



# SKYSQL VS AMAZON RDS

May 2023

# TABLE OF CONTENTS

- 1 **INTRODUCTION**
- 2 **SKYSQL BENEFITS AND CAPABILITIES**
- 4 **RDS BENEFITS AND CAPABILITIES**
- 5 **SKYSQL ARCHITECTURE**
- 7 **RDS ARCHITECTURE**
- 8 **HIGH AVAILABILITY**
  - 8 STORAGE
  - 8 FAILOVER
  - 9 PROXY
  - 10 READ SCALING
  - 10 WRITE SCALING
- 11 **PRICING**
- 12 **EVALUATE SKYSQL FOR YOURSELF**

# INTRODUCTION

---

This paper compares and contrasts two options for cloud-based databases. Many IT operations and development teams looking to create new applications or migrate existing databases to the cloud consider Amazon's RDS offering. However, MariaDB SkySQL offers multiple topologies and technical and business advantages to customers in terms of ease of use, flexibility, availability, scalability, cost and overall capabilities.

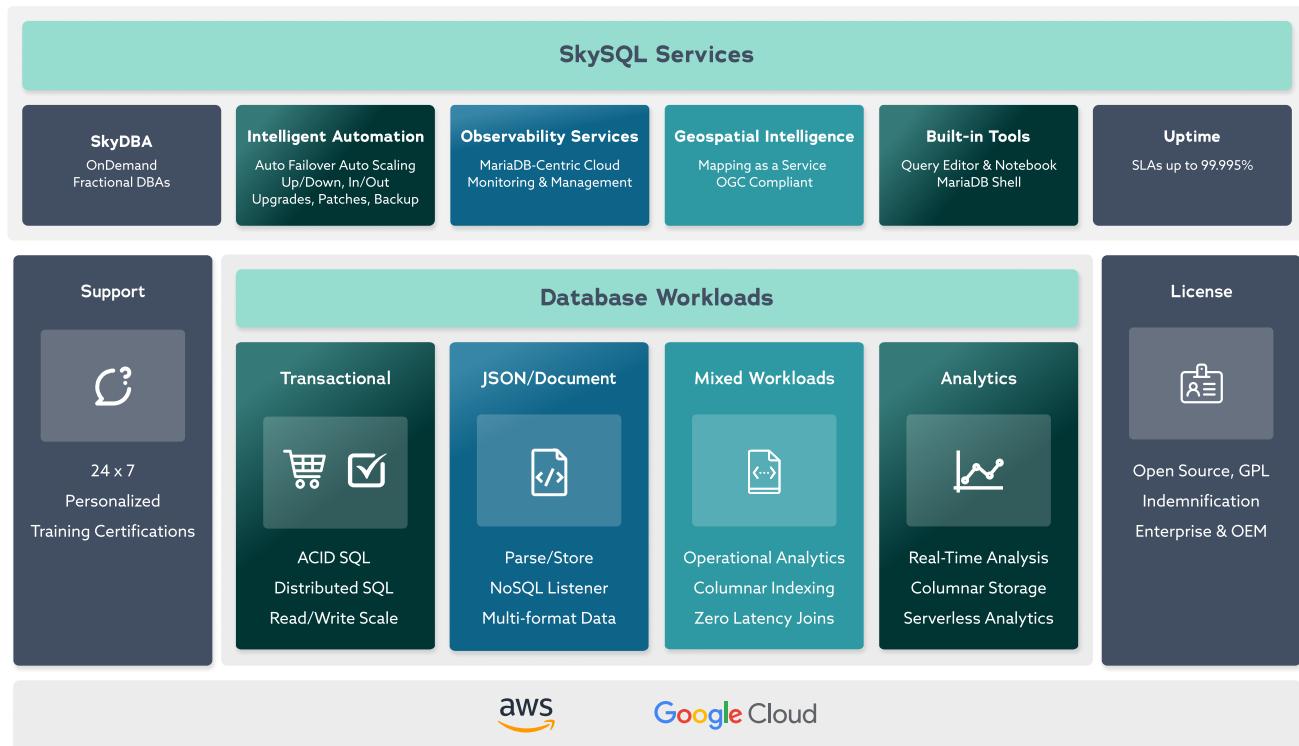
MariaDB SkySQL cloud database service is a second generation, fully managed, database-as-a-service (DBaaS) for everyday applications – perfect for small to medium size companies or departments within a large enterprise that need to start small and grow to any scale. SkySQL includes options for everything from purely transactional applications like reservation systems to more analysis-heavy workloads such as those found in finance or pricing. Most modern applications are mixed workloads, meaning neither purely transactional nor analytical; for example, online retail where product lookup would trigger a range of other product offerings before a purchase transaction. SkySQL runs typical read/write operations at high transactions per second (TPS) rates and as long-running analytical queries over terabytes of data with unparalleled price/performance. The power of SkySQL is that it runs multiple types of databases on multiple clouds with a single API and one administration console. SkySQL is an excellent option for applications with varied or unpredictable workloads and scale requirements, easy to use for IT and non-IT users, and affordable with transparent pricing.

