Couchbase AWS Deployment Guide

Introduction

Couchbase is an industry leading full-featured NoSQL Database Platform. The Couchbase AWS marketplace offer Couchbase Server and Sync Gateway is an IaaS solution. The documentation for Couchbase can be used to understand the technology and features of the product. This document augments the aforementioned documentation website as it relates to Couchbase features and recommendations. The Couchbase offer includes two products and you must choose at least one by selecting more than zero nodes for the desired product. Each product has two licensing options, BYOL and Hourly. The Couchbase offer uses the Enterprise Editions, however; any non-production deployment is free (by choosing the BYOL option). Each Couchbase Server node has all of the Couchbase Server services running (i.e. data, index, full text search, eventing, query and analytics). The offer is intended for deployments in simple use-cases and/or nonproduction environments. To customize most aspects of a deployment use the GitHub Cloud Formation Templates (see the "Typical AWS Deployment" Section). In general, for any complex deployment; utilizing Couchbase Training and Professional Services is recommended.

Intended use

Use Couchbase in any scenario where a NoSQL Database Platform, Caching data store is needed. The marketplace offer (using the UI) is for simple deployments and non-production environments. More complex configurations should use the command line and GitHub Cloud Formation Templates.

Use cases

Couchbase use cases may be found at the Couchbase website.

Typical AWS deployment

A typical Couchbase deployment creates the following resources:
- An AWS::IAM::Role named CouchbaseRole.
- An AWS::IAM::InstanceProfile named CouchbaseInstanceProfile (generated from the above IAM Role).
- An AWS::EC2::SecurityGroup named CouchbaseSecurityGroup that enables SSH and Couchbase Ports.
- An AWS::AutoScaling::AutoScalingGroup named ServerAutoScalingGroup to manage Couchbase Server nodes.
- An AWS::AutoScaling::LaunchConfiguration that corresponds with your provided parameters for the Couchbase Server part of the deployment.
- An AWS::AutoScaling::AutoScalingGroup named SyncGatewayAutoScalingGroup to manage Couchbase Sync Gateway Nodes.
- An AWS::AutoScaling::LaunchConfiguration that corresponds with your provided parameters for the Couchbase Sync Gateway part of the deployment.

- An EBS volume per node for data and an EBS volume per node for the OS that is 8 Gib (with DeleteOnTermination set).

Additionally, it is also important to note:
- VMs use Amazon Linux as the Operating System.
- Memory is allocated forty percent to the data service, eight percent each to the other applicable services for a total of 72% of memory (to adjust this use the CFTs referenced earlier in this section).
- Each node has a public IP address so that the Couchbase Web Console can be reached from any node.
- A deployment requires an SSH key that will be used by the deployment to access the virtual machines.
- SSL is not enabled by default for the Couchbase Web Console; refer to the Couchbase documentation for steps to enable it.
- Backup is not enabled by default, however; Couchbase is somewhat FT by design if the cluster is sized correctly.
- The Couchbase documentation has guidance on deploying from the AWS marketplace as well.

Time to deploy

A deployment usually takes less than six minutes but no more than ten.

Prerequisites and requirements

You will need an AWS account and use an IAM user with at least the following permissions:
- aws-marketplace:Subscribe
- aws-marketplace:ViewSubscriptions
- ec2:AuthorizeSecurityGroupEgress
- ec2:AuthorizeSecurityGroupIngress
- ec2:CreateSecurityGroup
- ec2:DescribeImages
- ec2:DescribeInstances
- ec2:DescribeKeyPairs
- ec2:DeleteSecurityGroup
- ec2:DescribeSecurityGroups
- ec2:DescribeSubnets
- ec2:DescribeVpcs
- ec2:DescribeAccountAttributes
- ec2:RunInstances
- ec2:StartInstances
- ec2:StopInstances
- ec2:TerminateInstances
- iam:AddRoleToInstanceProfile
- iam:CreateInstanceProfile
- iam:DeleteInstanceProfile

- iam:DeleteRole
- iam:DeleteRolePolicy
- iam:PassRole
- iam:PutRolePolicy
- iam:RemoveRoleFromInstanceProfile

