AWS RDS Storage

# AWS RDS Storage

* RDS storage uses Elastic Block Store (EBS) volumes for database and log storage.
* RDS automatically stripes across multiple EBS volumes to enhance IOPS performance, depending on the amount of storage requested

## RDS Storage Types

* RDS storage provides three storage types: Magnetic, General Purpose (SSD), and Provisioned IOPS (input/output operations per second).
* These storage types differ in performance characteristics and price, which allows tailoring of storage performance and cost to the database needs
* MySQL, MariaDB, PostgreSQL, and Oracle RDS DB instances can be created with up to 6TB of storage and SQL Server RDS DB instances with up to 4TB of storage when using the Provisioned IOPS and General Purpose (SSD)  
  storage types.
* Existing MySQL, PostgreSQL, and Oracle RDS database instances can be scaled to these new database storage limits without any downtime.

### Magnetic (Standard)

* Magnetic storage, also called standard storage, offers cost-effective storage that is ideal for applications with light or burst I/O requirements.
* They deliver approximately 100 IOPS on average, with burst capability of up to hundreds of IOPS, and they can range in size from 5 GB to 3 TB, depending on the DB instance engine.
* Magnetic storage is not reserved for a single DB instance, so performance can vary greatly depending on the demands placed on shared resources by other customers.

### General Purpose (SSD)

* General purpose, SSD-backed storage, also called gp2, can provide faster access than disk-based storage.
* They can deliver single-digit millisecond latencies, with a base performance of 3 IOPS per Gigabyte (GB) and the ability to burst to 3,000 IOPS for extended periods of time up to a maximum of 10,000 PIOPS.
* Gp2 volumes can range in size from 5 GB to 6 TB for MySQL, MariaDB, PostgreSQL, and Oracle DB instances, and from 20 GB to 4 TB for SQL Server DB instances.
* Gp2 is excellent for small to medium-sized databases.

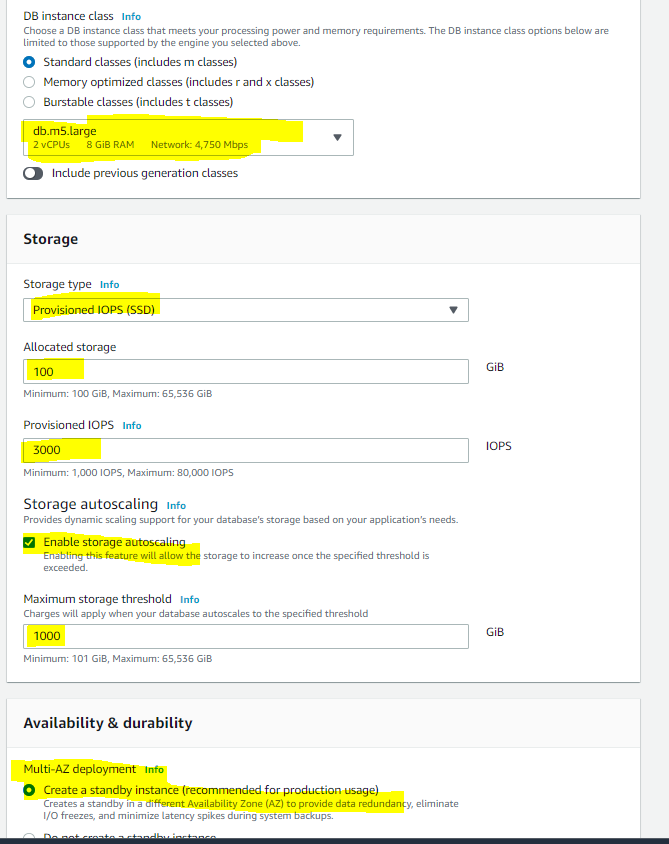
### Provisioned IOPS

* Provisioned IOPS storage is designed to meet the needs of I/O-intensive workloads, particularly database workloads, that are sensitive to storage performance and consistency in random access I/O throughput.
* Provisioned IOPS storage is a storage type that delivers fast, predictable, and consistent throughput performance.
* For any production application that requires fast and consistent I/O performance, Amazon recommends Provisioned IOPS (input/output operations per second) storage.
* Provisioned IOPS storage is optimized for I/O intensive, online transaction processing (OLTP) workloads that have consistent performance requirements.
* Provisioned IOPS helps performance tuning.
* Provisioned IOPS volumes can range in size from 100 GB to 6 TB for MySQL, MariaDB, PostgreSQL, and Oracle DB engines. SQL Server Express and Web editions can range in size from 100 GB to 4 TB, while SQL Server Standard and Enterprise editions can range in size from 200 GB to 4 TB.
* Dedicated IOPS rate and storage space allocation is specified, when a DB instance is created. RDS provisions that IOPS rate and storage for the lifetime of the DB instance or until its changed.
* RDS delivers within 10 percent of the provisioned IOPS performance 99.9 percent of the time over a given year.

