Cloud Backups and Archiving Deployment Runbook

Trent J Broderick

University of Maryland Global Campus

November 24, 2020

**Executive Summary**

This document is intended for the management team of BallotOnline. The goal of developing this book is to provide deeper knowledge around a reference architecture for a potential deployment of a backup and archiving environment. The executive team must pay close attention to all of the service deployed to create this secure architecture as each piece while needed, may incur additional costs. There are several methods of ensuring costs are contained and will also be explored in this runbook.

The ask from the executive team is to leverage this runbook while working with a cloud vendor and or cloud migration technician as a reference architecture. It is extremely important that BallotOnline ensures that VPCs, VPNs, S3 Buckets, Synchronization services as well as CloudWatch and monitoring events are implanted in this architecture. Through leveraging these systems with CloudWatch, BallotOnline will remain fiscally responsible while leveraging cloud resources.

**Overview**

This runbook documents a cloud architect and engineers’ journey that creates a basic deployment for a backup and archiving environment in AWS. This deployment contains elements of Virtual Private Cloud, Networking, S3 Bucket Configuration, Cloud Synchronization Service, and setting up CloudWatch alarms (UMGC, 2020). This runbook contains screenshots as well as commentary on personal experiences while setting up this environment.

The requirements to rerun this book are an AWS account, a NetApp Account, and access to the SMB server where the files for synchronization are needed. It is encouraged that if this is to be repeated, the screenshots and personal commentary are leveraged to ensure a smooth deployment.

**Cloud Service Offering for Backups and Archiving**

Amazon offers two main services for backups and archiving, Amazon S3 and Amazon Glacier (Amazon, 2020). Both services are extremely valuable to BallotOnline from both short- and long-term storage. S3 offers the ability to store files in S3 buckets that are cross regional, highly available object stores for frequently accessed storage while Glacier stores files that may not need to be accessed for years or even less frequently. S3, and Glacier is compatible with on-prem services as well as their other services such as Elastic Block Storage, Deep Archive, Amazon Backup and Snapshots (Amazon 2020).

**AWS Virtual Private Cloud (VPC)**

Amazon Virtual Private Cloud is a service that allows Amazon users to launch Amazon Resources in a Virtual Network. Customers can configure IP Tables, Access Control Lists or use a virtual routing appliance to allow specific types of traffic in and out of the environment (Amazon, 2020). A majority of secure architectures with any vendor include some type of Virtual Private Cloud or Layer Two set up to give customers complete access of the networking to help protect against networking related threats.

Graphical user interface, text, application

Description automatically generated**Configuration and Deployment of VPC and Networking**

Figure : AWS VPC Dashboard (Amazon 2020

Configuring and managing a Virtual Private Cloud is generally simple to do. Within the AWS Console, navigate to the VPC service, click on your VPC and create VPC. Granted to size of the network, I deployed the VPC with a /26 CIDR Block to provide 62 usable IP Addresses (UMGC, 2020).

Graphical user interface, website

Description automatically generated

Figure : AWS VPC Post Creation Dashboard (Amazon, 2020)

Once the VPC was created, the next step was to configure the networking for the VPC. Networking is extremely important to configure here because if it is left unchecked, not only will the instances and services within the VPC not communicate with each other, but also those services may not have access to the internet or other Amazon microservices needed to keep the environment updated and running.

Graphical user interface, table

Description automatically generated Next, I split the VPC CIDR Block into several public and private subnets. In this model, secure infrastructure can remain in the private subnet where it will not interact with the public internet where servers and services in the public subnets will be able too once the routing tables are configured (UMGC, 2020).

Figure : AWS Subnet Dashboard (Amazon, 2020)

Following the creation of the subnets, I created routing tables to allow the provate instance to communicate within themselves as well as the public instances. Graphical user interface, text, application

Description automatically generated

Graphical user interface, application

Description automatically generatedFigure : AWS Route Table Dashboard (Amazon, 2020)

Figure 5: AWS Internet Gateway Dashboard (Amazon, 2020)

Now the public instances need the ability to talk to the public internet. In a VPC, there is no immediate public internet access. Amazons answer is to set up what it calls an Internet Gateway and attatch it to the subnets where instances that need to be exposed to the public internet will live. In the VPC dashboard, navigate to Internet Gateways, Create IGW and attach it too the correct subnets (Amazon, 2020).

