

[Cloud SQL](https://cloud.google.com/sql/) (<https://cloud.google.com/sql/>)

[Documentation](https://cloud.google.com/sql/docs/) (<https://cloud.google.com/sql/docs/>)

[SQL Server](https://cloud.google.com/sql/docs/sqlserver/) (<https://cloud.google.com/sql/docs/sqlserver/>) [Guides](#)

# Overview of the high availability configuration

---

[MySQL](https://cloud.google.com/sql/docs/mysql/high-availability) (<https://cloud.google.com/sql/docs/mysql/high-availability>) | [PostgreSQL](https://cloud.google.com/sql/docs/postgres/high-availability) (<https://cloud.google.com/sql/docs/postgres/high-availability>) | **SQL Server**

## Beta

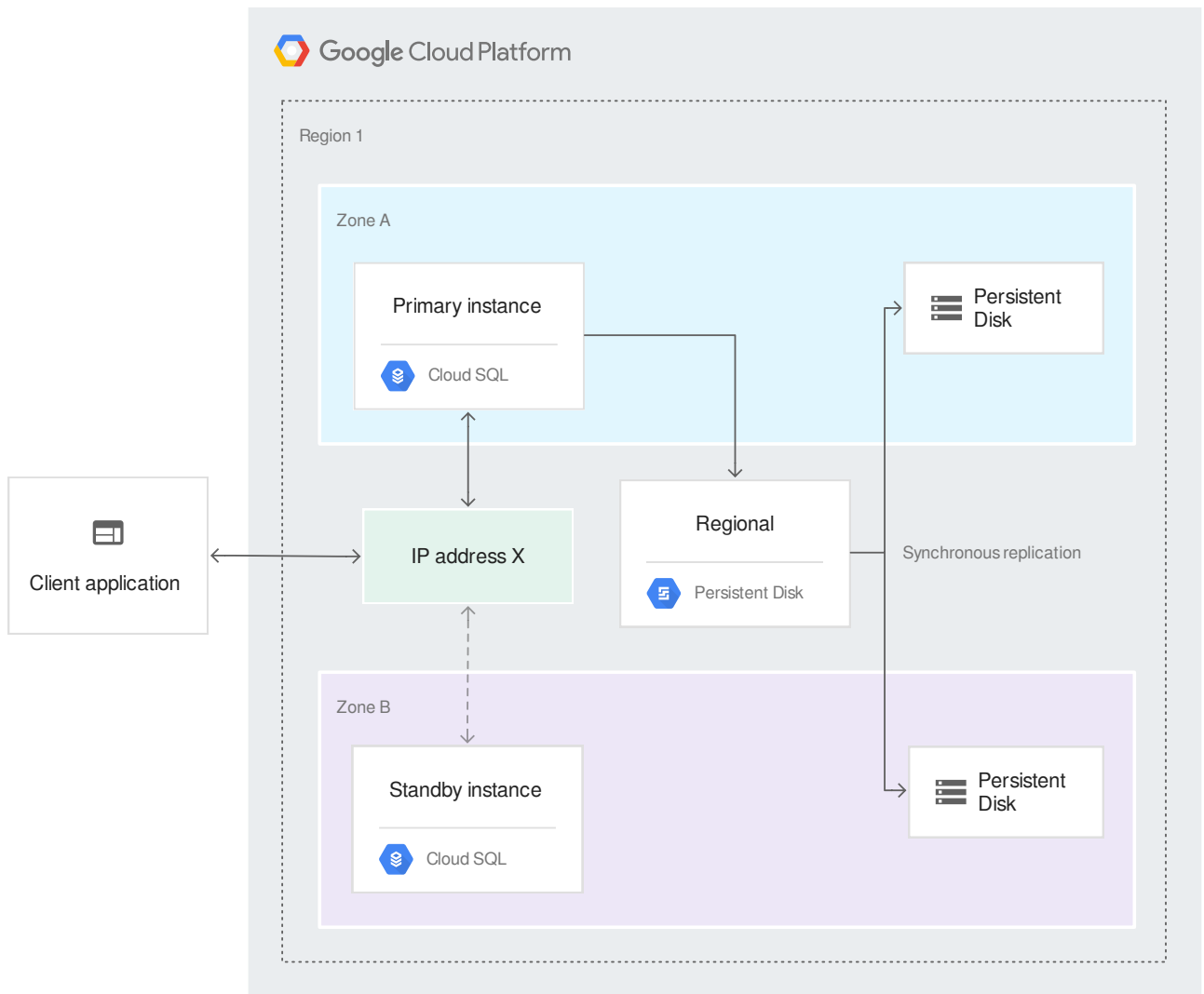
This feature is in a pre-release state and might change or have limited support. For more information, see the [product launch stages](https://cloud.google.com/products/#product-launch-stages) (<https://cloud.google.com/products/#product-launch-stages>).