Amazon Relational Database Service (Amazon RDS) allows users to deploy instances of PostgreSQL, MySQL, MariaDB Community Server, Oracle or SQL Server on Amazon's cloud infrastructure. RDS is optimized for read scaling of purely transactional workloads. The primary advantage of RDS is its ability to set up multiple brands of relational databases on Amazon Web Services (AWS). It is a good option for legacy applications with limited scale or availability needs that have predictable workloads that cannot undergo any schema or query optimization changes and do not require multicloud or mixed workload support.

# SKYSQL BENEFITS AND CAPABILITIES

MariaDB SkySQL is fully managed service offered by MariaDB plc, the originators of MySQL and MariaDB, and database experts focused on database innovation that harnesses native cloud technologies.. SkySQL enables teams to reduce the time and complexity of their intelligent application development and database deployment and simplify their overall cloud architecture. Rather than deploy multiple types of databases and supporting tools for monitoring and management with different management consoles and APIs, SkySQL provides one service available on both AWS and Google Cloud. SkySQL empowers users to ensure capacity meets workload demand while reducing the risk of downtime and cost overruns. Its ability to self-service or automatically scale up/down a database instance and scale out/in multiple nodes for an HA instance ensures performance and capacity meets customer demand and provides the flexibility to decrease capacity when not in use. Instead of overprovisioning and paying extra costs, SkySQL enables just-in-time (JIT) provisioning.

As shown in Figure 1 below, SkySQL includes several utilities that would be additional, separately paid for services from AWS or Google including MaxScale, an advanced database proxy and load balancer that includes a query editor for provisioned servers, SkySQL Observability Service (OS) for monitoring all your MariaDB databases no matter where they are deployed (on SkySQL, on premises, or Community Server-based including RDS), and Serverless Analytics powered by Apache Spark SQL with notebook based on Apache Zeppelin for zero-ETL analytics over SkySQL databases, Amazon S3 and other external data sources.



SkySQL supports modern cloud applications with 24x7x365 uptime to guarantee business continuity (up to 99.995% SLA for Power Tier) when things go wrong. This highly resilient architecture can tolerate multiple hardware, software and network outages without a loss of service. This resilience includes several redundancies and failure heuristics from the connection to the storage and backup of data.

Organizations choose SkySQL to reduce the cost and complexity of their architecture and support their mixed transactional and analytical workloads. As opposed to a series of point database solutions, SkySQL can be used as a single database that handles all of an application's needs in support of developers, engineers, architects, analysts and scientists.

SkySQL includes an intuitive GUI with self-service options for scaling instances up and down (vCPUs, memory, storage, IOPS) and scaling the number of replicas (nodes) up and down. Price estimates are provided for scaling to avoid surprises. Additionally, you can start and stop services based on various operational parameters with notifications of these events. All the above functionality is offered as an auto-scaling option, implemented with a policy engine that monitors resource utilization and notifies users if certain predefined limits are passed, and action is taken to scale up and scale down. Fail-safe mechanisms help prevent runaway scaling.

