# GENESYS

# Setting up Genesys Multicloud CX Private Edition

## High availability and disaster recovery

7/31/2025

# Contents

High availability (HA) and disaster recovery (DR) are two important factors in establishing a resilient infrastructure. This article describes the two supported architecture types for HA and DR, as well as the HA and DR modes supported by the private edition services.

## Related documentation:
- 
- 

## RSS:

- For private edition

Modern software environments demand two major types of agility:

- The ability to autoscale—that is, to rapidly increase processing power to handle a growth in interaction volume
- **Resiliency**—that is, the ability to fail over after losing one or more services—or even a whole data center or region

The second type of agility—the ability to bounce back from a failure—is broadly divided into two types of activity, each with its own requirements:

- **High availability (HA)** is the use of built-in redundancy to handle the failure of a service within a single region or data center
- **Disaster recovery (DR)** is the ability to continue processing after losing a whole region or data center, by failing over to another region or data center

Genesys Multicloud CX private edition allows you to set up a highly available and resilient infrastructure whether you are using a cloud deployment or hosting it in a private data center.

Note, however, that these two types of deployments require somewhat different architectures, as discussed below.

## Important
Before you continue, review the platform section of the private edition architecture page for an in-depth discussion of key components of the private edition architecture, such as **unit pairs** and **Availability Zones**.

## Key architectural distinctions

Both the cloud and private data center architectures use multiple geographical regions that are hosted within a single unit group. And in both types of environment, all of the unit pairs in a deployment are fully meshed with each other.

But the cloud deployment's ability to use Availability Zones makes its redundancy features more robust, as shown in the following table:

| Deployment type | Redundancy type | Characteristics |
| --- | --- | --- |
| Cloud | Availability Zones within regions | Multiple data centers in a small geographical area—can share a single Kubernetes cluster |
| Private data center | Physically discrete data centers | Data centers cannot share a Kubernetes cluster |

## Cloud architecture

One of the most important advantages of a cloud architecture is the enhanced redundancy through the use of Availability Zones (AZs). As discussed in the platform section of the private edition architecture page, an AZ is a discrete location within a region that is designed to operate independently from the other Availability Zones in that region. Because of this separation, any given Availability Zone is unlikely to be affected by failures in other Availability Zones.

In the cloud architecture, high availability is achieved by deploying instances within different Availability Zones.

