Merrilton Bank

Cloud Infrastructure Design

D088 Final Assessment

Ryan Peterson

3-26-2022

Version 1.0

Western Governors University

CONTENTS

A. AUTHENTICATION PROCESS

Amazon Cognito facilitates development by assisting you in managing identities for customer-facing applications. Authentication for all users and devices will occur using Amazon AWS Cognito with User Pools. AWS Cognito permits device mapping to users and restricts the number of registered devices per user to 10 per the specifications. The International Mobile Equipment Identity (IEMI) number connects the user to their account. Is the device new or already registered? When users reach their limit, they must access their account via a known device to unregister unneeded devices.

AWS Cognito with User Pools will prompt users for extra verifications through SMS or a Time-Based One-Time Password generator such as Authy or Google Authenticator when logging in from an unregistered device. Unable to install Google Authenticator? Register your phone to receive an SMS text message with your one-time code. Lesser-known browsers or devices will also trigger additional verification requirements.

B. REMOTE ACCESS

In today's hybrid workplace, the distributed workforce requires various convenient authentication options. Remote access will utilize AWS Single Sign-On (AWS SSO) as the main access point for all employees to access the system. One of the most critical features of AWS SSO is that it allows you to create or connect your workforce identities in AWS only once and manage access centrally across your entire AWS organization. You have the option of restricting access only to your AWS accounts or cloud applications entirely. You can define, customize, and assign fine-grained access rights from a single interface). The AWS accounts, EC2 instances, and cloud applications allocated to your workforce users are accessible through a user portal.

We will configure AWS SSO instead of standard AWS account access management via AWS IAM. Employees will use bastion hosts as a method of access control to on-premises environments as well through this method. To further tighten access to AWS SSO, we will use SecurID to gain access to a variety of convenient authentication methods. Push-to-approve and one-time passcodes on mobile devices and passwordless solutions such as biometrics and FIDO-based authentication are examples of these approaches. In addition, SecurID provides the world's most widely used hardware token, which provides the highest level of protection for access to the most sensitive data.

The SecurID will be issued to all employees for use with remote access, and the dynamic token will provide the credential required for multi-factor authentication (MFA) to function. When the user accesses the bank resources either remotely or on-premises, they will be prompted for the SecurID to be inserted into a USB port on the computer and passed to the MFA login screen after entering their username and password.

C. APPLICATION SECURITY

A username and password combination with multi-factor authentication (MFA) via Google Authenticator, Authy, or an SMS one-time passcode texted to their registered device will be used by end-users to access their bank account information applications on mobile devices. The authentication process will also enforce MFA for end-users accessing their bank accounts via a browser on a tablet, laptop, or desktop computer. Those users will also be required to input their Google Authenticator or Authy passcode or get a one-time passcode via SMS text message pushed to their registered mobile device. Data in flight will be encrypted using TLS in both directions between the cloud and the end user's browser or mobile device used to access their account.

We will make several efforts to secure employee devices. The first step is to employ the Mobile OS encryption included with all current iOS and Android OS versions. Data in transit between the device application and the cloud will be secured using TLS in both directions. To authenticate a user, they need to provide a multi-factor authentication (MFA) code by SecurID. Employee laptops and desktops will have the same application security. The application will use a username and password along with a SecurID token or one-time code to authenticate the computer's VPN connection to the cloud services. The TLS security tunnel built during the VPN process will encrypt data in both ways.

D. NETWORK SECURITY

Employees will be required to use VPN to access cloud resources from registered devices for network security reasons. When the employee begins the VPN process to access cloud resources, the system prompts for a username and password combination associated with their account. If they are successful, the system asks for their MFA credentials, and they must supply the SecurID credentials they were issued. After successful authentication, the employee's device and the cloud resources communicate through a secure TLS tunnel to encrypt data in transit. If authentication fails, the system terminates the connection, and the employee must restart the process; if the employee fails three times in the authentication process, the system locks the account, and the employee must contact the helpdesk to have the account unlocked.

E. INTERNAL APIS

E1. BANK FRAUD SERVICES INTERNAL API

The fraud detection API will be written in Python and will utilize the PyOD library for anomaly detection. There will be a distributed database containing all transaction information, and GPU-enabled EC2 instances along with in-house servers will be used to analyze training data and continuously refine the models. PyOD has support for over 30 different modeling techniques, and these will continuously be compared to ensure the most accurate techniques possible are in use at any given time.

The models themselves will run as AWS Lambda functions. Each time a new transaction attempt is made, a GraphQL request will be triggered by the Lambda function to pull in data from the database. Since the anomaly detection models will continuously be refined, GraphQL will be more performant than a standard REST API, as the parameters needed by the models will likely be changing over time. GraphQL provides better response times for pulling in data from disparate sources compared to standard REST.