Now, more than ever, businesses must support a globally distributed remote workforce and customer base. SkySQL is available on both AWS and Google Cloud in many regions and delivers maximum performance by keeping data close to application endpoints and end users that need it.

Organizations choose SkySQL to develop new and more powerful applications and accelerate their cloud migration and multicloud efforts. Legacy database license costs and maintenance are expensive. They depend on SkySQL's compatibility with databases like Oracle, MySQL, PostgreSQL and Sybase to reduce the time and cost of migrating existing applications to the cloud.

The principal benefit of SkySQL is that it can handle multiple types of workloads on both Google Cloud and AWS with unparalleled scale and high availability.

# RDS BENEFITS AND CAPABILITIES

---

Amazon RDS provides a lower administrative burden compared to self-hosted databases. RDS can run several different relational database engines as is, without any efforts to enhance the features or functionality of either the proprietary or open source database engines it runs. RDS automates the rollout of patches and provides recommendations based on best practices by analyzing network and instance types.

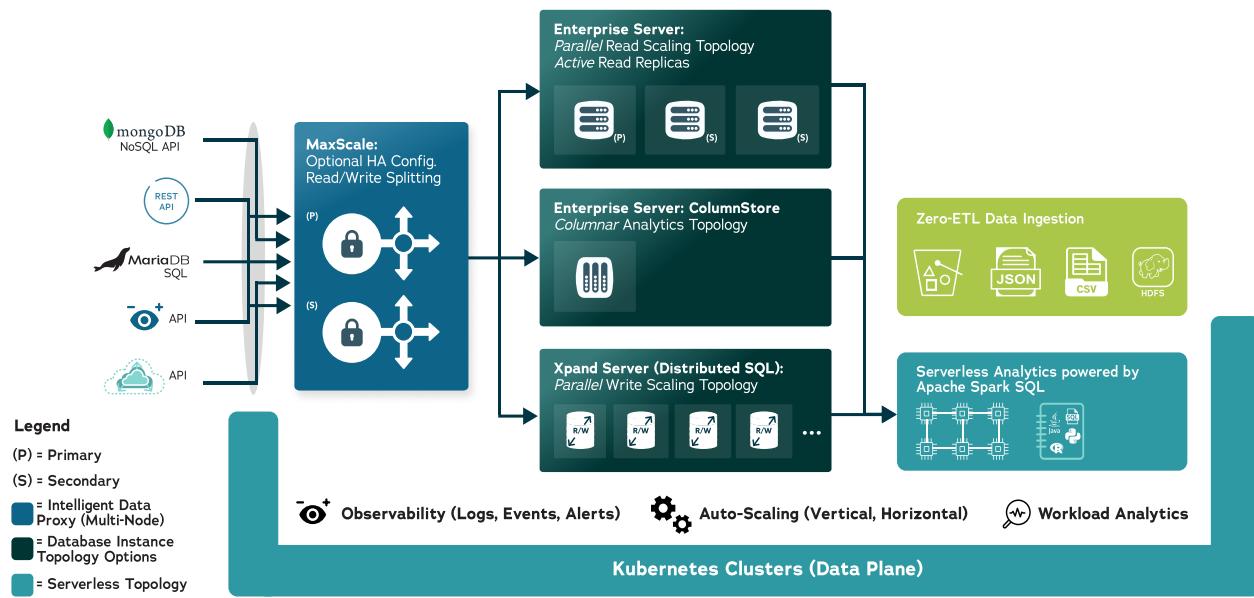
RDS provides push-button compute scaling up to its capabilities and claims that deployment completes in a few minutes. AWS has put a lot of effort into scaling database storage and is able to scale storage on the fly without downtime. This scale is done strictly at AWS's custom layer. It is similar to how non-cloud deployed databases are sometimes scaled at the SAN layer.

RDS does not offer a high availability setup, but by having a hot standby read replica in another availability zone (AZ), it is possible to failover within two minutes on many applications. These "multi-AZ deployments" cannot be used for scalability purposes. RDS also provides backups that are saved up to 35 days and user-generated snapshots.