The Couchbase offer is straight forward and there are only a few technical prerequisites that are needed:
- A basic understanding of the Couchbase product you want to use.
- How to operationally manage Amazon Linux virtual machines (in order perform infrastructure and software maintenance).
- How to create an SSH key in a region (see appendix).
- Optionally if modifying the CFTs, CFT knowledge, an understanding of bash and an understanding of Python may be helpful.
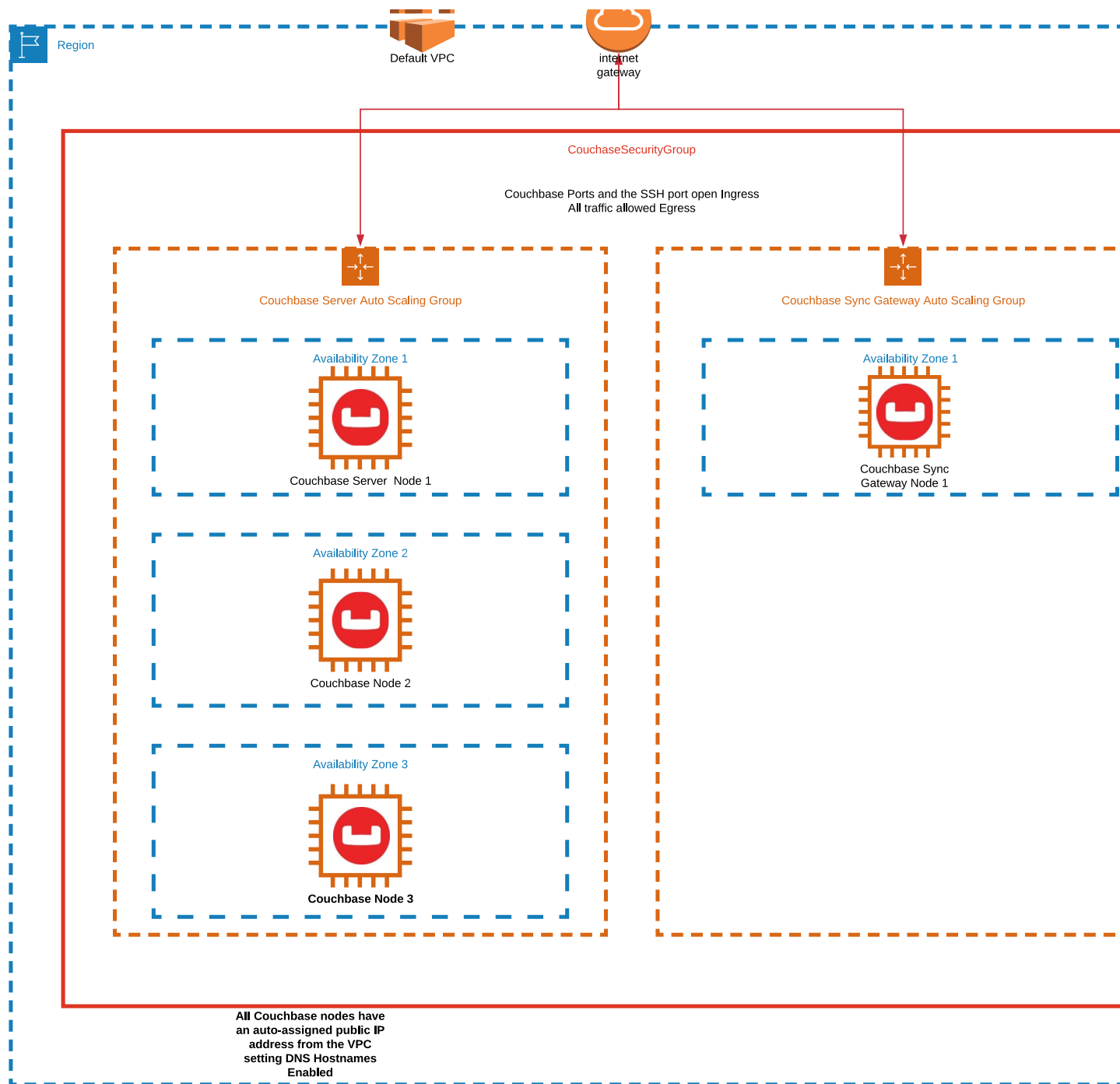
Architectural Diagrams

