Data Backup in Mexico

March 2024

Notices

1. Informational Guidance. Customers are responsible for making their own independent assessment of the information contained in this Guide: Options to Maintain a Data Backup in Mexico (the “Guide”). This Guide is (a) informational only, (b) non-binding, (c) represents current AWS product offerings and practices, which are subject to change without notice, (d) does not create any commitments or assurances from AWS and its affiliates, suppliers, or licensors, and (e) provided “as is” without warranties, representations, or conditions whatsoever.
2. Agreement. All use of services provided by Amazon Web Services (AWS) is governed by the AWS Customer Agreement available at http://aws.amazon.com/agreement/ (or other definitive written agreement as may be agreed between AWS and customers governing the use of AWS's services) (as applicable, the ”Agreement“). This Guide is not part of, nor does it modify or supplement, any agreement between AWS and its customers.
3. Confidentiality. The existence and contents of this Guide and any activities described herein are confidential information and is subject to the Nondisclosure Agreement (“NDA”) between AWS and customers. No part of this Guide may be disclosed without AWS's prior written consent.

Contents

Abstract

This Guide is intended for Mexican financial institutions regulated by the National Banking and Securities Commission (Comisión Nacional Bancaria y de Valores (CNBV)) and the Mexican Central Bank (Banco de México), which use Amazon Web Services (AWS) services. Specifically, this Guide describes options to maintain a data backup within Mexican territory, describes AWS backup and replication services, and provides AWS reference architectures.

Overview

Amazon Web Services (AWS) offers IT services in categories ranging from compute, storage, database, and networking, to artificial intelligence and machine learning. Financial services institutions (FSIs) can use AWS services to modernize and automate their technology infrastructure, meet rapidly changing customer behaviors and expectations, and drive business growth. Through continuous innovation, AWS provides strong security systems, a breadth and depth of services, extensive industry expertise, and an expansive partner network.

The CNBV supervises different sectors of the financial ecosystem in Mexico, including but not limited to banks, broker dealers, credit unions, financial technology institutions (FinTechs), and community and rural financial corporations (SOFIPOs and SOCAPs). The specific regulatory requirements related to outsourcing of technology services might vary depending on the sector.

Under Mexican regulation, and in certain scenarios, FSIs from different sectors are required to maintain a copy of accounting and transactional records within Mexican territory to ensure operational continuity if their databases are serviced by a third party whose data centers are located outside of Mexico. The backup copy can be maintained on-premises or on third-party infrastructure, as long as it is in Mexico and available for use in case of a contingency. These records must be kept in a format that permits their consultation, operation, and use; irrespective of whether the service contracted with the third party is available.

Table 1 – Examples of backup requirements by sector.

Sector	Reference
Banks and SOFOM	Annex 52 (I)(e)
Electronic payment funds institutions (IFPE)	Article 49 (IV)
Crowdfunding institutions (IFC)	Article 86 (IV)
Brokerage houses	Annex 12 (I)(e)
SOFIPO	While there is no regulation under which SOFIPOs are expressly required to comply with the requirement to maintain a data backup in Mexico, a number of SOFIPOs have received notifications from the regulator indicating that they are subject to this requirement.

Data backup in Mexico when using AWS

In this Guide, we describe architectures that FSI customers might consider implementing to maintain a data backup of their AWS environments in Mexico. Each customer's architecture is different and the AWS services they choose to use are determined by their needs. Customers might consider several options of architectures to comply with Mexican regulation, depending on their needs and architecture.

AWS hybrid cloud services deliver a consistent AWS experience wherever customers need it—from the cloud, on premises, and at the edge. AWS hybrid cloud services include:

- AWS Local Zones: A type of AWS infrastructure deployment that places compute, storage, database, and other select services closer to population, industry, and IT centers, enabling customers to deliver applications that require single-digit millisecond latency to end-users. AWS has a Local Zone in Queretaro, Mexico.

- AWS Outposts: A family of fully managed solutions that delivers AWS infrastructure and services to virtually any on-premises or edge location for a consistent hybrid experience. Outposts allows customers to extend and run native AWS services on-premises, and is available in a variety of form factors, from 1U and 2U Outposts servers up to 42U Outposts racks, which can be a single or multiple deployments. With Outposts, customers can run certain AWS services locally and connect to a broad range of services available in the local AWS Region, as well as run applications and workloads on-premises using familiar AWS services, tools, and APIs. Outposts supports workloads and devices requiring low latency access to on-premises systems, local data processing, data residency, and application migration with local system interdependencies.
- AWS Snow Family: Purpose-built devices to cost effectively move petabytes of data offline. Snow devices are field-tested for the most extreme conditions and can deliver high security and ruggedization into compute and storage-compatible devices. The range of Snow Family device options is designed to optimize for space- or weight-constrained environments, portability, and flexible networking options.

Figure 1 shows—at a high level—a number of options customers might consider for maintaining a data backup in Mexico.