Lastly, to access the now available subnetes, routing portocals and a NAT gateway need to be created. NAT Gateways allow instances to end infromation out of the VPC and to the public internet. AWS manages the NAT Gateway so there is very little configuration to be done. Once the NAT Gateway is set up, Public IPs need to be auto assigned to instances to allow them to be addressable when they are deployed (UMGC, 2020). I had an error when configuring the NAT gateway due to my student account permissions.

Graphical user interface, text, application, table

Description automatically generatedGraphical user interface, text, application

Description automatically generated

Figure 6: AWS NAT Gateway Creation (Amazon, 2020) Figure 7: AWS Subnet Dashboard (Amazon 2020)

**AWS S3 Bucket**

The S3 bucket is an integral part of this set up and configuration as it is the storage system for BallotOnline storage. S3 object storage is a highly available, secure solution that can be placed within VPCs and well as across regions (Amazon, 2020). There are several tools such as Elastic Beanstalk and CloudWatch to help ensure the buckets are sized properly and BallotOnline always has the proper amount of storage for its architecture. There are also additional services that ensures that objects stored in S3 that may not be used properly such as infrequent or no retrieval use are placed in Glacier or S3 Infrequent use buckets to help make storage more cost effective (Amazon 2020).

**Configuration and Deployment in AWS**

Configuring and deploying an S3 bucket is simple through the AWS Console. Through the AWS Console, navigate to the S3 Storage Service and click create bucket. Please note, all buckets must have unique names and can be placed in specific regions (UMGC, 2020). Once the properties and permissions are configured, the bucket can be deployed.Graphical user interface, text, application, email

Description automatically generated

Figure 8: AWS S3 Bucket Dashboard (amazon 2020)

**Cloud Synchronization Services – NetApp Cloud Sync**

During a cloud migration, services such as NetApp Cloud Sync allow for two environments to be connected in order to share and sync data together. NetApp has a very robust platform that allows for internal, collocated and cloud environments to connect into each other. When leveraging NetApp with Amazon, the pre-made template made synchronization painless and easy (UMGC, 2020)

**Configuration and Deployment of Cloud Sync Service**

After creating a NetApp Cloud Central account, you must create a Sync Relationship. This relationship is what builds the connection between the internal server when documents needed are stored and the Amazon Cloud Environment. The source in this case is a SMB Server and the target is AWS S3. In order to select the target and source, click and drag the source to the source bubble and the target to the target bubble. When it is complete, the NetApp Sync relationship will be established and appear as it does in the screenshot below. Graphical user interface, text

Description automatically generated

Figure 9: NetApp CloudSync Dashboard (NetApp, 2020)

From this point, you must install the Data Broker stack in your AWS account to migrate information form one server to S3. NetApp has an easy click through template that will automate the building of the connection between the servers (UMGC, 2020). The issue I ran into here was the sizing of instances was too big for what my account was permitted to deploy. After resizing the instances, setting up the Data Broker stack was very quick.Graphical user interface, website

Description automatically generated

Figure 10: AWS CloudFormation Stacks Dashboard (NetApp, 2020)

Graphical user interface, text

Description automatically generatedOnce the stack as completely deployed, NetApp will activate the Data Broker relationship between the server and S3 Bucket. Then, once the files that have been needed to be copied over have been, the user is greeted with a “Synced Successfully” message and the data migration will be complete.

Figure 11: NetApp DataBroker Menu (NetApp, 2020)

Graphical user interface, application

Description automatically generated

Figure 12: NetApp CloudSync Dashboard (NetApp, 2020)

**Data Backups and Archiving to Cloud Using Cloud Sync Service**

For data that can be stored in highly available cloud environments, the Data Backup and Archiving utility provided by NetApp will allow BallotOnline to keep a continuous funnel of complete backups across both the cloud and internal systems of BallotOnline together. Some additional bonuses to this type of back-up system include the ability to turn up additional Cloud Infrastructure if internal systems experience down time, ensure the user continues to have a flawless experience on the platform. In the event BallotOnline needs to scale internal infrastructure, the cloud can serve as a staging area for the infrastructures data until the racks are cabled and burned in.