This page is an overview of the high availability (HA) configuration for Cloud SQL instances. To configure a new instance for HA, or to enable HA on an existing instance, see [Enabling and disabling high availability on an instance](https://cloud.google.com/sql/docs/sqlserver/configure-ha) (<https://cloud.google.com/sql/docs/sqlserver/configure-ha>).

## HA configuration overview

The HA configuration, sometimes called a *cluster*, provides data redundancy. A Cloud SQL instance configured for HA is also called a *regional instance* and is located in a primary and secondary zone within the configured region. Within a regional instance, the configuration is made up of a *primary instance* (master) and a *standby instance*. Through [synchronous replication](https://cloud.google.com/compute/docs/disks/#repds) (<https://cloud.google.com/compute/docs/disks/#repds>) to each zone's persistent disk, all writes made to the primary instance are also made to the standby instance. In the event of an instance or zone failure, this configuration reduces downtime, and your data continues to be available to client applications.

Regional PD support for Cloud SQL and the Cloud SQL HA configuration are GA with full [SLA](https://cloud.google.com/sql/sla) (<https://cloud.google.com/sql/sla>) coverage. An HA-configured instance is charged at double the price of a standalone instance. This includes CPU, RAM, and storage. For more information, see [the pricing page](https://cloud.google.com/sql/docs/sqlserver/pricing) (<https://cloud.google.com/sql/docs/sqlserver/pricing>).