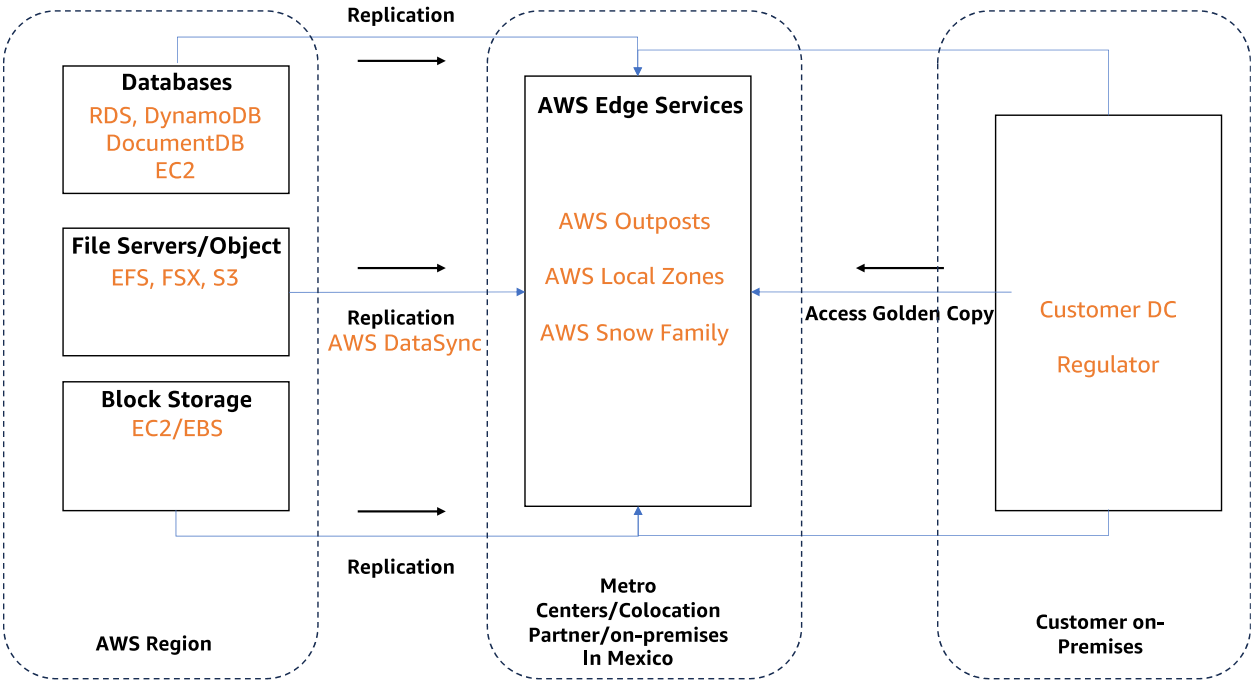

Figure 1 – Options for data backups in Mexico

Options for data replication

Customers might also consider a number of options to replicate data into Mexico from different services located in an AWS Region outside of Mexico, including:

- Database replication
- File based and object storage replication
- Partner solutions for Amazon Elastic Compute Cloud (Amazon EC2)

Database replication

In this section, we explore replication strategies for Amazon Relational Database Service (Amazon RDS), Amazon DynamoDB, Amazon DocumentDB, and Amazon EC2. In addition to the options presented in this section, see Appendix C for a summary of other options available for customers to maintain a copy of their data within Mexican territory.

Data source is Amazon RDS

Amazon RDS is a collection of managed services that makes it simple to set up, operate, and scale databases in the cloud. RDS supports seven database engines: Amazon Aurora, PostgreSQL, MySQL, MariaDB, Oracle Database, and SQL Server.

To back up an Amazon RDS database to AWS edge services, our customers can use AWS Database Migration Service (AWS DMS) or AWS DataSync.

- **Option 1**: AWS DMS is a managed migration and replication service that helps customers move their database and analytics workloads to AWS quickly, securely, and with minimal downtime and little to zero data loss. AWS DMS supports migration between over 20 database and analytics engines, such as Oracle to Amazon Aurora MySQL-Compatible Edition, MySQL to Amazon RDS for MySQL, Microsoft SQL Server to Amazon Aurora PostgreSQL-Compatible Edition, MongoDB to Amazon DocumentDB (with MongoDB compatibility), Oracle to Amazon Redshift, and Amazon Simple Storage Service (Amazon S3).

  AWS DMS provides ongoing replication of data, keeping the source and target databases in sync. AWS DMS can use many of the most popular databases as a target for data replication, available engines are listed in the AWS DMS User Guide. Customers consuming Amazon RDS services in an AWS Region can maintain a copy of their data in Mexico on the following AWS edge services:

  - Local Zones
  - Outposts
  - AWS Snowball Edge

  With the given source on Amazon RDS, AWS DMS can replicate the data for a select list of engines with a homogeneous database replication using the same target engine whether on Local Zones, Outposts, or Snowball Edge on top of Amazon EC2, a database replication solution using AWS Schema Conversion Tool, also available in AWS for no additional charge. This option allows customers to keep a reliable copy of data in Mexico and provides a ready to use database that can be consumed for read operations in the Local Zone in Mexico. During ongoing replication, it's critical to identify the network bandwidth between your source database system and your DMS replication instance. Make sure that the network doesn't cause any bottlenecks during ongoing replication.

- **Option 2**: AWS DataSync is an online data transfer service that simplifies, automates, and accelerates moving data between storage systems and services. Customers can use DataSync to export Amazon RDS snapshots. When customers export a database snapshot, RDS extracts data from the snapshot and stores it in an S3 bucket. The data is stored in an Apache Parquet format that is compressed and consistent. For more information, customers can see Exporting DB snapshot data to Amazon S3. After the data is stored in an S3 bucket, customers can synchronize those buckets to AWS edge services to maintain a backup in Mexico:
  - Local Zones
  - Outposts
  - Snowball (file interface): the file interface exposes a Network File System (NFS) mount point for each bucket on a Snowball Edge device. Customers can mount the file share from their NFS client using standard Linux, Microsoft Windows, or macOS commands. Customers can also use standard file operations to access the file share.
  - AWS Snowcone Edge DataSync agent: customers can use AWS OpsHub to create an AWS DataSync agent on their Snowcone device. They can then use DataSync to transfer files between their device and Amazon S3, Amazon Elastic File System (Amazon EFS), or Amazon FSx for Windows File Server in the AWS Cloud.