Region

Default VPC

internet gateway

CouchaseSecurityGroup

Couchbase Ports and the SSH port open Ingress
All traffic allowed Egress

Couchbase Server Auto Scaling Group

Availability Zone 1

Couchbase Server Node 1

Availability Zone 2

Couchbase Node 2

Availability Zone 3

Couchbase Node 3

Couchbase Sync Gateway Auto Scaling Group

Availability Zone 1

Couchbase Sync
Gateway Node 1

**All Couchbase nodes have
an auto-assigned public IP
address from the VPC
setting DNS Hostnames
Enabled**

Figure 1 A typical Couchbase Marketplace Deployment

Security

The management of the deployed infrastructure is the responsibility of the user. It is expected
that a purchaser of the software is able to maintain the environment. The pillar of the security
guidance is to restrict all credentials (i.e. the SSH keys and cluster credentials) to only those who

need it and to encrypt everything. By default, EBS encryption is enabled. However, using more robust third-party options for data encryption at rest is recommended.   SSL is disabled by default; between the nodes and for the Couchbase Web Console. However, this should be enabled before adding any critical data. Handling certificates is well documented and provides complete guidance.

The CouchbaseRole IAM Role (see appendix) allows the deployment to tag a key instance as the Rally Server (see the Deployment Guidance) section for more about the Rally Server. There is nothing special about this server outside of the name and it has no additional permissions compared to other nodes.

The deployment requires one SSH key pair that allows access to all the Couchbase nodes on cluster creation. The CouchbaseSecurityGroup defines the open ports for the deployment. All ports are Couchbase ports outside of the SSH port that is open to the public internet for simplicity and pre-production environments. For production ready environments whitelist certain IPs for access to SSH. An even more secure approach, is to deploy a secure Bastion Host into the VPC for SSH access and block port 22 from the public internet. With this option there the private SSH keys should be stored on the bastion alone, since it is the only location that SSH to the nodes would occur.

Minimally AWS CloudTrail and optionally, AWS CloudWatch should be used to monitor your IAM access and attempts to your cluster. AWS CloudTrail event history is enabled by default, however; for more data and alerting consider using CloudTrail insights and CloudWatch alerts together.

Planning Guidance

Created AWS Resources and Licensing

A Couchbase deployment consists only of EC2 instances and EBS storage. The minimum recommended Couchbase Server nodes in a cluster is three. As mentioned in the "Introduction" Section Couchbase has two licensing options BYOL and Hourly. BYOL gives you more options for the support model and licensing period. With an Hourly license all of your solution charges are billed by AWS and you receive Silver Support. The Couchbase Server and Sync Gateway listing shows up-to-date Hourly pricing information under the "Pricing Information" Section. For a full picture of your AWS infrastructure costs; use the AWS Simple Monthly Calculator to accurately estimate your costs.

Sizing

For EBS data volumes gp2 is a good balanced choice. For general purpose workloads us the latest M series instances. If your workload is more compute intensive use the latest C instances. For high memory use cases utilize the latest R or C series instances. More information on sizing in the Couchbase documentation.

Deployment Guidance

The Couchbase documentation also provides step by step instructions on deploying a Couchbase cluster from the Marketplace.

Deployment Assets

Please see the "Typical AWS deployment" Section.

Operational Guidance

To manage a Couchbase deployment explore the Couchbase documentation.

Health Check

The out of the box health check solution is the Couchbase Web Console. The Couchbase Web Console gives important status information of your cluster including, node health, memory residency and other important metrics. Additionally, setting alerts is recommended. To monitor the infrastructure AWS or third-party monitoring tools that fit your needs are recommended.

Backup and Recovery

Depending on your use-case Backup and Recovery may be viable for you.

Emergency Maintenance

The Couchbase offer deploys over multiple Availability Zones for Resilience and High Availability at the infrastructure level. Couchbase as a product seamlessly handles Cluster Availability and Recovery. However, it is important to correctly size a cluster according to your Data Set and have at least have a fair understanding of the product features. Depending on your requirements consider scheduling backups as well. For any production ready solution, it makes sense to contact Couchbase Training and Professional Services early in your Software Development Lifecycle to make sure your solution follows all best practices.