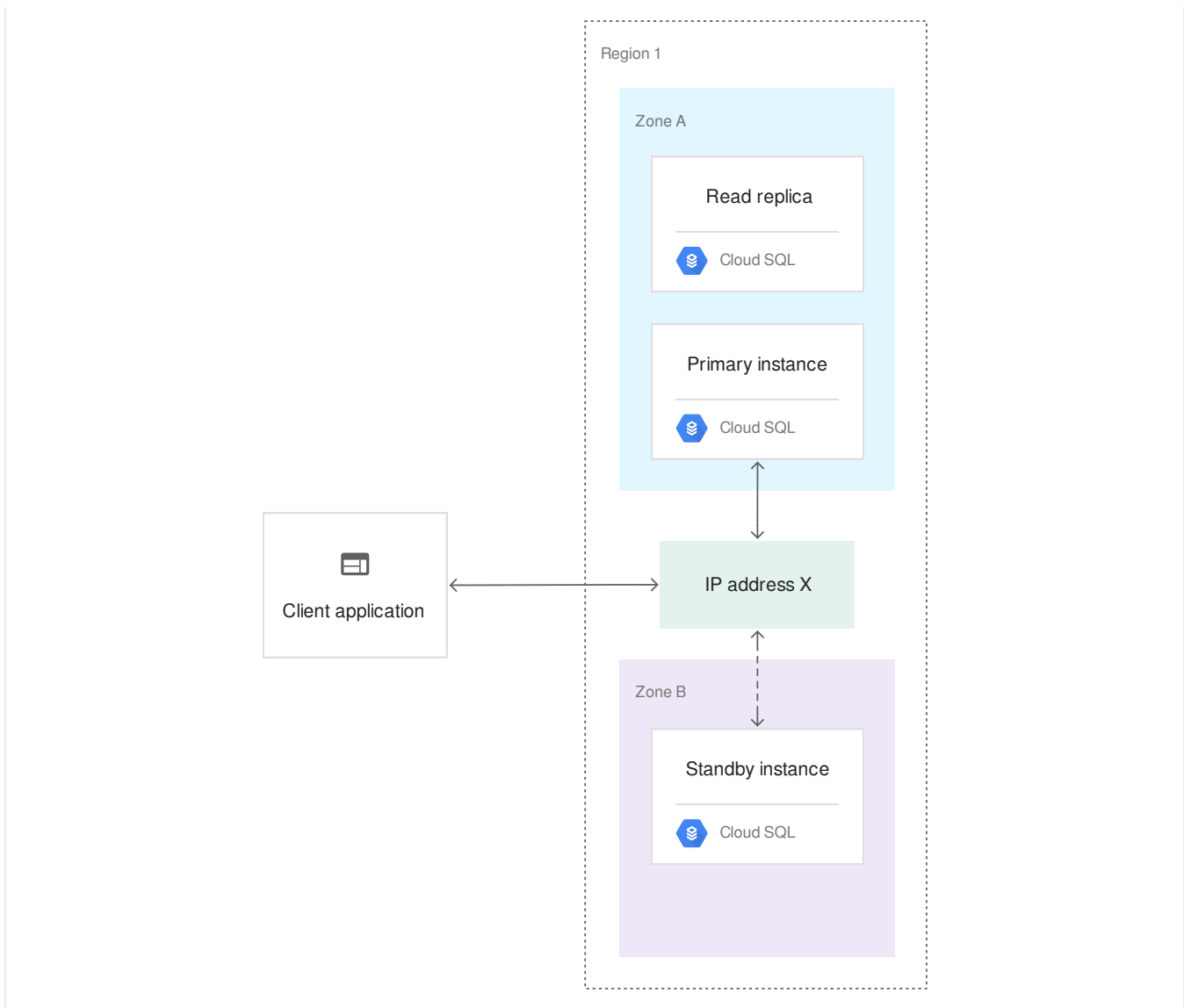


## Failover overview

If an HA-configured instance becomes unresponsive, Cloud SQL automatically switches to serving data from the standby instance. This is called a *failover*. To see if failover has occurred, check your operation log's failover history.

Click the tabs to see how failover affects your instance.

NORMAL	FAILOVER	FAILBACK



## Process

The following process occurs:

- The primary instance or zone fails.

Each second, the primary instance writes to a system database as a heartbeat signal. If multiple heartbeats aren't detected, failover is initiated. This occurs if the primary instance is unresponsive for approximately 60 seconds or the zone containing the primary instance experiences an outage.

- The standby instance now serves data upon reconnection.

Through a shared static IP address with the primary instance, the standby instance now serves data from the secondary zone.

## Requirements

For Cloud SQL to allow a failover, the configuration must meet the following requirements:

- The primary instance must be in a normal operating state (not stopped, undergoing maintenance, or performing a long-running operation).
- The secondary zone and standby instance must both be in a healthy state. When the standby instance is unhealthy and/or replication to the secondary zone is interrupted, failover operations are blocked. After Cloud SQL repairs the standby instance and the secondary zone is available, replication resumes and Cloud SQL allows failover.

★ **Note:** If both the primary and standby instances are unhealthy, Cloud SQL does not allow failover.

## Backups and restores

Configuring an instance for HA does not affect your need for backups or how you create them.

## Applications and instances

There is no difference in working with non-HA and HA instances, so your application does not need to be configured in any particular way. When failover occurs, any existing connections to the primary instance and read replicas are closed, and it will take approximately 2-3 minutes for connections to be reestablished. Your application reconnects using the same connection string or IP address, so you do not need to update your application after failover.

To see exactly how your applications are affected by failover, you should manually initiate failover (<https://cloud.google.com/sql/docs/sqlserver/configure-ha#test>).

## Maintenance downtime

Maintenance events affect primary instances configured with HA in the same way as any other instance. You can expect primary instances to be down during this time. To minimize impact to your service, you can set a maintenance window to control when downtime occurs.

## What's next

- [Enabling and disabling high availability on an instance](https://cloud.google.com/sql/docs/sqlserver/configure-ha) (https://cloud.google.com/sql/docs/sqlserver/configure-ha).
- [Initiate failover](https://cloud.google.com/sql/docs/sqlserver/configure-ha#test) (https://cloud.google.com/sql/docs/sqlserver/configure-ha#test).
- [Learn more about managing your database connections](https://cloud.google.com/sql/faq#connections) (https://cloud.google.com/sql/faq#connections).
- [Learn more about regions and zones](https://cloud.google.com/sql/docs/instance-locations) (https://cloud.google.com/sql/docs/instance-locations) in Cloud SQL.

---

*Except as otherwise noted, the content of this page is licensed under the [Creative Commons Attribution 4.0 License](https://creativecommons.org/licenses/by/4.0/) (https://creativecommons.org/licenses/by/4.0/), and code samples are licensed under the [Apache 2.0 License](https://www.apache.org/licenses/LICENSE-2.0) (https://www.apache.org/licenses/LICENSE-2.0). For details, see our [Site Policies](https://developers.google.com/terms/site-policies) (https://developers.google.com/terms/site-policies). Java is a registered trademark of Oracle and/or its affiliates.*

*Last updated January 22, 2020.*