Database snapshots can be moved to an NFS and hosted in an EC2 instance.

- **Option 3**: Customers can back up their Amazon RDS databases to .csv files to Amazon S3 in an AWS Region, then those .csv files can be copied to Mexico on Local Zones, Outposts, or Snowball Edge. CSV files are simple to read and consumable for anyone. To do this, customers can replicate their RDS schemas and databases to S3 in-Region using AWS DMS, and then use DataSync to move those files to an NFS server in:
    - Local Zones
    - Outposts
    - Snowball (file interface)
    - Snowcone Edge DataSync agent
- **Option 4**: Customers can back up an Amazon RDS read replica from an AWS Region to Outposts using the AWS Outposts RDS read replica function, recently announced for MySQL and PostgreSQL. See: Amazon RDS on AWS Outposts now supports read replicas for MySQL and PostgreSQL.

Note: It's important for customers to consider that they can use solutions that automate their backup processes. See Appendix A for architecture details.

Data source is Amazon DynamoDB

Amazon DynamoDB is a serverless, NoSQL, fully managed database service designed to run high-performance applications at any scale. DynamoDB offers built-in security, continuous backups, automated multi-Region replication, in-memory caching, and data import and export tools. When using DynamoDB, customers might consider Option 1 to keep a copy of their data in Mexico:

- **Option 1**: customers can export data from a DynamoDB table at any time within their point-in-time recovery window to an S3 bucket using DynamoDB table export. DynamoDB table export is a fully managed solution for exporting DynamoDB tables at scale and is faster than other workarounds involving table scans. After the table is exported to Amazon S3, customers can use DataSync to move those Amazon S3 objects to Mexico on:
    - Local Zones
    - Outposts
    - Snowball (File Interface)

- Snowcone Edge DataSync Agent

  For Local Zones where Amazon S3 isn't available, customers can use DataSync to synchronize S3 objects created from a DynamoDB table export to an NFS hosted in Amazon EC2.

  DataSync allows customers to move data to and from AWS storage services. Because DataSync supports Amazon S3 on Outposts, customers can automate data transfer between their Outposts and AWS Regions, choosing what to transfer, when to transfer, and how much network bandwidth to use.

There are other methods to export data from DynamoDB to Amazon S3 using Amazon EMR, AWS Glue, and AWS Data Pipeline. For more information, see: How can I back up a DynamoDB table to Amazon S3? For AWS Glue, there's an automated solution to export data from DynamoDB to Amazon S3. See Appendix A for architecture details.

Data source is Amazon DocumentDB

Amazon DocumentDB is a fully managed, native JSON document database that makes it simple and cost effective to operate critical document workloads at virtually any scale without managing infrastructure. Amazon DocumentDB simplifies customers' architectures by providing built-in security best practices, continuous backups, and native integrations with other AWS services. When using Amazon DocumentDB customers might consider the following option to keep a copy of their data in AWS edge services in Mexico:

- **Option 1**: Amazon DocumentDB customers can use the mongodump, mongorestore, mongoexport, and mongoimport utilities to back up or move data in and out of their Amazon DocumentDB cluster. The mongoexport tool exports data from Amazon DocumentDB to JSON, CSV, or TSV file formats. The mongoexport tool is the preferred method to export data that needs to be human or machine readable. Customers can find more information about dumping, restoring, importing, and exporting data in the Amazon DocumentDB User Guide. After data is exported, it can be copied to Amazon S3 using multipart upload. Afterwards, customers can use DataSync to move the JSON, CSV, or TSV files to an NFS server in:
  - Local Zones
  - Outposts
  - Snowball Edge

Amazon DocumentDB copies can be moved to an NFS and hosted in an EC2 instance.

Data source is a database hosted in Amazon EC2

Amazon EC2 offers a broad and deep compute environment, with over 500 instances and the choice of the latest processor, storage, networking, operating system, and purchase model to help customers best match the needs of their workload. Amazon EC2 can be used to host databases supported in a variety of operating systems.

- **Option 1**: When using Amazon EC2 to host a database instead of AWS database services, customers must follow the database engine vendor procedures to extract the data stored there. Customers can export their data and upload those exports into Amazon S3. Then the exported data can be moved to Mexico using DataSync replicated with an NFS server hosted in Amazon EC2 and using the following AWS edge services:
    - Local Zones
    - Outposts
    - Snowball (file interface)
    - Snowcone Edge DataSync agent
- **Option 2**: AWS DMS is a managed migration and replication service that helps customers move their database and analytics workloads to AWS quickly, securely, and with minimal downtime and zero data loss. AWS DMS supports migration between over 20 database and analytics engines, such as Oracle to Amazon Aurora MySQL-Compatible Edition, MySQL to Amazon Relational Database (RDS) for MySQL, Microsoft SQL Server to Amazon Aurora PostgreSQL-Compatible Edition, MongoDB to Amazon DocumentDB, Oracle to Amazon Redshift, and Amazon S3.

  AWS DMS provides ongoing replication of data, keeping the source and target databases in sync. DMS can use many of the most popular databases as a target for data replication. The available engines available are listed in the user guide. Customers consuming Amazon RDS services can maintain a copy of their data in Mexico in the following AWS edge services:
    - Local Zones
    - Outposts

- Snowball

  The Amazon RDS copy can be moved to any chosen engine using AWS DMS. After that, the database can be run in Local Zones or other AWS edge services using any database engine available in AWS Marketplace. The engine can be hosted in an EC2 instance. Amazon EC2 is available in AWS edge services. The target database will be accessible using traditional SQL ports with SQL language.