The models will return a score that reflects the likelihood that a given transaction is fraudulent. Based on that score, a transaction will either be approved, conditionally approved but flagged with a fraud warning sent, or be blocked outright. Customers will be able to receive push notifications or SMS about fraud warning and blocked transactions, where they can either mark the transactions as suspicious, or override the block and allow the transactions to process. This interactive nature will help to improve the models over time.

E2. BRANCH LOCATION INFORMATION INTERNAL API

The branch location API will be a relatively simple design. Like most of the rest of our systems, the backend will be written in Python using the AWS Location Services SDK and apps will use native functions to provide location services. The branch location API will run as a Lambda function that receives location data as a RESTFUL JSON payload and it will return a list of branch locations in the same manner.

E3. ALL LOG DATA INTERNAL API

Log data from all source will be fed into a FluentD server for standardization and then forwarded to a log aggregation system built with the ELK Stack (ElasticSearch, Logstash, Kibana). Logs from the following sources will be utilized:

- AWS CloudWatch
- AWS CloudTrail
- RSysLog on all servers
- Records of all API requests and responses

F. EXTERNAL APIS

F1. SINGLE SING-ON (SSO) EXTERNAL API

All employees will have remote access via AWS Single Sign-On (AWS SSO). One of the key features of AWS SSO is the ability to govern access across your whole AWS enterprise. Only your AWS accounts or cloud apps can be accessed. One interface defines, customizes, and assigns access rights). Your workforce's AWS accounts, EC2 instances, and cloud apps are accessible.

An alternative to AWS IAM will be AWS SSO. Employees will use bastion hosts to restrict on-premises access. We'll use SecurID to get useful AWS SSO authentication methods. Biometrics and FIDO-based authentication are examples of passwordless solutions. SecurID also offers the world's most widely used hardware token, giving the highest level of data security.

F2. CREDIT SCORE PROVIDER EXTERNAL API

It will be necessary for the banking mobile application to interact with the Experian Connect API in order to return the user's useless VantageScore credit score from the Experian Credit Bureau in order to offer the user with their credit score. It will run in AWS Lambda Functions and will be called whenever a user requests a credit score from within the application. Once the application has been invoked, it will submit a request to Experian over their Experian Connect API and receive back a response, which will then pass through the Lambda function and be shown to the user within the application.

F3. LOCATION DATA LOOK-UP EXTERNAL API

Mobile applications for Android and iOS platforms will use the AWS Location Service API to locate the closest branch office to the end-user. When a user clicks on a link in the app to locate the nearest branch office, the app calls the API. The app passes the data to the API in JSON format, and the returned response will provide the device with the user's position and accuracy information, among other things. In turn, the mobile application will use the returned data to generate a list of branches arranged from closest to farthest away from the user's present location.

G. DEPLOYMENT PLAN

G1. CLOUD BACKEND TIMELINE

From the date of approval of the beginning date, a one-year implementation timeline will be followed until the project is completed. A series of discovery meetings will be performed in order to determine the most effective cloud migration strategy for current apps and services. These sessions will serve as the starting point for the development process. This team will be tasked for developing the cloud infrastructure and testing the applications that will be deployed to it once it has completed its work.

G2. RESOURCES

The application development team will start working on mobile apps for Android and iOS devices, allowing banking customers to access their accounts from anywhere. We will provide the application developers with the various APIs needed to integrate them into the application. We estimate that the application development will take six months from start to finish. For the next six months of the implementation plan, independent auditing firms will conduct audits of the apps, security, and regulatory compliance. Additionally, auditor-identified items will be addressed and repaired in order to achieve certified compliance with all applicable rules.

G3. ESTIMATED COSTS

The design, development, and implementation of the project will cost $1.5 million. Because these are the only upfront Capital Expenditures required for the project, this cost can be amortized over three years at a rate of $500,000 per year. The expected annual operating cost of AWS services is $900,000, which can be deducted from corporate taxes as Operating Expenditures.

G4. REDUNDANCIES

Any redundancies between the previous and future designs of the system will be eliminated as a cost-saving measure. The most notable exception to this rule is the presence of a hot site for disaster recovery in the US-East 1 AWS Region. However, all other systems will be redesigned in a scalable and distributed fashion to allow the entire systems to expand with the bank's growth or contract during market downturns.

H. MAINTENANCE STRATEGY

The Change Advisory Board (CAB) will oversee patch management and redevelopment efforts following the ITIL v4 Change Management method. All modifications will be ranked for business continuity and must have a failover and failback strategy. The CAB will also assess all fixes and improvements to see if they are required for business continuity and will contact the respective teams for more information if needed.

Any final updates to cloud-based or mobile apps must be submitted to the CAB for evaluation and risk assessment. Once approved by the CAB, a change is executed during the change window stated in the request and validated upon deployment.