For detailed explanation on refer post @ [EBS volume Types](https://jayendrapatil.com/aws-ebs-volume-types/)

### Adding Storage and Changing Storage Type

* DB instance can be modified to use additional storage and converted to a different storage type.
* **However, storage allocated for a DB instance cannot be decreased**
* MySQL, MariaDB, PostgreSQL, and Oracle DB instances can be scaled up for storage, which helps improve I/O capacity.
* **Storage capacity nor the type of storage for a SQL Server DB instance can be changed due to extensibility limitations of striped storage attached to a Windows Server environment.**
* During the scaling process, the DB instance will be available for reads and writes, but may experience performance degradation
* Adding storage may take several hours; the duration of the process depends on several factors such as load, storage size, storage type, amount of IOPS provisioned (if any), and number of prior scale storage operations.
* While storage is being added, nightly backups are suspended and no other RDS operations can take place, including modify, reboot, delete, create Read Replica, and create DB Snapshot



## Performance Metrics

* Amazon RDS provides several metrics that can be used to determine how the DB instance is performing.
  + **IOPS**
    - the number of I/O operations completed per second.
    - it is reported as the average IOPS for a given time interval.
    - RDS reports read and write IOPS separately on one-minute intervals.
    - Total IOPS is the sum of the read and write IOPS.
    - Typical values for IOPS range from zero to tens of thousands per second.
  + **Latency**
    - the elapsed time between the submission of an I/O request and its completion
    - it is reported as the average latency for a given time interval.
    - RDS reports read and write latency separately on one minute intervals in units of seconds.
    - Typical values for latency are in the millisecond (ms)
  + **Throughput**
    - the number of bytes per second transferred to or from disk
    - it is reported as the average throughput for a given time interval.
    - RDS reports read and write throughput separately on one minute intervals using units of megabytes per second (MB/s).
    - Typical values for throughput range from zero to the I/O channel’s maximum bandwidth.
  + **Queue Depth**
    - the number of I/O requests in the queue waiting to be serviced.
    - these are I/O requests that have been submitted by the application but have not been sent to the device because the device is busy servicing other I/O requests.
    - it is reported as the average queue depth for a given time interval.
    - RDS reports queue depth in one minute intervals. Typical values for queue depth range from zero to several hundred.
    - Time spent waiting in the queue is a component of Latency and  
      Service Time (not available as a metric).

## Amazon RDS Storage Facts

* First time a DB instance is started and accesses an area of disk for the first time, the process can take longer than all subsequent accesses to the same disk area. This is known as the “**first touch penalty**”. Once an area of disk has incurred the first touch penalty, that area of disk does not incur the penalty again for the life of the instance, even if the DB instance is rebooted, restarted, or the DB instance class changes. Note that a DB instance created from a snapshot, a point-in-time restore, or a read replica is a new instance and does incur this first touch penalty.
* RDS manages the DB instance and it reserves overhead space on the instance. While the amount of reserved storage varies by DB instance class and other factors, this reserved space can be as much as one or two percent of the total storage
* Provisioned IOPS provides a way to reserve I/O capacity by specifying IOPS. Like any other system capacity attribute, maximum throughput under load will be constrained by the resource that is consumed first, which could be IOPS, channel bandwidth, CPU, memory, or database internal resources.
* Current maximum channel bandwidth available is 4000 megabits per second (Mbps) full duplex. In terms of the read and write throughput metrics, this equates to about 210 megabytes per second (MB/s) in each direction. A perfectly balanced workload of 50% reads and 50% writes may attain a maximum combined throughput of 420 MB/s, which includes protocol overhead, so the actual data throughput may be less.
* Provisioned IOPS works with an I/O request size of 32 KB. Provisioned IOPS consumption is a linear function of I/O request size above 32 KB. An I/O request smaller than 32 KB is handled as one I/O; for e.g. 1000 16 KB I/O requests are treated the same as 1000 32 KB requests. I/O requests larger than 32 KB consume more than one I/O request; while, a 48 KB I/O request consumes 1.5 I/O requests of storage capacity; a 64 KB I/O request consumes 2 I/O requests

## Factors That Impact Storage Performance

* Several factors can affect the performance of a DB instance, such as instance configuration, I/O characteristics, and workload demand.
* System related activities also consume I/O capacity and may reduce database instance performance while in progress:
  + DB snapshot creation
  + Nightly backups
  + Multi-AZ peer creation
  + Read replica creation
  + Scaling storage
* System resources can constrain the throughput of a DB instance, but there can be other reasons for a bottleneck. Database could be the issue if :-
  + Channel throughput limit is not reached
  + Queue depths are consistently low
  + CPU utilization is under 80%
  + Free memory available
  + No swap activity
  + Plenty of free disk space
  + Application has dozens of threads all submitting transactions as fast as the database will take them, but there is clearly unused I/O capacity

## AWS Certification Exam Practice Questions

1. When should I choose Provisioned IOPS over Standard RDS storage?
   1. If you have batch-oriented workloads
   2. **If you use production online transaction processing (OLTP) workloads**
   3. If you have workloads that are not sensitive to consistent performance
2. Is decreasing the storage size of a DB Instance permitted?
   1. Depends on the RDMS used
   2. Yes
   3. **No**
3. Because of the extensibility limitations of striped storage attached to Windows Server, Amazon RDS does not currently support increasing storage on a \_\_\_\_\_ DB Instance.
   1. **SQL Server**
   2. MySQL
   3. Oracle
4. If I want to run a database in an Amazon instance, which is the most recommended Amazon storage option?
   1. Amazon Instance Storage
   2. **Amazon EBS**
   3. You can’t run a database inside an Amazon instance.
   4. Amazon S3
5. For each DB Instance class, what is the maximum size of associated storage capacity?
   1. 1TiB
   2. 2TiB
   3. 8TiB
   4. **16TiB**(New limit – 65TB. The limit keeps on changing so please check the latest always), IOPS max is 80,000