Auto Scaling of Barracuda Web Application Firewall using CloudFormation Template on Amazon Web Services

https://campus.barracuda.com/doc/46897279/

The Barracuda Web Application Firewall can be deployed on AWS using the CloudFormation Template. The Barracuda Web Application Firewall integrates with various AWS services to provide Auto Scaling capabilities that enable the Barracuda Web Application Firewall deployment to scale up/down based on auto scale metrics such as CPU utilization and bandwidth.

Deployment using the CloudFormation template also enables you to bootstrap the configuration of the Barracuda Web Application Firewall for AWS. The initial deployment will allow you to specify the service configuration during launch. Later, when new instances appear, they will automatically synchronize the configuration from the previously deployed Barracuda Web Application Firewall instances and serve traffic with complete configuration.

You can define the scaling polices for your instances and set the minimum and maximum number of instances to be used on demand. Auto Scaling can be used for applications that have stable demand as well as for applications that experience hourly, daily, or weekly variability in usage. For more information on AWS Auto Scaling, refer to the AWS article.

A CloudFormation Template (CFT) is a declaration of the Amazon Web Services resources that creates a stack. The Barracuda CloudFormation Template will deploy the Barracuda Web Application Firewall for AWS with the basic service configuration and set up the necessary AWS services (Auto Scale Groups and Launch Configurations, CloudWatch Alarms, SNS Email Notifications, IAM Roles and S3 Buckets) for successful Auto Scaling and bootstrapping.

The Barracuda Web Application Firewall for AWS provides the CFT for:

- Bring Your Own License (BYOL)
  - Basic Bootstrapping
  - Backup Bootstrapping
- Pay-As-You-Go/Hourly
- Metered/Usage Based

For the Barracuda Web Application Firewall for AWS CloudFormation Templates, see CloudFormation Templates on GitHub.

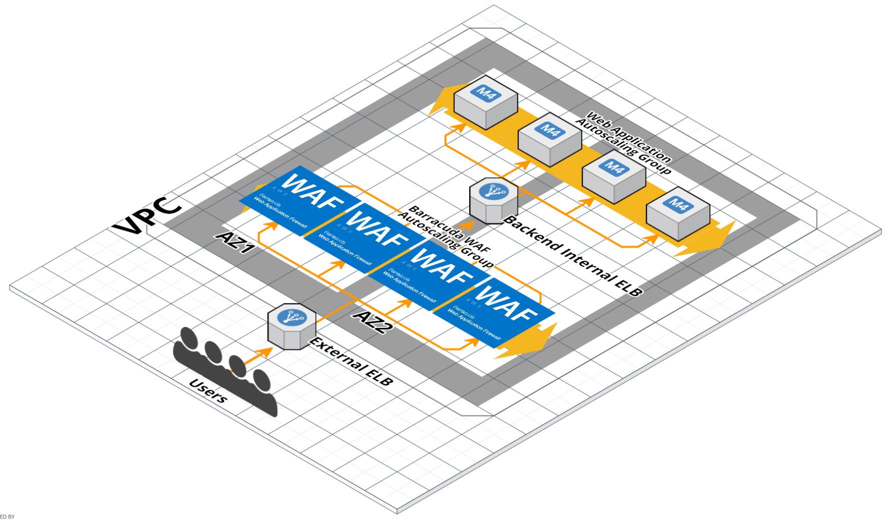
To learn more about the supported bootstrapping methods, see Bootstrapping Methods.

Metered/Usage Based

To successfully launch and use the metered AMI, the account being used to launch the instance must have the ability to use the “**AWSMarketplaceMeteringFullAccess**” IAM role with

"**Allow All**" permissions. If these permissions are not available, the instance will not be launched successfully.

The CFT deploys the Barracuda Web Application Firewall for AWS into a pre-existing VPC deployment to secure the servers. A typical deployment would look like this:



POWERED BY CLOUDCRAFT.CO

AWS Services Required for the Auto Scaling Setup

The following are the AWS services required for the auto scaling setup:

- Virtual Private Cloud (VPC)
- Elastic Compute Cloud (EC2)
- CloudFormation
- Simple Storage Service (S3)
- SNS
- CloudWatch
- Identity and Access Management (IAM)

Best Practices

Following are the best practices that should be considered to have an optimal deployment:

Instance Type

Before launching an auto scaling group in your region, ensure that you always check if the selected instance type is supported in all the available deployments. Refer to the list to know the instance types supported in your region.

Number of Instances

It is ideal to run the CloudFormation Template (CFT) with the **Minimum Instances** set to one (1), as clustering may not work as designed if two instances boot up at the same time. This issue will be fixed in a later release. You can add more instances by modifying the Auto scaling configuration on the AWS EC2 console. Auto Scaling groups are available in the Amazon EC2 Management Console.

Alarms

The CloudFormation Template includes predefined alarms that are triggered based on the set metrics. These alarms may scale up new instances, or scale down older instances based on the alarm configuration. It is recommended that you deploy a test setup, perform the auto scaling test, and modify the alarms as per your requirement. For instance, there may be cases where you want the instances to scale up later, or scale down earlier based on your testing. In such cases, you can either modify the CFT before uploading it, or modify the alarm in the Amazon CloudWatch Management Console after deploying the CFT.

The Barracuda CloudFormation Template includes alarms that are based on the following principles for optimal usage:

- **Scale Up Early** – The “Scale Up” alarms are set to trigger early when the traffic flow/CPU increases. Scaling up the instance(s) early when the instance encounters high traffic flow or CPU usage ensures the service works without interruption and the requests are load balanced between the Barracuda Web Application Firewall instances for optimal performance.
- **Scale Down Slowly** – To prevent thrashing of instances being brought up and down continually, the alarms are designed to scale down slowly. The time taken to scale down an instance should ideally be at least an hour or more, to ensure that there are sufficient instances available to handle the available load smoothly, and auto scaling does not happen until it is absolutely required.

Next Step

Continue with the How Barracuda CloudFormation Template Works in PAYG/Hourly Instance article.

Figures

© Barracuda Networks Inc., 2024 The information contained within this document is confidential and proprietary to Barracuda Networks Inc. No portion of this document may be copied, distributed, publicized or used for other than internal documentary purposes without the written consent of an official representative of Barracuda Networks Inc. All specifications are subject to change without notice. Barracuda Networks Inc. assumes no responsibility for any inaccuracies in this document. Barracuda Networks Inc. reserves the right to change, modify, transfer, or otherwise revise this publication without notice.