

Couchbase on AWS

Best Practices for Running Couchbase on AWS

Couchbase Cloud-native Data Platform

v1.0 - November 2018

This document outlines the architectural versatility of the Couchbase Data Platform, specifically the tight integration with Amazon Web Services. It also highlights the Couchbase Autonomous Operator and its industry leading native integration with Amazon EKS.

Why Choose Couchbase on AWS?

Couchbase has a long and productive history with AWS. Couchbase was one of the original launch partners on the [AWS Marketplace in 2012](#), and has continued to work closely with Amazon to provide a first-class integration experience with AWS.

Some of the highlights of Couchbase on AWS are:

- Memory-first architecture
- High availability, fault tolerance, and load balancing
- Multi-master and multi-region support for optimal performance
- Flexible data model with full SQL on JSON
- Deployments that can take less than five minutes
- Automated deployment on Kubernetes using the Couchbase Autonomous Operator, with native support for Amazon EKS¹

Couchbase Deployment Options on AWS

The AWS platform is uniquely suited to accommodate the versatility of Couchbase [products](#). Currently, Couchbase offers two products on AWS: Couchbase Enterprise Server and Couchbase Sync Gateway. Both products have two production pricing options: Bring Your Own License (BYOL), where you procure a license from Couchbase [directly](#); and Hourly, where you're billed through your AWS account. Pre-production environments are always free (excluding your EC2 costs), so you can feel free to get started right away (just use the BYOL option).

The following sections cover the various options for deploying Couchbase on AWS.

Infrastructure as Code (IaC) Deployments

CloudFormation Templates Through the Couchbase GitHub Repository (EC2)

For more complex deployments, you can use the official CloudFormation Templates (CFTs) on [GitHub](#) to deploy Couchbase from the AWS command line interface. This method provides more flexibility and allows for the use of advanced Couchbase features, such as [Multi-Dimensional Scaling](#) (MDS) and [Cross Data Center Replication](#) (XDCR).

¹ The Couchbase Autonomous Operator is currently in developer preview for Amazon EKS

Note: When using the CFTs, the billing and pricing options are the same as using the AWS Marketplace.

Terraform

Terraform is a popular third-party IaC orchestration solution, and Couchbase has a fully-featured solution for AWS available on [GitHub](#). The Terraform solution supports MDS, XDCR, simple deployments, and has a full test suite. It's a good choice if Terraform is a technology you're more familiar with.

Kubernetes and Containers

Couchbase Autonomous Operator and Amazon EKS

Couchbase is the first NoSQL company to [release an operator](#) for Kubernetes. The goal of the Couchbase Autonomous Operator is to fully manage one or more Couchbase deployments so that you don't need to worry about the operational complexities of running Couchbase. Not only does the Operator automatically administer your Couchbase cluster, it also manages the cluster according to best practices.

The Operator achieves this level of automation by extending the Kubernetes API to create, configure, and manage instances of complex stateful applications on behalf of a Kubernetes user. It builds upon the basic Kubernetes resource and controller concepts, but also includes domain or application-specific knowledge to automate common tasks better managed by computers. Whether you're using a single-provider, hybrid cloud, or cross-cloud strategy, the Couchbase Autonomous Operator is deployment agnostic, thereby enabling seamless multi-cloud deployment and migration.

If you already have an open source Kubernetes environment, you can start using the Operator in [production](#). If you're using Amazon EKS (the Managed Kubernetes Service offering on AWS), an official GitHub repository is available to help you [get started](#) using the Operator (currently in developer preview for Amazon EKS).

AWS Marketplace

The [AWS Marketplace](#) offers a great starting point for those wanting to start or modify their foray into Couchbase on AWS. The AWS Marketplace allows for a clean and simple click-to-deploy web interface to get instances deployed quickly. The Couchbase Enterprise Edition product on AWS Marketplace is based on a CloudFormation Template (CFT), which is the native AWS IaC solution.

The Couchbase Enterprise Edition product can be deployed for any combination of Couchbase Server or Couchbase Sync Gateway nodes from one product listing

Couchbase Couchbase Enterprise Edition [Continue to Subscribe](#)

Overview **Pricing** Usage Support

Pricing Information

Use this tool to estimate the software and infrastructure costs based on your configuration choices. Your usage and costs might be different from this estimate. They will be reflected on your monthly AWS billing reports.