Support

Review all the support options that fits your need. However, the Hourly billing option comes with Silver Support. BYOL comes with the support that matches your purchased license.

Appendix

Create an SSH key via the AWS Command Line Interface by running the following commands on a Linux/Unix based OS.

```
REGION=`aws configure get region`
KEY_NAME="couchbase-${REGION}"
KEY_FILENAME=~/.ssh/${KEY_NAME}.pem
aws ec2 create-key-pair \
  --region ${REGION} \
  --key-name ${KEY_NAME} \
```

```
  --query 'KeyMaterial' \
  --output text > ${KEY_FILENAME}
chmod 600 ${KEY_FILENAME}
echo "Key saved to ${KEY_FILENAME}"
```

CouchbaseRole is an AWS::IAM::Role that has an inline policy called CouchbasePolicy that is defined as:

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Action": [
        "ec2:CreateTags",
        "ec2:DescribeTags",
        "ec2:DescribeInstances",
        "autoscaling:DescribeAutoScalingGroups"
      ],
      "Resource": "*",
      "Effect": "Allow"
    }
  ]
}
```

CouchbaseSecurityGroup is an AWS::EC2::SecurityGroup with the rules as shown in the screenshot below:

Edit inbound rules

Type	Protocol	Port Range	Source	
Custom TCP I ⇕	TCP	18091 - 18096	Custom ⇕	0.0.0.0/0
Custom TCP I ⇕	TCP	9100 - 9105	Custom ⇕	0.0.0.0/0
Custom TCP I ⇕	TCP	8091 - 8096	Custom ⇕	0.0.0.0/0
SSH ⇕	TCP	22	Custom ⇕	10.0.0.0/16
Custom TCP I ⇕	TCP	21100 - 21299	Custom ⇕	0.0.0.0/0
Custom TCP I ⇕	TCP	4369	Custom ⇕	0.0.0.0/0
Custom TCP I ⇕	TCP	9998 - 9999	Custom ⇕	0.0.0.0/0
Custom TCP I ⇕	TCP	9110 - 9122	Custom ⇕	0.0.0.0/0
Custom TCP I ⇕	TCP	4984 - 4985	Custom ⇕	0.0.0.0/0
Custom TCP I ⇕	TCP	11207 - 11215	Custom ⇕	0.0.0.0/0

Figure 2 CouchbaseSecurityGroup Inbound Rules

Edit outbound rules

Type ⓘ	Protocol ⓘ	Port Range ⓘ	Destination ⓘ
All traffic ⇅	All	0 - 65535	Custom ⇅ 0.0.0.0/0

Note: The virtual machines have the default SSH port (22) open by default, however; access to the virtual machines are protected by the SSH key required to deploy.

Glossary

IaaS – Infrastructure as a Service which uses VMs as a replacement for bare-metal servers with all of the flexibility of an on-premise deployment.
KV – Key Value where data is stored with by a key and its corresponding value.
ASG – An Auto Scaling Group contains a collection of Amazon EC2 instances that are treated as a logical grouping for the purposes of automatic scaling and management.
IAM - Identity and Access Management.
CFT – Cloud Formation Template which is the Infrastructure as Code native solution of AWS.
AWS – Amazon Web Services.
EBS - Amazon Elastic Block Store (EBS) is an easy to use, high performance block storage service designed for use with Amazon Elastic Compute Cloud (EC2) for both throughput and transaction intensive workloads at any scale.
Couchbase Server – The Couchbase Enterprise Database product.
Couchbase Sync Gateway – A secure web gateway for data access and synchronization over the web.
Couchbase Web Console - Couchbase Web Console is a browser-based, interactive graphical facility that supports the management of Couchbase Server
BYOL – Bring your own license, i.e. buy a license directly from Couchbase.
Hourly – AWS billing your account on an hourly basis for Couchbase software.
EC2 - Amazon Elastic Compute Cloud.
VM - (EC2) is the Amazon Web Service you use to create and run virtual machines in the cloud.
IP address – Internet protocol address.