Appendix A includes a reference architecture that supports data exports from Amazon RDS to Amazon S3. The reference architecture can be used as a base for these types of backups. In Creating a Data Snapshot Using mysqldump, customers can find more information on how to perform snapshots for MySQL, where after data is exported, it can be copied to Amazon S3 using multipart upload.

File-based replication

Cloud file storage is a method for storing data in the cloud that provides servers and applications access to data through shared file systems. This compatibility makes cloud file storage ideal for workloads that rely on shared file systems and provides integration without code changes. AWS provides fully managed file system services that help customers address the diverse needs of their file-based applications and workloads. AWS offers customers the following file system services optimized for their applications and use cases:

- Amazon EFS provides a simple, serverless, set-and-forget elastic file system that allows customers to create and configure shared file systems quickly for AWS compute services.
- Amazon FSx makes it simple and cost effective to launch, run, and scale feature-rich, high-performance file systems in the cloud. Amazon FSx allows customers to choose between four widely used file systems: NetApp ONTAP, OpenZFS, Windows File Server, and Lustre.

For file-based replication, customers should first identify which AWS file-based services contain the files that they want to back up in Mexico. The mechanism to have a backup in Mexico will depend on the services used.

Data source is Amazon FSx or Amazon EFS

For Amazon FSx or Amazon EFS, there's a mechanism to batch replicate to Mexico with AWS DataSync, which will be covered in the section following this one.

- **Option 1**: Customers can replicate files from Amazon FSx or EFS to on-premises using DataSync, which can batch replication to in-country infrastructure to run task periodically to keep an update copy of files from source in-Region to a target in Mexico on a file system running on:
    - Local Zones
    - Outposts
    - Snowball

If customers are using Outposts, it's possible to use Amazon S3 on AWS Outposts plus replication with DataSync from an AWS Region to Outposts, this mechanism is reviewed in the following section.

Customers who choose to use the Local Zone in Mexico can use DataSync to move data from Amazon EFS or Amazon FSx from AWS Regions to the Local Zone in Mexico using Amazon Linux AMI plus NFS or any Marketplace independent software vendor (ISV) solutions that support NFS server on top of Amazon EC2.

Data transfer services in AWS

There are different services that can be used to transfer data in and out of AWS infrastructure to on-premises storage. It's also possible to use services to transfer data from an AWS Region to AWS edge services (for example, Local Zones).

AWS DataSync

DataSync is a secure, online service that automates and accelerates moving data between on-premises and AWS storage services. DataSync can copy data between NFS shares, Server Message Block (SMB) shares, Hadoop Distributed File Systems (HDFS), self-managed object storage, Snowcone, S3 buckets, Amazon EFS file systems, and Amazon FSx for Windows File Server file systems.

In this Guide, we focus on mechanism options customers can consider for transferring data from Amazon S3, EFS, or Amazon FSx to their on-premises environments in Mexico using DataSync. DataSync supports transferences using the internet, virtual private network (VPN) such as AWS Site-to-Site VPN, or AWS Direct Connect to replicate data between on-premises environments and AWS infrastructure. Data is encrypted in-transit using TLS, irrespective of which connection method customers choose to use.

For more information about the benefits, key components, and how to use DataSync, see Appendix D on data transfer services.

A DataSync hands-on lab is available. DataSync reference architectures can be found in Appendix A: AWS Reference Architectures.

DataSync allows customers to copy data between AWS storage services and the AWS Local Zone in Mexico. The DataSync target might be an NFS running on top of Amazon EC2 and Amazon Elastic Block Store (Amazon EBS) (both services are available in any Local Zone). Customers can deploy a file server using Amazon Linux 2 AMI on EC2 in Local Zones to comply with this requirement.

When to use AWS DataSync to move data

The following are scenarios in which it could be useful to use DataSync to move data from a Region to a Local Zone or from on-premises to a Local Zone:

- Source data is on-premises NFS or SMB.
- On-premises infrastructure is available for running a DataSync agent.
- Data will be transferred to (or from) Amazon EFS, Amazon S3, or Amazon FSx for Windows File Server.
- Need to support both one-time and incremental transfers.
- Source or destination – Customers can transfer from their on-premises servers to AWS and the other way around. Customers can also use DataSync to transfer data between AWS services. In the case of the Local Zone in Mexico, customers can use DataSync to replicate data from the AWS Region to an NFS running on an EC2 instance.
- Need online migration for active data sets or replication for business continuity.

Customers only pay for the amount of data transferred per gigabyte according to the source Region. Customers can refer to AWS DataSync pricing for more details on pricing. See Appendix D to learn more about data transfer options.

Using AWS Outposts

Outposts has been available in Mexico since July 2020. Outposts are ideal for workloads with low latency, local data processing, or data localization needs. Mexican financial institutions can move data, including personal information, outside the country, so long as they comply with local regulation. As previously mentioned, certain financial institutions are required to maintain a copy of certain records in national territory. The data must be located in Mexico and be accessible at all times.

Outposts data remains available when there are disconnection-causing events in AWS Regions. If there is a disconnection-causing event, instances running on an Outpost continue to run and are accessible from on-premises networks through the Outpost local gateway (LGW). Local workloads and services might be impaired or fail if they rely on services in the Region. Mutating requests (like starting or stopping instances on the Outpost), control plane operations, and service telemetry (for example, Amazon CloudWatch metrics) will fail while the Outpost is disconnected from the Region. For further information, see the AWS Outposts High Availability Design and Architecture Considerations whitepaper.

When planning to use Outposts, it's important for customers to consider the AWS services available in Outposts.