The principal benefit of RDS compared to SkySQL is that it runs other transactional relational database engines besides MariaDB. However, the RDS database engines are meant only for transactional workloads and any analytical or mixed workloads would need separate analytics services such as Amazon Redshift. Similarly, for applications that make NoSQL calls to MongoDB, for example, a separate NoSQL service would be needed such as DynamoDB. In summary, for real-world complex applications or where there are multiple types of applications and a desire for a single platform, AWS RDS may not be the optimal solution.

# SKYSQL ARCHITECTURE

As depicted in Figure 2, MariaDB SkySQL enables novice and sophisticated database users alike to go through either the integrated and intuitive portal or use the SkySQL DBaaS API to create, manage and delete MariaDB databases. SkySQL runs on AWS and Google Cloud, and can be used on both clouds or in combination with MariaDB Enterprise and Xpand on premises or in private cloud for hybrid and multicloud setups. SkySQL is equipped to handle everyday applications through load balancing, failover and fault tolerance tools, and datastores with both analytical and transactional capabilities.



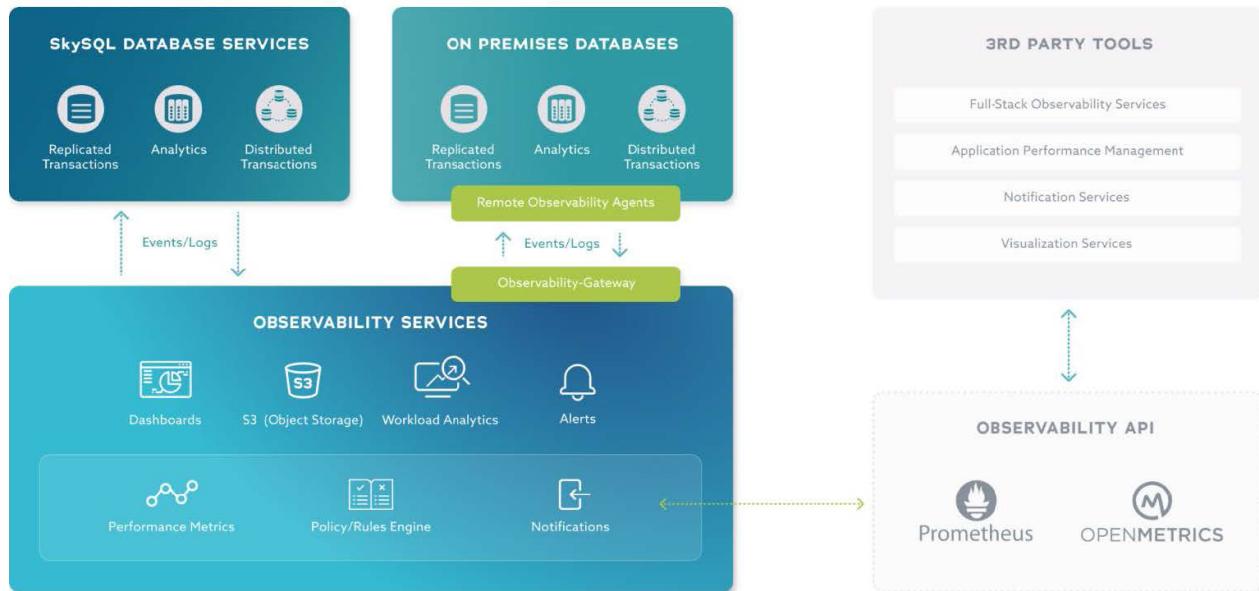
There are two provisioned databases offered on SkySQL. The first, MariaDB Enterprise Server, is a hardened, more robust version of MariaDB Community Server. It is best suited for traditional transactional applications. MariaDB Enterprise may be the best choice for robust applications with low-latency tolerance when they initially move to the cloud. The ColumnStore storage engine provides data warehousing functionality collocated with MariaDB Enterprise Server's InnoDB storage engine. SkySQL Serverless Analytics powered by Apache Spark SQL provides data science and ad hoc analytics on high-speed, zero-ETL ingestion and integration of structured or semi-structured data from any SkySQL database or other sources, including S3, HDFS or disparate files in various formats, including Avro, Parquet and .CSV.

