**AWS SA Pro Cert Notes**

**Introduction:**

* 2.7 minutes per question
* 170 minutes
* 80 questions
* scenario based questions
* FAQ’s – cloudformation, direct connect, kinesis, Lambda
* Read Exam blue print
* Take practice exam
* Make a note of technologies in the exam.
* Key Tech: KMS, STS, Import/Export, cloudformation, cloudfront, cloudhsm, cloudsearch, cloudwatch, datapipeline, direct connect, dynamodb, EBS, EC2, ELB, EMR, ElastiCache, Elastic Beanstalk, Elastic Transcoder, Glacier, IAM, Kinesis, Opsworks, RDS, Redshift, Route 53, S3, SES, SNS, SQS, SWF, Storage Gateway, VPC.
* Test 4 things – comprehension, knowledge of AWS, How you cope under pressure, time management.
* Exam strategy – look at sample questions on AWS.com. Read question first.

**High Availability & Business Continuity:**

**\*Read DR whitepaper - http://d36cz9buwru1tt.cloudfront.net/AWS\_Disaster\_Recovery.pdf**

* **DR** – preparing for and recovering from a disaster
  + Usually uses a N+1 approach
  + S3
  + Glacier
  + EBS
  + Direct Connect
  + Storage Gateway
  + EC2/VM Import Connector
  + Route53
  + Elastic Load Balancing
  + VPC
  + RDS
  + DynamoDB
  + Redshift
  + CloudFormation
  + ElasticBeanstalk
  + OpsWork
* **RTO** – time it takes to recover from an outage or disruption
* **RPO** – maximum period of time in which data might be lost from an IT service due to a major incident.
* **Backup/Restore** – Cheapest/longest RPO/RTO
* Pilot Light \
* Warm Standby |-> in SAA Notes
* MultiSite /
* **Know different RTO/RPOS for different AWS services.**
  + **S3**
  + **Glacier**
  + **EBS**
  + **DynamoDB –** offers cross region replication
  + **RDS –** can haveread replica in another region
  + **RedShift –** can copy to another region

DR & BC for Databases

* SQL Server – AlwaysOn Availability Groups, SQL Mirroring
* MySQL – Asynchronous replication
* Oracle – Oracle Data Guard, Oracle RAC
* RDS Multi-AZ Failover
* Automatic failover in case of
  + Loss of availability in **primary AZ**
  + Loss of connectivity to **primary DB**
  + Storage or host failure to **primary DB**
  + Software patching
  + Rebooting of **primary DB**
* Oracle, PostgreSQL, MySQL, and MariaDB use Amazon’s failover technology
* SQL Server DB uses **SQL Server mirroring**
* Amazon Aurora instances stores copies of the data in a DB cluster across multiple **AZ’s**
* MySQL
  + MySQL 5.6 (**NOT 5.1 or 5.5)**
  + Can use both **MyISAM and InnoDB** however InnoDB is supported by AWS
* PostgreSQL
  + PostgreSQL 9.3.5 or newer
* MariaDB
  + All current versions
* Oracle MSSQL
  + All current versions
* **Synchronous replication**
* If application doesn’t require transaction support, consider using DynamoDB if it doesn’t need ACID compliance (Atomicity, Consistency, Isolation, Durability)

**Read Replicas** – elastically scale out beyond the capacity constraints of a single DB instance for read-heavy database workloads

* Can only have **5 read replicas** of a DB at a moment.
* Can be created by console or CreateDBInstanceReadReplica API
* **Asynchronous replication**
* Used for i/o read heavy workloads or serving read traffic when source DB is unavailable **due to i/o suspension for backups or scheduled maintenance** or business reporting or data warehousing scenarios.
* Scaling = read replicas / BC = Multi-AZ
* If Multi-AZ not enabled
  + Snapshot will be of your **primary database** causing a brief I/O suspension for around 1 minute
* If Multi-AZ enabled
  + Snapshot will be of your **secondary database** and won’t have any performance hits.
* Get a new **DNS** endpoint address.
* Can promote a read replica but it does break the replication link between both DB’s.
* Can create read replica’s with RDS (MySQL, PostgreSQL, MariaDB) in different **regions**
  + Improve DR capabilities
  + Scaling read operations into a region closer to you
  + Make it easier to migrate from a DC to another.
  + Can create an Amazon Aurora DB cluster as read replicate in a different region.
  + Can be encrypted or unencrypted DB clusters
  + Read replica must be encrypted if the source DB is encrypted.
  + SQL Server and Oracle doesn’t have it in different regions.
  + **Read replicas can’t be Multi-AZ currently**
  + Can be built off Multi-AZ’s
  + Can have read replicas of read replicas with **only MySQL**
  + DB snapshots and automated backups **can’t** be taken of read replicas.