Estimating your costs

Choose your region and fulfillment option to see the pricing details. Then, modify the estimated price by choosing different instance types.

Region:

Fulfillment Option:

Software Pricing Details

Couchbase Server Enterprise Edition (Hourly Pricing)	\$0.4 /hr >
<small>running on m4.large</small>	
Couchbase Sync Gateway Enterprise Edition (Hourly Pricing)	\$0.796 /hr >
<small>running on m4.xlarge</small>	
Infrastructure Pricing Details	
Estimated Infrastructure Cost	\$0.2 EC2/hr >

The table shows current software and infrastructure pricing for services hosted in US East (N. Virginia). Additional taxes or fees may apply.

EC2 Instance type	Software/Hr	EC2/Hr	Total/Hr
<input type="radio"/> m4.large	\$0.2	\$0.1	\$0.3
<input checked="" type="radio"/> m4.xlarge	\$0.4	\$0.2	\$0.6
<input type="radio"/> m4.2xlarge	\$0.8	\$0.4	\$1.2
<input type="radio"/> m4.4xlarge	\$1.6	\$0.8	\$2.4
<input type="radio"/> m4.10xlarge	\$4	\$2	\$6
<input type="radio"/> m4.16xlarge	\$6.4	\$3.2	\$9.6

Figure 1: Couchbase Multi-AMI offering on the AWS Marketplace.

The Couchbase Server Enterprise Edition product is simple and straightforward. The base CFT deploys all services (data, query, index, and fts) on the instances.

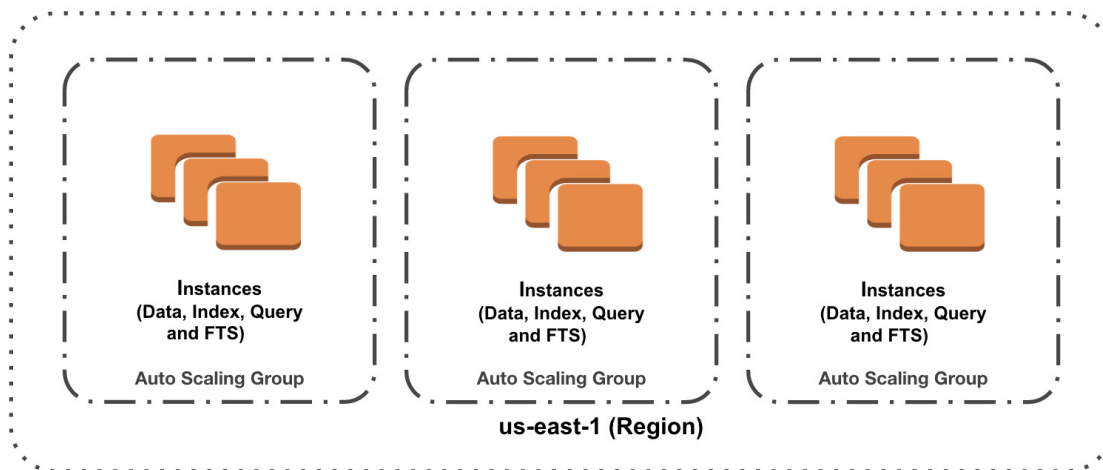


Figure 2: Couchbase Server Marketplace Deployment

If you require a bit more flexibility, consider using one of the previously mentioned IaC options (or the IaaS approach covered in the next section).

AWS EC2 - Infrastructure as a Service (IaaS)

This is the classic way to deploy Couchbase to AWS. In this scenario, you install Couchbase on your own EC2 instances running a supported OS, and then set up the entire environment from scratch. Refer to the [official Couchbase documentation](#) for detailed instructions.

Best Practices for Deploying Couchbase on AWS

Choosing EC2 Instance Types

AWS provides several EC2 [instance types](#). The designation of the instance type gives insight into the suitability for your use-case. Using r5.4xlarge as an example (16 cores and 128GB RAM); “r” indicates the instance family, “5” indicates the instance generation, and “4xlarge” indicates the instance size. The instance type also specifies what type of storage (and its bandwidth that can be used).