The second is MariaDB Xpand, a distributed SQL database capable of parallel reads and writes that provides the best option in terms of scalability and reliability. Xpand enables continuous availability by slicing data into multiple partitions and ensuring that there are multiple replicas of each slice. Xpand distributed SQL scales linearly with self-healing and self-tuning, adding and removing capacity as demand requires, ensuring performance and availability while saving on the cost of overprovisioning.

Regardless of which SkySQL database service is launched, a single IP address for transparent availability to applications can be connected to a distributed or clustered database just as if they were connecting to a single database. When a multi-node configuration is spun up, MaxScale is automatically brought online with its own redundancy to avoid having it as a single point of failure either.

SkySQL uses Kubernetes and an advanced operator to perform self-healing when cloud instances fail, automatically recovering standalone databases and returning replicated databases to full capacity. If a cloud instance fails, Kubernetes will create a new node. Through the combination of MaxScale and Kubernetes, SkySQL provides load balancing, failover and read-write splitting. When a database node goes down SkySQL will automatically replay the failed transactions on the new primary instance. This is done in seconds and is invisible to the client application. SkySQL can deploy up to 1152 CPUs with a maximum of 512 GiB of RAM, 9TB of storage and 65,000 IOPS.

The MariaDB SkySQL Observability Service , shown in Figure 3 below, is a fully managed, cloud-native service purpose-built for MariaDB databases and services running on premises, on AWS or Google Cloud, or a hybrid on-prem and cloud. SkySQL fully instruments MariaDB Community Server and Enterprise Server, MariaDB Xpand for distributed SQL, and their storage engines – including InnoDB and ColumnStore. SkySQL OS delivers real-time performance and capacity insights through templated, customizable dashboards fed by rules-based alerts on a full range of metrics covering all aspects of MariaDB operations, down to individual queries. SkySQL OS is free for SkySQL databases and all databases covered under a MariaDB subscription. You can also try SkySQL OS with your existing MariaDB Community Server deployments at \$45 monthly per database node.



# RDS ARCHITECTURE

---

Amazon RDS allows for the setup of a single primary read-write instance with five read replicas across multiple zones in a single region. RDS scales to a maximum of 64TB, 64 vCPUs and 1,024 GiB of RAM.

RDS for MariaDB architecture uses a standard instance of MariaDB Community Server with either general-purpose or provisioned IOPS. When using general-purpose AWS claims to provide three IOPS per provisioned gigabyte of storage and the ability to burst up to 3,000 IOPS above this baseline. For provisioned IOPS, RDS allows up to 80,000 IOPS per instance.

Load balancing and failover are provided by [Amazon RDS Proxy](#) at an additional cost. Replication is handled by the underlying database technology (in the case of MariaDB). RDS [can handle disaster recovery failover to a standby replica on another availability zone](#). AWS claims this failover takes between 60 and 120 seconds. However, AWS RDS uses DNS and, as a result, failover can take several minutes over and above the listed two-minute failover.

Monitoring RDS database instances can be done through Amazon CloudWatch and Performance Insights, separate services. Both have a free tier for limited logging and storage periods, seven-day performance history retained and a few customized metrics and alarms above which they become additional paid services. The advantage of these services is that they work with the RDS CLI and plug into the AWS Management Console.

RDS is a good choice for applications that do not require 24x7x365 availability and do not anticipate scaling to large data sizes or large numbers of transactions. It may also be a good choice for firms with a number of small applications that use Microsoft, Oracle, PostgreSQL and MySQL that don't want to undertake any schema or query migration. RDS is only available on AWS so it is not a good option for teams utilizing a multicloud strategy.