Outposts is designed to operate with a constant and consistent connection between customers' Outposts and an AWS Region. Outposts extends an Amazon Virtual Private Cloud (Amazon VPC) from a Region to an Outpost with the VPC components that are accessible in the Region, including internet gateways, virtual private gateways, Amazon VPC transit gateways, and VPC endpoints. An Outpost is homed to an AWS Availability Zone in the Region and is an extension of that Availability Zone that customers can use for resiliency. Customers can run their workloads using both the Region and AWS Outposts. Outposts supports workloads and devices requiring low latency access to on-premises systems, local data processing, data residency, and application migration with local system interdependencies.

Outposts is a good fit to maintain a backup in Mexico in the following situations:

- When customers are already using Outposts for other purposes. Using Outposts exclusively to keep a backup of the data might not be a cost-effective solution.
- When customers are using Outposts for processing main workloads and want to also access data locally in Mexico.

Amazon S3 on Outposts delivers object storage to customers' on-premises Outposts environments to meet data localization needs. Using the Amazon S3 APIs and features available in AWS Regions today, S3 on Outposts makes it simpler for customers to store and retrieve data on their Outposts, as well as secure the data, control access, tag, and report on it.

Amazon S3 on Outposts provides a new S3 storage class, named OUTPOSTS, which uses the S3 APIs and is designed to durably and redundantly store data across multiple devices and servers on customers' Outposts. DataSync, a service that makes it easier to move data to and from AWS storage services, supports S3 on Outposts. Thus, customers can automate data transfer between their Outposts and AWS Regions, choosing what to transfer, when to transfer, and how much network bandwidth to use. When using DataSync to access S3 on Outposts, customers must launch the agent in a VPC that's allowed to access their S3 access point and activate the agent in the parent Region of the Outpost.

To learn more about working with Amazon S3 on Outposts endpoints, see Working with Amazon S3 on Outposts and the AWS DataSync User Guide.

Connectivity between on-premises and AWS

AWS Direct Connect

Direct Connect is a networking service that provides an alternative to using the internet to connect to AWS. Using Direct Connect, data that would have previously been transported over the internet is delivered through a private network connection between customers' facilities and AWS. In many circumstances, private network connections can reduce costs, increase bandwidth, and provide a more consistent network experience than internet-based connections. AWS services including Amazon EC2, Amazon VPC, Amazon S3, and DynamoDB can be used with Direct Connect. Direct Connect is available in Mexico. Customers can use the AWS Direct Connect tab on the AWS Management Console to create a new connection. When requesting a connection, customers will be asked to select a Direct Connect location, the number of ports, and the port speed. Customers can work with a Direct Connect Delivery Partner if they require assistance extending their office or data center network to a Direct Connect location.

AWS Direct Connect Delivery Partners have passed additional validation from the AWS Service Delivery Program for enabling customers to access Direct Connect. Additionally, Direct Connect Delivery Partners have established interconnect monitoring and are authorized to provision capacities greater than 500 Mbps.

There are Direct Connect Delivery Partners in Mexico—including C3ntro Telecom, Alestra, Telmex, MCM Telecom, and Transtelco—who can help customers establish dedicated connection from on-premises to AWS. The updated list of Direct Connect Delivery Partners can be found on the AWS Direct Connect Delivery Partners site.

AWS site-to-site VPN

By default, instances that customers launch into an Amazon VPC cannot communicate with customers' own (remote) networks. However, customers can enable access to their remote network from their VPC by creating an AWS Site-to-Site VPN connection, and configuring routing to pass traffic through the connection. These types of connections go over the internet, meaning they are not dedicated, and the maximum supported throughput is 1.25 Gbps comparted to 100 Gbps as the maximum speed for Direct Connect (available in specific locations).

Enabling the Mexico-Queretaro Local Zone

The list of AWS services that are available in the Local Zone in Queretaro, Mexico, can be found in AWS Local Zones features. In order to enable a Local Zone for their environments, customers must go to the console and enable it under the Amazon EC2 Zones tab. Customers do not incur any additional cost when they enable the Mexico Local Zone for their environments. However, billing might apply after customers begin to build services on top of the Local Zone. For additional information, see Getting Started with AWS Local Zones.

Appendix A – AWS reference architectures

This section presents architecture options that customers can use as reference for backing up and transferring data from AWS to on-premises.