The recommended choice for typical production workloads is the R family (memory optimized) with r5.4xlarge being a good balance of price and performance. A reasonable all-around choice is the M family which is classified as general purpose. For very small workloads, consider the T family. For compute-intensive workloads, consider the C family.

Although the choice always depends on the specific use-case, a good place to start is at least 4 cores that support block storage (see the next section).

Storage Options

[Elastic Block Storage](#) (EBS) is the recommended and easiest to manage choice. EBS volumes persist after the instance it is attached to is terminated. EBS has many volume options to meet the needs of most use-cases. Couchbase recommends considering the gp2 volume type as a balance between performance and price. For higher performance needs, the io1 volume type should be considered. If none of these options meet your storage performance requirements, then instance storage should be used. When deciding to use instance storage, there is an increase in the complexity of administration (and cost) that should be considered.

Networking

Use [EC2-VPC](#) instead of [EC2-Classic](#). Virtual Private Clouds (VPCs) give you network isolation, which leads to simpler configuration. By using the default VPC (provided by default per region), you simplify network configuration even further.

Deploy nodes across all availability zones that are available within a region. Configure your nodes with public DNS entries and let AWS’ support for split horizon DNS handle communication within the VPC.

For Couchbase Server, configure your security groups to allow XDCR ports to communicate to the public and for remote client access (if needed). All other Couchbase [ports](#) should be accessible within the VPC only.

Couchbase Deployment Examples

Couchbase IaC-based deployments use [Auto Scaling Groups](#) for groups of instances (e.g. data nodes). Make sure to disable the auto-scaling feature of the Auto Scaling Groups after your

deployment is complete. Then create Couchbase server groups that correspond to these Auto Scaling Groups to provide a durable deployment.

The following sections describe some cluster topologies that you may want to consider.

Multi-Dimensional Scaling (MDS)

Leverage multiple AWS instance types to improve resource utilization and reduce operational costs by enabling [Multi-Dimensional Scaling](#) to run different services (e.g., data, index, and query) on different instance types – optimizing resource allocation.

The following is a MDS deployment that spans one region:

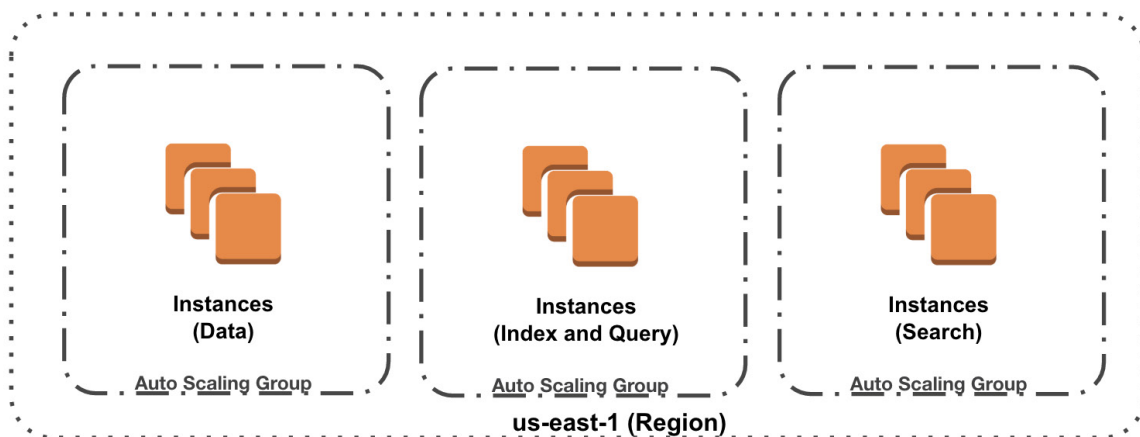


Figure 3: Instances are spread across all availability zones

Rack-Zone Awareness

Leverage multiple AWS availability zones for high availability and fast failover by configuring [Server Group Awareness](#) to replicate all data to multiple AZs within a region – using proxies or routers to redirect traffic without delay.