H1. PATCH MANAGEMENT

Short-term patch management focuses on any security vulnerabilities discovered in our application's systems. Upon discovery of a zero-day, a fix will be applied as quickly as possible by hot-patching any critical systems after sufficient testing and sign-off from the CAB.

We will adopt a three-week agile sprint cycle for a patch management plan for dealing with the long term. If no conflicts or mistakes are detected, patches not dealing with security will be tested in a sandbox for three cycles before being rolled out to the production application in a blue-green style.

H2. UPDATES AND REDEVELOPMENT

The application's Product Managers will oversee application updates, and all updates will follow a three-week sprint cycle for both short-term and long-term application updates. Long-term projects to introduce new functionalities will be implemented in a Minimum Viable Product (MVP) fashion during each sprint cycle, with iterative improvements after testing within the application's testing and staging environments. Once tested, all new functionalities will be introduced to the production environment by first validating user acceptance with A/B testing and then rolled out wide with a blue-green deployment.

When new cloud technologies emerge and new cloud management features become available, we will investigate, test, and integrate them to add value to the application. Cloud architects will use a three-week sprint cycle to update the backend for consistency. Before the production release, the architects will test any new cloud management changes for functionality and security in a test environment.

I. DISASTER RECOVERY PLAN

The proposal will use AWS regions and availability zones. It will replicate across availability zones, regions, and datacenters. Data replication from the datacenter to AWS for backups is enabled. The design will achieve performance, reliability, security, and scalability via AWS Regions and Availability Zones. The disaster recovery strategy will leverage AWS regions US-East 1 (Northern Virginia), US-East 2 (Ohio), and US-West 2 (California) (Oregon). The system will mirror all data to the Atlanta datacenter and route all traffic through it when operational. Whereas traffic from branches and remote sites will be directed to the nearest data center if Atlanta's datacenter fails, traffic from branches and remote sites will be directed to the US-East 1 hot site.

J. REGULATORY COMPLIANCE

AWS is famous for its shared responsibility model for security. AWS is responsible for the host operating system, virtualization, and physical security of the facilities. The client is responsible for administering the guest operating system, any associated applications, and the AWS security group firewall.

You'll have the most comprehensive set of compliance controls accessible when you use AWS. As a result, AWS provides the greatest number of security standards and compliance certifications of any product. PCI-DSS (Payment Card Industry Data Security Standard), FIPS 140-2 (Financial Information Processing Standard), and NIST 800-171 are among these standards and certifications. This enables the systems to comply with almost any regulatory agency on the planet.

Our design will adhere to all FDIC, PCI DSS, SOX, and municipal or governmental rules, regulations, and standards. Quarterly independent audits will confirm adherence to security protocols. The auditors will ensure that all data is encrypted and only personnel who need it may access it. Data in flight is also encrypted and secured using TLS and VPN connections to the cloud. Primary account numbers are stored but rendered illegible without adequate verification, CVV2 numbers are not stored, and PINs get hashed with SHA-512. An external auditor will annually check all papers for regulatory compliance. Independent auditors will monitor compliance with SEC, SOX, and FDIC rules.

K. SOURCES

“Multi-Factor Authentication - SecurID.” *SecurID*, www.securid.com, 12 Nov. 2021, https://www.securid.com/en-us/products/multi-factor-authentication/.

“AWS Single Sign-On | Workforce Identity Service | AWS.” *Amazon Web Services, Inc.*, aws.amazon.com, https://aws.amazon.com/single-sign-on/. Accessed 13 Mar. 2022.

Medan, Eran, and Yuri Duchovny. “Role-Based Access Control Using Amazon Cognito and an External Identity Provider | Amazon Web Services.” *Amazon Web Services*, aws.amazon.com, 15 Aug. 2020, https://aws.amazon.com/blogs/security/role-based-access-control-using-amazon-cognito-and-an-external-identity-provider/.

“Cloud Compliance - Amazon Web Services (AWS).” *Amazon Web Services, Inc.*, aws.amazon.com, https://aws.amazon.com/compliance/. Accessed 14 Mar. 2022.

yzhao062. “GitHub - Yzhao062/Pyod: A Comprehensive and Scalable Python Library for Outlier Detection (Anomaly Detection).” *GitHub*, github.com, 5 Mar. 2022, https://github.com/yzhao062/pyod.

“Pyod 0.9.8 Documentation.” *Pyod 0.9.8 Documentation*, pyod.readthedocs.io, https://pyod.readthedocs.io/en/latest/. Accessed 26 Mar. 2022.

“Amazon Location Service.” *Amazon Web Services, Inc.*, aws.amazon.com, https://aws.amazon.com/pm/location/. Accessed 26 Mar. 2022.

“Experian Connect - Credit Report and VantageScore for Consumers and Small Businesses.” *Experian Connect - Credit Report and VantageScore for Consumers and Small Businesses*, www.experian.com, https://www.experian.com/connect/. Accessed 26 Mar. 2022.