Database snapshots and data transfer to on-premises

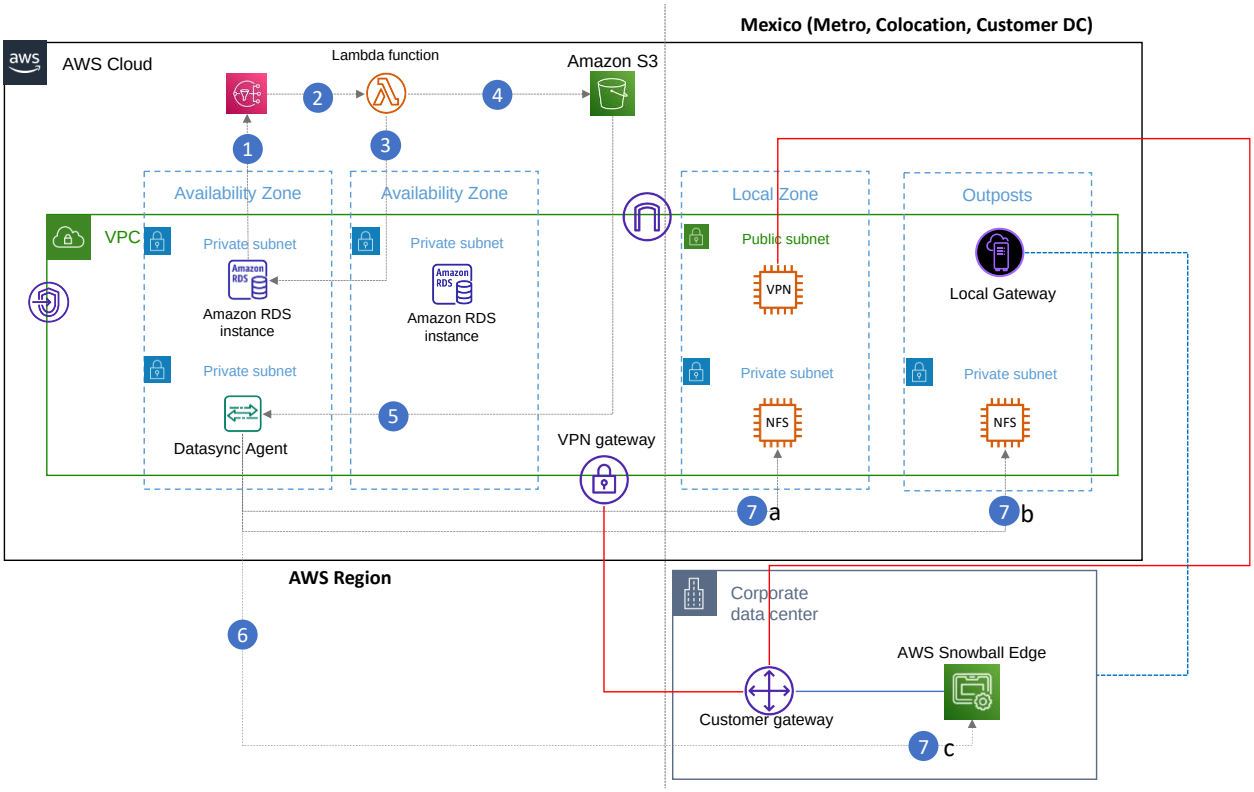

Figure 2 – Database snapshots and data transfer architecture.

This architecture handles the following:

1. After a snapshot of Amazon RDS is taken (manually or through automation), the back-up process starts. That generates an event that is sent to Amazon Simple Notification Service (Amazon SNS).
2. An AWS Lambda function is invoked to capture the SNS event and to export the snapshot API from Amazon RDS (start_export_task API).
3. A request to export the snapshot to Amazon S3 is made.
4. Export is finished and sent to S3.

5. A DataSync task is run periodically to take data from S3 and send it to on-premises storage through the DataSync agent.
6. Files are transferred by the DataSync agent. Connectivity between Amazon VPC or the DataSync VPC endpoint and on-premises can be done through VPN such as AWS Site-to-Site VPN or Direct Connect. For more information, customers can refer to Transferring files from on premises to AWS and back without leaving your VPC using AWS DataSync.
7. Files arrive to a server on-premises, which allows data to be located in Mexico and be accessible even when there's an event in the Region. Customers must plan mechanisms to use the data they have backed up.
    a. Storing the data in Local Zones
    b. Storing the data in Outposts
    c. Storing the data in Snowball Edge

Note: A similar architecture can be found in RDS Snapshot Export to S3 Pipeline—for purposes of a scenario in which customers want to maintain a data backup in Mexico, Amazon Athena can be removed as it isn't necessary to perform analysis on the exported snapshots.

Amazon EFS to on-premises

Amazon EFS to on-premises data transfer requires at least one EC2 instance with a file system mounted over EFS, which is constantly writing and reading to and from the shared file system.

