



# Multiconnect

[/v1/account/shards](#)

The standard WhatsApp Business API Client solution runs on a single Docker container. In case you want to split the load and have multiple servers send and receive messages to WhatsApp, you can make use of our multiconnect solution on top of it.



The Multiconnect solution **requires an existing high availability setup** first. Please follow the [High Availability documentation](#) to set it up, then continue below.

This document covers:

- [Setting Up Multiconnect](#)
- [Retrieving Shard Information](#)
- [Message Send Rates](#)
- [AWS Deployment Details](#)



## Setting Up Multiconnect

Once you have your cluster set up according to the [High Availability documentation](#), use the following request to turn on multiconnect.



**Remember** you need to have shard number + X Docker containers of the Coreapp running before continuing.

Multiconnect does not guarantee High Availability. Have more Corepps than shards running for High Availability.

## Request

```
POST /v1/account/shards
{
  "cc": "your-country-code",
  "phone_number": "your-phone-number",
  "shards": 1 | 2 | 4 | 8 | 16 | 32,
  "pin": "your-pin"
}
```

## Parameters

Name	Required	Description
cc	Yes	Country code for the phone number registered for this WhatsApp Business API Client as a string (e.g., "1")
phone_number	Yes	Phone number registered for this WhatsApp Business API Client without the country code or plus symbol (+) as a string (e.g., "6315550000")
shards	Yes	<b>Options:</b> 1, 2, 4, 8, 16, 32 Number of shards you want to have as an integer
pin	No	The existing 6-digit PIN for two-step verification as a string (e.g., "123456") This is only required if you have <a href="#">two-step verification</a> enabled on this account.

## Response

```
201 Created      : You successfully changed shard number
403 Forbidden    : You could hit this if server is temporarily unavailable, retry
```

If you see an error when setting the shards, please try again.

## Retrieving Shard Information

### Request

```
GET /v1/account/shards
```

### Response

```
{
  "account": {
    "shards": number-of-shards
  }
}
```

## Message Send Rates

With Multiconnect deployed the following message send rates are expected:

Message Rate (per sec)	Setup	AWS	Message Type
100 - 150	Active shards: 8 DB connection encryption: disabled DB Storage capacity: 32G	DB: db.m4.2xlarge EC2: c4.large	Text Message Template

## AWS Deployment Details

Template URLs:

- **Enterprise:** [https://s3.amazonaws.com/wa-ent-cfn/wa\\_ent.yml?versionId=C3JDtTfqFxGm4QAd\\_tMm33UHbDCGvts3](https://s3.amazonaws.com/wa-ent-cfn/wa_ent.yml?versionId=C3JDtTfqFxGm4QAd_tMm33UHbDCGvts3)
- **DB:** [https://s3.amazonaws.com/wa-ent-cfn/wa\\_ent\\_db.yml?versionId=1XJEwdOPecEsecG0rfQIZAh9sIKh9Hlv](https://s3.amazonaws.com/wa-ent-cfn/wa_ent_db.yml?versionId=1XJEwdOPecEsecG0rfQIZAh9sIKh9Hlv)
- **Lambda:** [https://s3.amazonaws.com/wa-ent-cfn/wa\\_ent\\_lambda.yml?versionId=qo\\_Tx6j6.M5WJjE4b3k22bpQz4YJHFV\\_](https://s3.amazonaws.com/wa-ent-cfn/wa_ent_lambda.yml?versionId=qo_Tx6j6.M5WJjE4b3k22bpQz4YJHFV_)
- **Network:** [https://s3.amazonaws.com/wa-ent-cfn/wa\\_ent\\_net.yml?versionId=5II\\_QAUA7H1Og9HXWdf7Ds1LYkrYTjsQ](https://s3.amazonaws.com/wa-ent-cfn/wa_ent_net.yml?versionId=5II_QAUA7H1Og9HXWdf7Ds1LYkrYTjsQ)

The template allows you to configure the number of active Coreapp container instances to be created. The template creates one additional Coreapp container instance to aid quick switchover in case of Coreapp failure.

The template creates the following number of instances per environment type for Multiconnect, by default, when High Availability is enabled:

- **Production:** EC2 instances: 3, Web container: 3, Coreapp container: 3, Master container: 3
- **Staging:** EC2 instances: 2, Web container: 2, Coreapp container: 2, Master container: 2

The template is configured to auto scale EC2 instances depending on the memory utilization. Memory utilization increases (or decreases) with an increase (or decrease) in number of “active” Coreapp container instances. Hence, when more Coreapp instances are created, EC2 instances auto-scale accordingly. However, maximum number of EC2 instances that can be created is capped as follows:

Active Coreapp Instances	Maximum EC2 Instances
2	3
4	4
8	5
16	8
32	15

## RDS Instance Sizing

API request rate & number of active Coreapp instances determine number of connections to the database. With 8 active Coreapp instances and an API rate of 100 messages/second, it requires about 700 DB connections (SSL is disabled) and 1200 DB connections (when SSL is enabled). However, with 32 active Coreapp instances and an API rate of 250 messages/second, it requires about 1,700 DB connections.

In the current release, we used `db.m4.2xlarge` for 8 active Coreapp instances (DB connection encryption disabled) and `db.m4.4x.large` for 32 active Coreapp instances (DB connection encryption enabled). The following table provides a guidance on RDS instance class selection and the number of maximum connections it can support:

RDS Instance	Maximum DB Connections
<code>db.t2.medium</code>	318
<code>db.t2.large</code>	636
<code>db.t2.xlarge</code>	1272
<code>db.t2.2xlarge</code>	2543
<code>db.r4.large</code>	1212
<code>db.r4.xlarge</code>	2424
<code>db.r4.2xlarge</code>	4848
<code>db.r4.4xlarge</code>	9696
<code>db.r4.10xlarge</code>	19391
<code>db.r4.16xlarge</code>	38783
<code>db.m4.large</code>	636
<code>db.m4.xlarge</code>	1272
<code>db.m4.2xlarge</code>	2543

db.m4.4xlarge	5086
db.m4.10xlarge	12716
db.m4.16xlarge	20345
db.m3.medium	298
db.m3.large	596
db.m3.xlarge	1192
db.m3.2xlarge	2384

## Configuration

- Active Coreapp instances set in a template only govern the number of Coreapp instances created. However, to activate the same, **Set Shards** (documented in the [Setting Up Multiconnect](#) section) must be used. Default value of shards is 1.
- Always ensure that the number of Coreapp instances is always greater than or equal to the shard number set in the API.
- To increase the number of shards:
  - Create or update the stack with desired number of active Coreapp instances.
  - Once successful, use **Set Shards** to activate the same number of active Coreapp instances/shards.
  - *Note:* **Set Shards** causes all Coreapp container instances to stop and be restarted automatically. There will be a downtime of about 45 seconds to 1 minute when **Set Shards** is executed.
- To decrease the number of shards:
  - Use **Set Shards** to reduce the same number of active Coreapp instances/shards.
  - Once all the Coreapp instances restarts successfully, update the stack with the same number of active Coreapp instances.
  - *Note:* Updating the stack might terminate currently active Coreapp instances that are serving the shard. However, other alive Coreapp instances will be assigned shortly. In other words, there could be an additional downtime (~35 seconds) during this procedure.



**WhatsApp Business API**

Overview

Getting Started

Guides

API Reference

Webhooks

**Availability and Scaling**

High Availability

**Multiconnect**

Message Templates

Guidelines

Troubleshooting

Changelog

FAQ