# HIGH AVAILABILITY

---

## Storage

RDS databases use an elastic storage volume (Amazon EBS) and replicate using the replication capability of the underlying database (this is different from Amazon Aurora which uses the storage layer for replication). RDS can synchronously replicate between the primary and an additional standby node in a different AZ. The result is there are two copies of data. In addition to the standby node, RDS can replicate from the primary to 15 additional read replicas. Regardless of the number of read replicas, RDS replicates data to three AZs at the storage layer.

- If the primary or primary AZ fails, inbound transactions are terminated and must be retried. The standby node is promoted as the new primary and a new secondary is created in the original AZ.
- If two AZs fail, applications can neither read nor write data.

SkySQL databases use persistent SSDs on Google Cloud and EBS for storage on AWS along with partitioning. By default, there are three copies of data with one copy per AZ. However, the number of copies is configurable from as little as one per region to three like RDS, to as many per AZ as there are database instances.

Unlike RDS, SkySQL databases can continue to read and write data if two AZs fail. If a single zone fails, its database instances will automatically be migrated to one of the remaining zones. If two zones fail, their database instances will automatically be migrated to the remaining zone.

## Summary

	RDS	SkySQL
Data redundancy - min	0	1
Data redundancy - max	1	5
Zone failures tolerated	1	2

## Failover

RDS databases support one primary and multiple read replicas.

RDS databases consist of a primary for reading and writing data (accessed via the cluster/writer endpoint) and up to 15 replicas for reading data (accessed via the reader endpoint). If the primary fails, RDS promotes one of the replicas and updates both cluster/writer and reader endpoints.

RDS multi-zone databases allow a standby synchronously replicated from the primary. There is a performance cost to this addition and standby instances cannot be used as read replicas.

SkySQL Enterprise Server databases consist of one to five database nodes in a single primary and up to four read replicas. SkySQL can also deploy Xpand databases, consisting of three to 18 database instances, all capable of reading and writing data. Like RDS, SkySQL databases are accessed via a read/write endpoint that is automatically

updated when database instances are added or removed. However, unlike RDS, SkySQL databases provide continuous availability because there are no standbys, thus no need to wait for a standby replica to be promoted. If a database instance fails, the read/write endpoint will stop using it and continue with connections to the remaining database instances in fractions of a second.

## Summary

	RDS	SkySQL
Continuous availability	Yes	Yes
Maximum database instances	1 primary, 1 standby, 15 replicas	18 primaries

## Proxy

RDS Proxy, a separate service, can be used to speed up automatic failover and simplify connection management.

Applications no longer have to wait for DNS changes after an automatic failover. If the primary fails, RDS Proxy switches to the newly promoted primary. In addition, RDS Proxy preserves idle client connections (rather than closing them).

When used with RDS, applications no longer have to implement connection management. RDS Proxy treats one instance as active and the other as a hot standby. If the active database instance fails, RDS Proxy switches to the hot standby.

SkySQL database deployments include MariaDB MaxScale, an advanced database proxy for MariaDB Enterprise Server and Xpand. Like RDS Proxy, it provides applications with a single endpoint. If a database instance fails, MaxScale stops routing queries to it and, because SkySQL databases are primary with active read replicas in the case of Enterprise Server database instances and multi-primary in the case of Xpand database instances, continues routing queries to the remaining database instances.

However, MaxScale goes further by providing zero interruption failover (ZIF). Rather than closing client connections, dropping sessions and rolling back in-flight transactions, MaxScale will migrate connections to a new database instance, re-create sessions and replay any in-flight transactions – completely hiding failures from applications and allowing them to continue without interruption.

## Summary

	RDS	SkySQL
Single endpoint	Yes	Yes
Connection migration	Partial (idle)	Full (all)
Session restore	No	Yes
Transaction replay	No	Yes

## Read scaling

RDS supports read scaling with up to 15 replicas, and up to 64TB per node of storage. However, data is asynchronously replicated to read replicas and may not be transactionally valid. RDS does not make effective use of the memory on additional nodes in terms of caching or other features. Moreover, read replicas can have a negative impact on performance compared to a single instance when statement caches and other optimizations are considered.