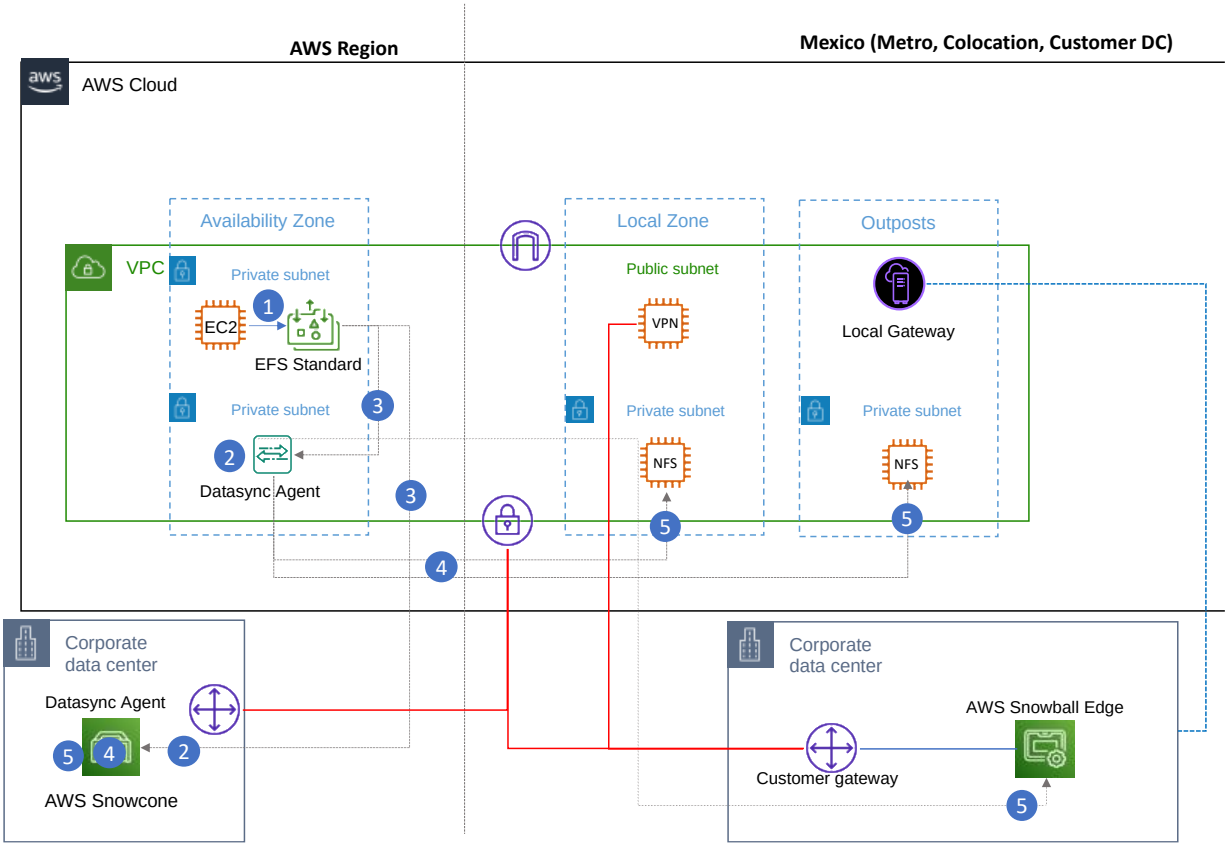

Figure 3 – EFS to on-premises architecture.

This architecture handles the following:

1. There is at least one EC2 instance with a file system mounted over Amazon EFS, which is constantly writing and reading to and from the shared file system.
2. A DataSync task is configured to be run periodically (this is configurable) with EFS as the source and an NFS server as the target.
3. After a DataSync task is run, file transfer starts.
4. Files are received by the DataSync agent and are sent to the NFS server.

5. Files are stored in the server located in Mexico. That data can be used and queried in case of an event in the AWS Region. Connectivity between VPC or the DataSync VPC endpoint and on-premises storage can be done through VPN or DirectConnect. For more information, customers can review Transferring files from on premises to AWS and back without leaving your VPC using AWS DataSync.

A similar architecture is mentioned in Getting started with Amazon Elastic File System. However, for purposes of a scenario in which customers want to maintain a data backup in Mexico, source and destinations must be changed to match exactly the explained architecture. A similar architecture can be implemented for transferring files from Amazon S3 to an NFS server (just changing EFS for S3).

Amazon DynamoDB to on-premises

DynamoDB to on-premises data transfer uses a data pipeline that launches an Amazon EMR cluster to read data from DynamoDB and export the data to an S3 bucket.

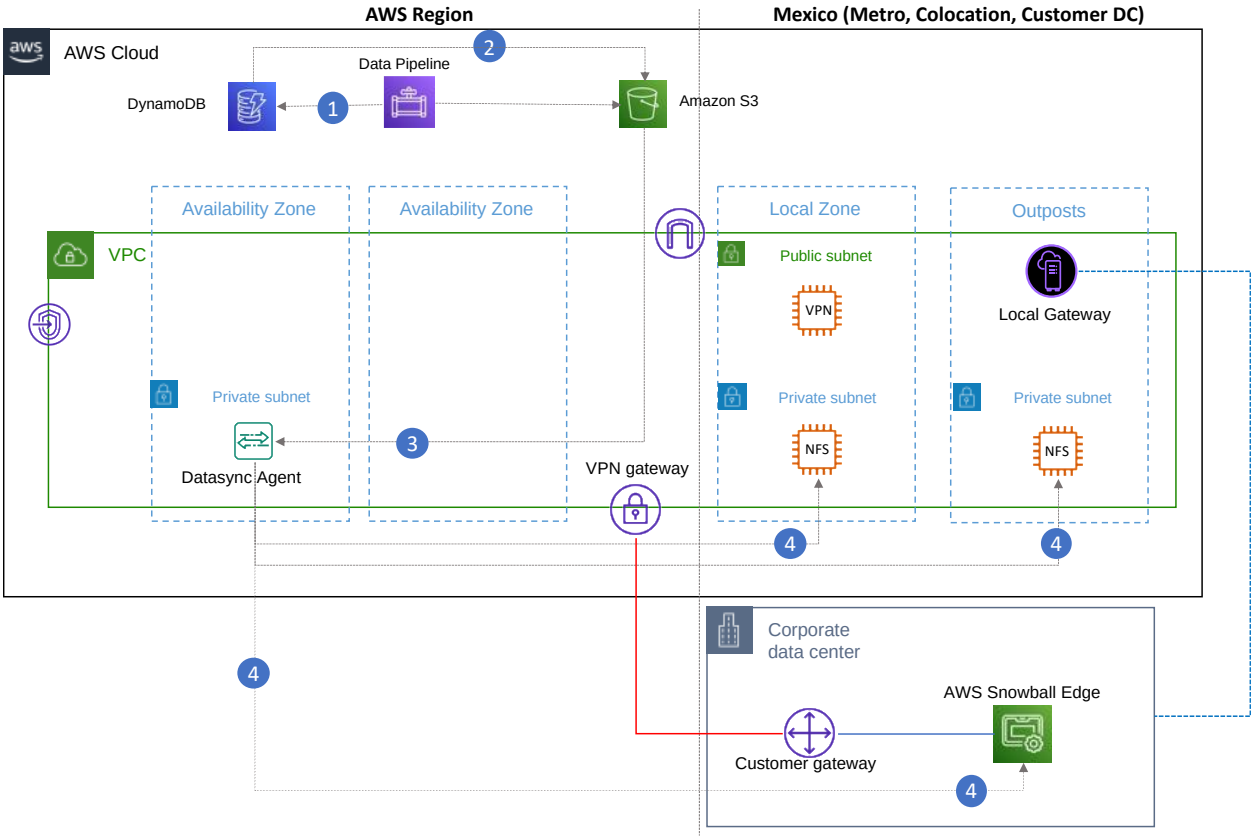

Figure 4 – DynamoDB to on-premises architecture

This architecture handles the following:

1. Data Pipeline launches an Amazon EMR cluster to perform the export.
2. Amazon EMR reads the data from DynamoDB and writes the data to an export file in an S3 bucket.
3. A DataSync task is configured to be run periodically (this is configurable) use the S3 bucket or folder where DynamoDB data was stored as the source and an NFS server as the target.
4. Files go through the DataSync Agent to be sent to the NFS server.