> ### Important
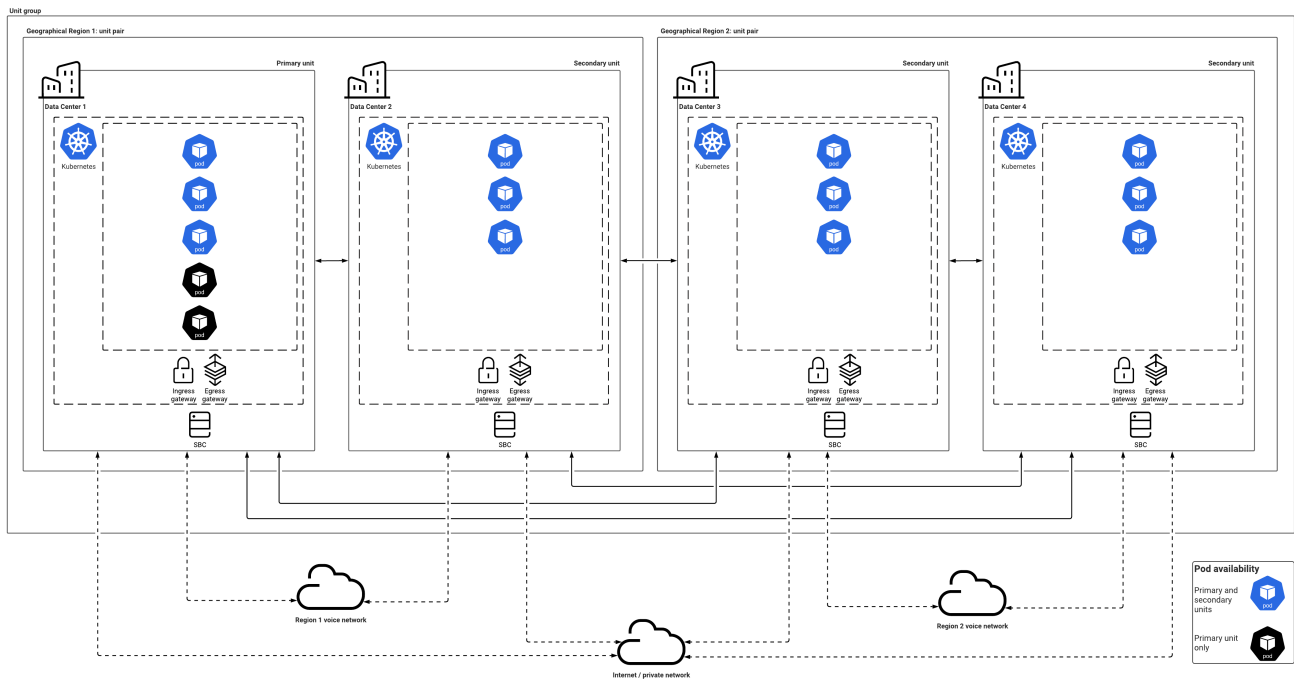> Black pod icons indicate services that can only be hosted in the primary unit.
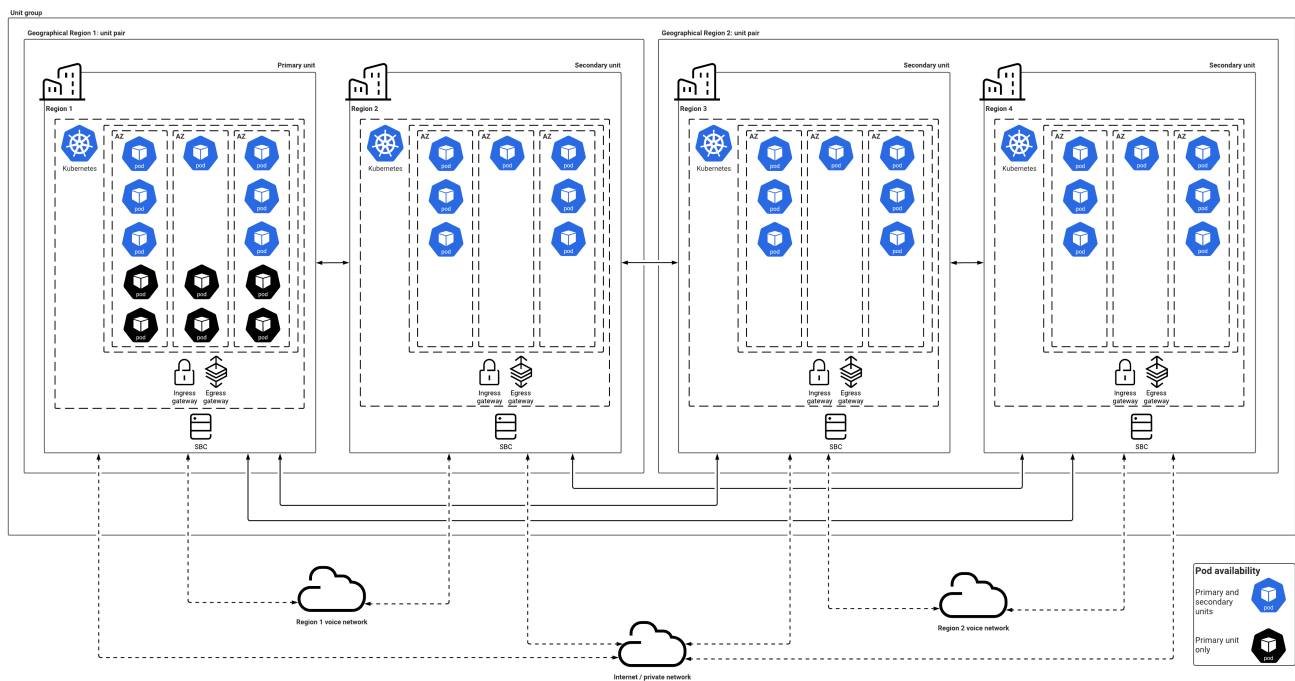
# Private data center architecture

> **Important**
>
> Black pod icons indicate services that can only be hosted in the primary unit.

