten for a fixed amount of time or indefinitely.

4. **Data Consistency Model:**

   Amazon S3 offers immediate consistency for PUTS of new objects and eventual consistency for overwrite PUTS and DELETES.

5. **Cross Region Replication:**

   This feature enables automatic, asynchronous copying of objects across buckets in different AWS Regions.

6. Object Storage:

You store your data in Amazon S3 as objects. An object is composed of the data you want to store along with metadata.

7. S3 Storage Classes:

Amazon S3 offers a range of storage classes designed for different use cases: S3 Standard, S3 Intelligent-Tiering, S3 Standard-IA, S3 One Zone-IA, and Amazon S3 Glacier, each having different availability, durability, and cost characteristics.

8. Versioning:

Once you enable Versioning for a bucket, Amazon S3 preserves existing objects anytime you perform a PUT, POST, COPY, or DELETE operation on them.

9. Lifecycle Management:

With S3 Lifecycle management, you can set up a policy to manage objects throughout their lifecycle.

10. Bucket Creation:

This is where you start with S3. A bucket is a container for objects stored in Amazon S3. Understanding how to create and configure a bucket is a must.

11. Bucket Policy:

This allows you to add a permissions policy to a bucket that authorizes or denies specific actions on it and its objects.

12. ACLs:

Access control lists (ACLs) are one of the ways to manage access permissions to your buckets and objects.

13. Pre Signed URLs:

This feature allows you to share specific objects with others by providing them with a URL that is valid for a specific time period.

14. Multipart Upload:

This feature allows you to upload large objects in parts, which can improve performance and allows pause and resume for the upload.

15. Encryption:

Amazon S3 offers several methods for encrypting your data at rest or in transit, such as SSE-S3, SSE-KMS, SSE-C, and client-side encryption.

16. S3 Events:

You can set up notification of specified bucket events, such as object creation or deletion.

17. Requester Pays:

An Amazon S3 bucket owner can specify that anyone who requests access to objects in a particular bucket must pay the data transfer and request costs.

18. Transfer Acceleration:

This feature uses the CloudFront Edge Network to accelerate your uploads to S3.

19. Storage Metrics and Inventory:

Amazon S3 provides storage metrics and inventory for analysis, reporting, or further action.

20. Tagging:

With S3 object tagging, you can manage and control access for Amazon S3 resources, apply AWS Identity and Access Management (IAM) policies to S3 resources, set up S3 Lifecycle policies, and customize storage metrics.

21. Analytics and Metrics:

Amazon S3 Analytics helps you analyze storage access patterns to help manage data for cost savings, organize data, and configure fine-tuned data access controls.

22. Glacier Deep Archive Storage Class:

This is the lowest-cost storage class and supports long-term retention and digital preservation for data that may be accessed once or twice in a year.

23. Glacier Vault Lock:

This allows you to deploy and enforce compliance controls for individual S3 Glacier vaults with a vault lock policy.

24. S3 Batch Operations:

This feature allows you to perform large-scale operations on Amazon S3 objects.

25. S3 Access Points:

These are unique hostnames with dedicated access policies that describe how data can be accessed using that endpoint.

26. S3 Block Public Access:

This provides settings for access points, buckets, and accounts to help you manage public access to Amazon S3 resources.

27. Intelligent-Tiering:

This storage class is designed to optimize costs by automatically moving data to the most cost-effective access tier, without performance impact or operational overhead.

28. S3 Pricing Factors:

Understand how factors like storage class, management features, data transfer, and requests affect the cost.

29. S3 Replication Time Control (RTC):

This provides a SLA for the replication of S3 objects, and metrics and events to monitor the status of replication.

30. S3 Same-Region Replication (SRR):

This automatically replicates new S3 objects across different buckets in the same AWS Region.

31. S3 Ownership Controls:

These controls provide bucket-level controls over object ownership.

32. S3 Strong Read-after-Write Consistency:

This feature ensures all access to S3 objects, whether new or overwritten, delivers the latest version of the object.

33. S3 Object Lambda:

This allows you to add your own code to S3 GET requests to modify and process data as it is returned to an application.

34. Public Block Access Settings for Accounts:

These settings allow the root account in an organization to enforce specific public block settings across all the AWS accounts.

35. DataSync:

AWS DataSync makes it simple and fast to move large amounts of data online between on-premises storage and Amazon S3.

36. S3 Console:

The S3 Console is the graphical interface to manage your S3 resources.

37. Bucket Logging:

This captures all events related to bucket activities and allows you to track usage and troubleshoot issues.

38. Tagging and Cost Allocation:

With S3 object tagging, you can categorize storage, and AWS generates a cost report with usage and costs aggregated by your tags.

39. Transfer Acceleration:

This uses Amazon CloudFront's globally distributed edge locations to speed up data transfer to your buckets.

40. Cross-Origin Resource Sharing (CORS):

This allows you to control how your content is shared across different domains.

41. Object Locking:

Amazon S3 Object Lock functionality allows you to store objects using a "Write Once Read Many" (WORM) model, which can prevent accidental deletion.

42. S3 Delete Markers:

When you delete a versioned object, S3 inserts a delete marker, which becomes the current version of the object. This enables recovery of the deleted version if necessary.

43. AWS Organizations Integration:

This allows centralized control across your AWS accounts, helping you to manage policies for your accounts.

44. S3 Select:

This feature allows applications to retrieve only a subset of data from an object by using simple SQL expressions.

45. Compliance and Audit Readiness:

S3 supports numerous security standards and compliance certifications to help you meet regulatory requirements.

46. Amazon Macie Integration:

Amazon Macie uses machine learning to automatically discover, classify, and protect sensitive data stored in Amazon S3.

47. Amazon Athena and Amazon Redshift Spectrum Integration:

With these integrations, you can run sophisticated analytics directly on your S3 data.

48. AWS PrivateLink Support:

This allows private connections between VPCs, AWS services, and on-premises applications, securely on the Amazon network.

49. AWS Snowball Integration:

AWS Snowball is a data transport solution that accelerates moving terabytes to petabytes of data into and out of AWS using storage appliances.

50. Amazon S3 Outposts:

This extends S3 on-premises to provide a single, S3-compatible endpoint to store and retrieve data.