SkySQL Enterprise Server uses MaxScale to provide load balancing and read/write splitting, removing the primary node as a bottleneck in parallel reads and thereby improving performance. Alternately, SkySQL running Xpand multi-primary (no replicas) distributes tables and indexes, so each database instance caches a different subset of data in memory. It effectively creates a distributed cache by combining the memory of multiple database instances, and because every database instance can read data (cached or not), adding additional database instances (i.e., scaling out) increases read performance. It not only improves read throughput, but it also improves read latency as well – increasing the total amount of data cached in memory so reads do not require disk I/O.

### Summary

	RDS	SkySQL
Max read instances	15	18
Distributed caching	No	Yes

## Write scaling

RDS is limited to a single write node and is limited to scaling up if higher write performance is needed (i.e., higher throughput and/or lower latency). Once the largest instance size is used, it is no longer possible to scale writes with RDS. The maximum instance size for RDS is 64 vCPU.

SkySQL enables several different topologies including distributed SQL which will allow 18 read/write nodes and enables the largest instance size to be used.

	RDS	SkySQL
Max write instances	1	18
Distributed SQL	No	Yes
vCPUs	128	1152

# PRICING

RDS and SkySQL have comparable compute costs, although SkySQL instance pricing may be slightly lower or higher depending on region, but differ dramatically in database proxy, storage and backup costs. SkySQL includes a powerful database proxy, MaxScale, ensuring automatic failover within seconds, while AWS RDS failover "typically takes 60-120 seconds," but can vary greatly. MaxScale also performs automatic replay of transactions that could not be completed due to an outage, and applications never become aware of a problem and simply continue without dropping user connections, shopping carts, etc. The proxy also has read/write split capability that maximizes read scale in primary/secondary topologies.

RDS charges \$0.115 per GB a month for storage as well as \$0.25 per one million I/Os per month. SkySQL simply charges zonal SSD list pricing of \$0.17 per GB a month for storage in Google Cloud and io1 EBS storage list pricing of \$0.125/GB per month plus \$0.065/provisioned IOPS/month in AWS. In addition, RDS includes only 700GB of backup storage for free then charges \$0.095 per GB a month, whereas SkySQL provides one week of backups for free.

	RDS (Oregon)	SkySQL in AWS (globally)	SkySQL in GCP (globally)
Compute:			
2 vCPUs x 4GB 1 primary, 2 replicas io1: 100GB SSD, 1000 IOPS	db.t3.medium + RDS proxy \$670.46 per month	Sky-2x4 with Scale Fabric \$664.95 per month	Sky-2x8 with Scale Fabric \$596.69 per month
Storage	\$0.125/GB per month	\$0.125/GB per month	\$0.17/GB per month (zonal SSD)
I/Os	\$0.20 per I/O-month	\$0.065/IOPS-month	Free
Backups (for one week)	Free up to 700GB additional \$0.095/GB per month	Free	Free
Replication (between AZs)	Free	Free	Free
Proxy	\$10.95/vCPU per month	Included with Scale Fabric	Included with Scale Fabric

While at first look SkySQL may appear more expensive if comparing only instance pricing, it has an overall lower total cost of ownership (TCO) and is all-inclusive, not charging extra for many items that carry a separate price tag in AWS. In addition, SkySQL pricing is more predictable and transparent with fewer variables that often lead to surprise charges for AWS customers.

# EVALUATE SKYSQL FOR YOURSELF

---

SkySQL provides the functionality you need for scalable, highly available, consistent data in the cloud, and is a cost-effective solution. We invite you to [try SkySQL today with a \\$500 credit](#). For larger applications, we're happy to work with you to find a trial solution that works for you. [Drop us a line.](#)



# THE SKY IS TRULY THE LIMIT

**SKYSQL.** The ultimate MariaDB Cloud is here.

## EASY TO USE + POWERFUL

SkySQL is the first database-as-a-service (DBaaS) to unlock the full power of MariaDB in the cloud. Engineered for mission-critical applications, SkySQL does away with common limitations and combines automation with the expertise and experience necessary to support production deployments in the cloud.

Get started with a \$500 credit:

[mariadb.com/skyview](http://mariadb.com/skyview)