## Planning for high availability

Private edition services scale automatically to meet demand. And when a service fails, private edition's high availability features enable an auto restart of that service.

For first-time deployments, you must plan:

- The number of nodes
- The number of pods that each node must run in your Kubernetes cluster

In order to reduce service disruptions, Genesys recommends that you run a minimum of three pod replicas for each service. Use the Sizing Calculator to determine the infrastructure requirements for achieving high availability in your contact center.

## Resiliency modes of private edition services

### High availability modes

Private edition services maintain high availability by using the following modes:

> ### Important
> Some services support more than one HA mode.

High availability modes

| Mode | Description |
|---|---|
| N = 2 (active-active) | The service is running on two nodes simultaneously. If one fails, the other takes over. |
| N = 1 (singleton) | The service is running on a single node. If that node fails, a new node is started to take over processing for that service. |
| N = N (N+1) | The service normally runs on N nodes. If a node fails, a new node is started to replace the failing node. |
| Cron jobs | Some services run as cron jobs, meaning that normal HA is not applicable |

## Disaster recovery modes

Private edition services achieve disaster recovery by using the following modes:

Disaster recovery modes

| Mode | Description |
|---|---|
| Active spare | A complete production replica is in place and serves traffic during normal operations |
| Limited active spare | A complete production replica is in place and serves traffic during normal operations, but the data is only used in case of disaster |
| Pilot light | The bare minimum configuration is in place to get the system back within a short time period. For example, there might be a read replica for a database. Application servers and web servers are deployed after the disaster. |
| Not supported | Disaster recovery is not supported for this service |

## Modes for each service

The following table displays the high availability and disaster recovery modes used by private edition services.

> ### Important
> Disaster recovery is not supported for services that are only available in the primary

unit.

| Service & Included Services | High Availability | Disaster Recovery | Where can you host this service? |
|---|---|---|---|
| | N = N (N+1) | Active-spare | Primary or secondary unit |
| — Designer | N = N (N+1)<br>*Or*<br>N = 2 (active-active) | Pilot light | Primary unit only |
| — Designer Application Server | N = N (N+1)<br>*Or*<br>N = 2 (active-active) | Active-spare | Primary or secondary unit |
| — Voice Platform Configuration Server | N = 1 (singleton) | Active-spare | Primary or secondary unit |
| — Voice Platform Media Control Platform | N = N (N+1) | Active-spare | Primary or secondary unit |
| — Voice Platform Reporting Server | N = 1 (singleton) | Active-spare | Primary or secondary unit |
| — Voice Platform Resource Manager | N = 2 (active-active) | Active-spare | Primary or secondary unit |
| — Voice Platform Service Discovery | N = 1 (singleton) | Active-spare | Primary or secondary unit |
| | N = N (N+1) | Active-spare | Primary or secondary unit |
| | N = N (N+1) | Active-spare | Primary or secondary unit |
| | N = N (N+1) | Not supported | Primary unit only |
| | N = N (N+1) | Not supported | Primary unit only |
| | N = N (N+1) | Not supported | Primary unit only |
| | IWD Data Mart is a Cronjob that runs on a per-tenant basis, so High Availability (HA) is not applicable. | | |
| | N = 1 (singleton) | Not supported | Primary unit only |

| Service & Included Services | High Availability | Disaster Recovery | Where can you host this service? |
|---|---|---|---|
| | N = N (N+1) | Not supported | Primary or secondary unit |
| — Genesys CX Insights | N = 2 (active-active) | Not supported | Primary unit only |
| — Reporting and Analytics Aggregates | N = 1 (singleton) | Limited active spare | Primary or secondary unit |
| | N = 1 (singleton) | Limited active spare | Primary or secondary unit |
| | N = 2 (active-active) | Pilot light | Primary unit only |
| | N = N (N+1) | Active-spare | Primary or secondary unit |
| | N = 1 (singleton) | Active-spare | Primary unit only |
| | N = N (N+1) | Active-spare | Primary or secondary unit |
| | N = N (N+1) | Not supported | Primary unit only |
| | N = N (N+1) | Active-spare | Primary or secondary unit |
| | N = N (N+1) | Active-spare | Primary or secondary unit |