Storage Gateway

* How can I backup by data?
  + To S3 via **API calls**
  + To Storage Gateway and then replicate to S3.
* **File Interface**
  + NFS(Network File System) protocol stored in S3.
  + S3 buckets as mount points. File Server. Used data is cached on the gateway.
* **Volume interface**
  + **Gateway-Cached Volumes**
    - iSCSI based block
    - store primary data in S3 but retain IA locally
    - Unlimited amount of storage. **File size max is 5 TB**.
  + **Gateway-Stored Volumes**
    - iSCSI based block
    - Store primary data locally and asynchronously backup point in time snapshots to S3.
* **Tape Interface**
  + **Gateway-Virtual Tape Library**
    - iSCSI based virtual tape solution
    - Backed by S3 or VTS (Virtual Tape Shelf) and then backed by Glacier.
    - VTS and VTL retrieval times are different.

Snowball

* **Import/Export Disk** – move large amounts of data in and out of AWS using a portable storage device using Amazon’s high speed internal network.
* **Snowball** – PB scale transport to transfer data in and out of AWS. 80 TB snowball in all regions.
  + TPM enclosures
  + 256 encryptions.
  + Data is erased
  + Need a client to connect to the snowball.
* **Snowball Edge** – contain 100TB of storage with transfer device capability. Comes with compute capability whereas snowball does not. Lamba functions can be run from this.
* **Snowmobile** – Diesel truck can transfer up to 100PB of data.

Know snowball, import/export, what snowball can do(export to S3, etc)

Automated Backups

* Have
  + RDS
    - MySQL you need **InnoDB**
    - Performance hit Multi-AZ is not enabled
    - If you delete an instance, **ALL** automated backups are deleted
    - Manual DB snapshots will **NOT** be deleted.
    - Stored on S3
    - When you do a restore, you can change the engine **type (SQL Standard to SQL Enterprise)**
  + Elasticache (**redis only)**
    - Entire cluster is snapshotted
    - Snapshot will degrade performance
    - Set your snapshot window during the least buy part of the day.
    - Stored on **S3**
  + Redshift
    - Stored on **S3**
    - Enables automated backups of your data warehouse cluster with a **1-day** retention period
    - Incremental backup that has changed so most snapshots only use up a small amount of your free backup storage.
* Don’t Have
  + EC2
    - Backups **degrade** performance
    - Need automated script by python or CLI
    - **Incremental** backups
      * Only charged for incremental storage
      * Contains base snapshot
    - Stored in **S3**
    - Not automated.

Summary:

* Domain 1.0 :**HA and BC**
  + Demonstrate ability to architect the appropriate level of availability based on stakeholder requirements
  + 1.2 Demonstrate ability to implement **DR** for systems based on **RPO** and **RTO**
    - **Know how to backup each service and look closely at RPO/RTO**
    - Backup/Restore
    - Pilot Light
    - Warm Standby
    - Multi-Site
    - S3
      * **11 9s**
    - EBS
    - Archives
      * **3 hours or longer**
    - **DynamoDB – always go for this first.**
    - **RDS**
    - **Redshift**
  + 1.3 Determine appropriate use of **Multi-AZ vs. Multi-region architectures**
    - Know different use cases
      * **Multi-AZ** for DR and BC
      * **Read Replicas** for scaling
    - Know different types of replication
      * **Synchronous** replication for **Multi-AZ**
      * **Asynchronous** replication **for read replicas.**
    - App doesn’t require **ACID**, use **DynamoDB**.
    - Know **4** different Storage Gateways
      * **File Gateway**
      * **Volume Gateway**
        + Cached
        + Stored – store data as **EBS snapshots in S3.**
      * **Tape Gateway**
        + Virtual tapes access are **instantaneous**
        + Virtual tape shelf can take **24 hours.**
    - Encrypted using SSL for transit and encrypted at rest in **Amazon S3** using **AES-256**
    - Snapshots can be scheduled
    - Bandwidth can be throttled **(good for remote sites)**
    - **Know Snowball and Import/Export**
      * Snowball can **import** to S3 and **export** from S3.
  + 1.4 Demonstrate ability to implement self-healing capabilities
  + **15%** of the exam