**AWS CloudWatch Monitoring**

AWS CloudWatch is a service that monitors the cloud environment for a number of different metrics and allows the user to configure alarms based on the needs for CloudWatch. CloudWatch covers usage, billing, potential damage or non-working servers, network traffic and many other useful metrics (Amazon, 2020). The main focus for BallotOnline is server usage and load as well as billing alerts to help ensure costs are staying under control.

**Configuration and Deployment for Monitoring**

To set up CloudWatch, you will want to set up alarms for the metrics that need to have a close eye kept to them. For this deployment, billing was extremely important to track. In the CloudWatch service, select Alarms and create new alarm, select the billing metric and create alarm (UMGC, 2020). Graphical user interface, text, application

Description automatically generated

Figure 13: AWS CloudWatch Alarms (Amazon, 2020)

Once the alarm is created, you can create a dashboard, this dashboard is where all the metrics being tracked will live. While creating the dashboard, I added billing alarms as well as server usage and credit balances. Dashboards are customizable and can be sectioned off for separate needs, i.e., billing dashboard, performance dashboard, etc (Amazon, 2020). In order to add additional trackers to the dashboard, go to Add Widget and select the type of monitoring from the drop list on the bottom of the screen (UMGC, 2020)

Graphical user interface, text, application

Description automatically generated

Figure 14: AWS CloudWatch Dashboards (Amazon, 2020)

**Limitations and Resources**

All limitations sprung from my AWS Student account status, in a full account with IAM access, Data Broker could have leveraged larger instances, and I could have manually created the IAM role to have greater control over what Data Broker can sync and not sync in a bucket. Outside of the resources provided by UMGC, there was no additional information needed as I have completed the AWS Solution Architect Associate course in Linux Academy and work for Equinix Metal that actively works with customer to sync and set up hybrid environments in conjunction with their current cloud ecosystem.

References

Amazon. (2020a). *Amazon CloudWatch: Complete visibility of your cloud resources and applications (2:02)*. Amazon Web Services, Inc. https://aws.amazon.com/cloudwatch/

Amazon. (2020b). *Amazon Web Services (AWS) - Cloud Computing Services*. Amazon Web Services, Inc. https://aws.amazon.com/

Amazon. (2020c). *Internet Gateways - Amazon Virtual Private Cloud*. Internet Gateways. https://docs.aws.amazon.com/vpc/latest/userguide/VPC\_Internet\_Gateway.html

Amazon. (2020d). *Introduction to Amazon S3 (4:31)*. Amazon Web Services, Inc. https://aws.amazon.com/s3/

Amazon. (2020e). *Introduction to AWS CloudFormation*. Amazon Web Services, Inc. https://aws.amazon.com/cloudformation/

Amazon. (2020f). *NAT gateways - Amazon Virtual Private Cloud*. NAT Gateways. https://docs.aws.amazon.com/vpc/latest/userguide/vpc-nat-gateway.html

Amazon. (2020g). *VPC Fundamentals and Connectivity Options*. Amazon Web Services, Inc. https://aws.amazon.com/vpc/

NetApp. (2020). *NetApp Cloud Sync*. NetApp CloudSync. https://cloudsync.netapp.com/

UMGC. (2020a). *AWS CloudWatch*. AWS CloudWatch Monitoring Guide. https://leocontent.umgc.edu/content/umuc/tgs/cca/cca630/2208/course-resource-list/aws-cloudwatch.html?ou=516113

UMGC. (2020b). *Cloud Sync Service Guide*. Cloud Sync Service. https://leocontent.umgc.edu/content/umuc/tgs/cca/cca630/2208/course-resource-list/cloud-sync-service.html?ou=516113

UMGC. (2020c). *Create a VPC With Private and Public Subnets*. Create a VPC With Private and Public Subnets. https://leocontent.umgc.edu/content/umuc/tgs/cca/cca630/2208/course-resource-list/create-a-vpc-with-private-and-public-subnets.html?ou=516113

UMGC. (2020d). *S3 Bucket Deployment Guide*. S3 Bucket Deployment. https://leocontent.umgc.edu/content/umuc/tgs/cca/cca630/2208/course-resource-list/-s3-bucket-deployment-guide.html?ou=516113