Cross Data Center Replication (XDCR)

Leverage multiple AWS regions to improve data locality, achieve higher availability, and enable faster disaster recovery with multi-master, multi-site deployment via built-in bidirectional [Cross Data Center Replication](#).

The following diagram shows a Couchbase deployment with Zone Awareness and Cross Data Center Replication:

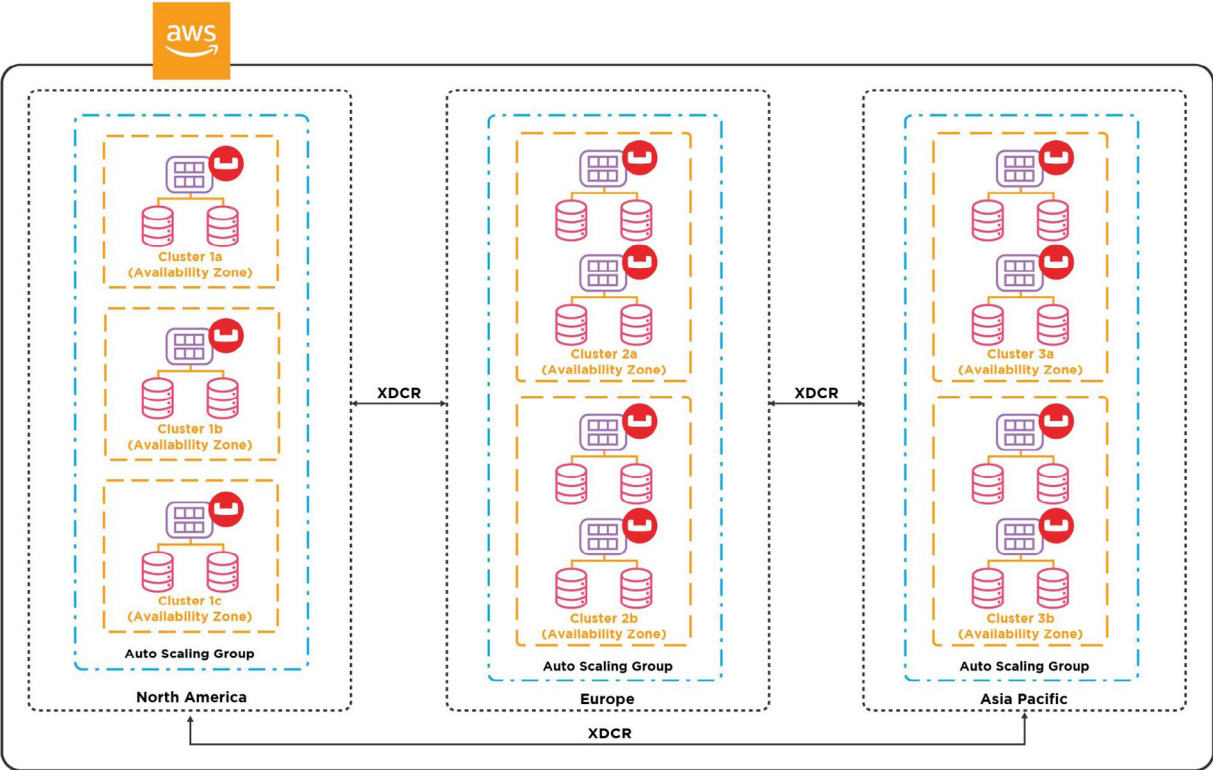


Figure 4: A possible XDCR deployment

Now that you've reviewed the options for deploying Couchbase on AWS, take some time and get your hands dirty with some [hands-on free learning](#).

About Couchbase

Couchbase's mission is to be the data platform that revolutionizes digital innovation. To make this possible, Couchbase created the world's first Engagement Database to help deliver ever-richer and ever-more-personalized customer and employee experiences. Built with the most powerful NoSQL technology, the Couchbase Data Platform was architected on top of an open source foundation for the massively interactive enterprise. Our geo-distributed Engagement Database provides unmatched developer agility and manageability, as well as unparalleled performance at any scale, from any cloud to the edge.