Appendix B – Additional resources

The following are additional resources to help banks, traditional credit institutions, and other financial institutions think about security, compliance, and designing a secure and resilient AWS environment.

- AWS Compliance Center Mexico
- AWS User Guide to Regulations Applicable to Credit Institutions in Mexico
- AWS Compliance Quick Reference Guide: AWS has many compliance-enabling features that customers can use for their regulated workloads in the AWS Cloud. These features allow customers to achieve a higher level of security at scale. Cloud-based compliance offers a lower cost of entry, simpler operations, and improved agility by providing more oversight, security control, and central automation.
- AWS Well-Architected Framework: The AWS Well-Architected Framework has been developed to help cloud architects build the most secure, high-performing, resilient, and efficient infrastructure possible for their applications. This framework provides a consistent approach for customers and partners to evaluate architectures and provides guidance to help customers implement designs that scale application needs over time. The AWS Well-Architected Framework consists of six pillars: operational excellence, security, reliability, performance efficiency, cost optimization, and sustainability.
    - AWS produced whitepapers on the six pillars of the AWS Well-Architected Framework: Operational Excellence Pillar; Security Pillar; Reliability Pillar; Performance Efficiency Pillar; Cost Optimization Pillar, and the Sustainability Pillar.
- Global financial services regulatory principles: AWS has identified five common principles related to financial services regulation that customers should consider when using AWS Cloud services and specifically when applying the AWS Shared Responsibility Model to their regulatory requirements. Customers can access a whitepaper on these principles at AWS Artifact. Customers must accept a non-disclosure agreement to access this whitepaper.

- National Institute of Standards and Technology (NIST) Cybersecurity Framework (CSF): the AWS NIST Cybersecurity Framework (CSF): Aligning to the NIST CSF in the AWS Cloud whitepaper demonstrates how public and commercial sector organizations can assess the AWS environment against the NIST CSF and improve the security measures they implement and operate (security in the cloud). The whitepaper also provides a third-party auditor letter attesting to the AWS Cloud offerings' conformance to NIST CSF risk management practices (security of the cloud). Banks and traditional credit institutions can use NIST CSF and AWS resources to elevate their risk management frameworks.
- Security, Identity, and Compliance Whitepapers

Appendix C – Database backup options

This table summarizes the database backup options.

Database service	Backup tool options	Comments
Amazon RDS	• Automated backups • Manual snapshots • Export snapshots to S3 • Manual exports/dumps depending on the database engine	Automated and manual snapshots help to restore data in an AWS Region. Exported snapshots can help to bring data to on-premises or to the Local Zone in Mexico, as well as manual database exports depending on the database engine.
DynamoDB	• On demand • Point-in-time recovery • Export snapshots to S3 with point-in-time recovery • Use AWS Glue to export data to S3.	Exporting options allow customers to bring their data on-premises or to the Mexico Local Zone. On-demand and point-in-time recovery options are only applicable to restore data in AWS Regions.
Amazon DocumentDB	• Manual snapshots • Automated snapshots • Using MongoDB tooling (mongoExport tool)	MongoDB is the preferred method to extract data in a human readable format and for bringing data to on-premises or to the Local zone in Mexico. Manual and automated snapshots are only applicable to restore data in AWS Regions.
EC2 hosted databases	• Using vendor database engine methods to export • Using third-party tooling for EC2 such as Veeam or Druva	Using third-party tools will have additional costs depending on the selected partner.

Appendix D – AWS DataSync

AWS DataSync employs an AWS-designed transfer protocol—decoupled from the storage protocol—to accelerate data movement. The protocol performs optimizations on how, when, and what data is sent over the network. Network optimizations performed by DataSync include incremental transfers, in-line compression, and sparse file detection, as well as in-line data validation and encryption.

Connections between the local DataSync agent and the in-cloud service components are multi-threaded, maximizing performance over customers' wide area networks (WANs). A single DataSync task is capable of fully utilizing 10 Gbps over a network link between a customer's on-premises environment and AWS.

DataSync removes many of the infrastructure and management challenges customers face when writing, optimizing, and managing their own copy scripts, or deploying and tuning heavyweight commercial transfer tools. With DataSync, customers can simplify their infrastructure, and stay in control with built-in monitoring and retry mechanisms to maintain successful data transfers.

DataSync comes with a built-in scheduling mechanism, allowing customers to periodically run data transfer tasks to detect and copy changes from their source storage system to the destination. Customers can schedule their tasks using the AWS DataSync console or AWS Command Line Interface (CLI) without writing scripts to manage repeated transfers. Task scheduling automatically runs tasks on customers' configured schedules with hourly, daily, or weekly options provided directly in the console.

Customer data is encrypted in transit between the DataSync agent and the DataSync service using Transport Layer Security (TLS). DataSync supports using default at-rest encryption for S3 buckets. DataSync also supports encryption of data at rest and in transit for Amazon EFS and Amazon FSx. DataSync supports using default encryption for S3 buckets, Amazon EFS file system encryption of data at rest, and Amazon FSx for Windows File Server encryption at rest and in transit.

DataSync verifies that customers' data arrives intact. For each transfer, the service performs integrity checks both in transit and at rest. These checks verify that the data written to customers' destinations matches the data read from their source, validating consistency.

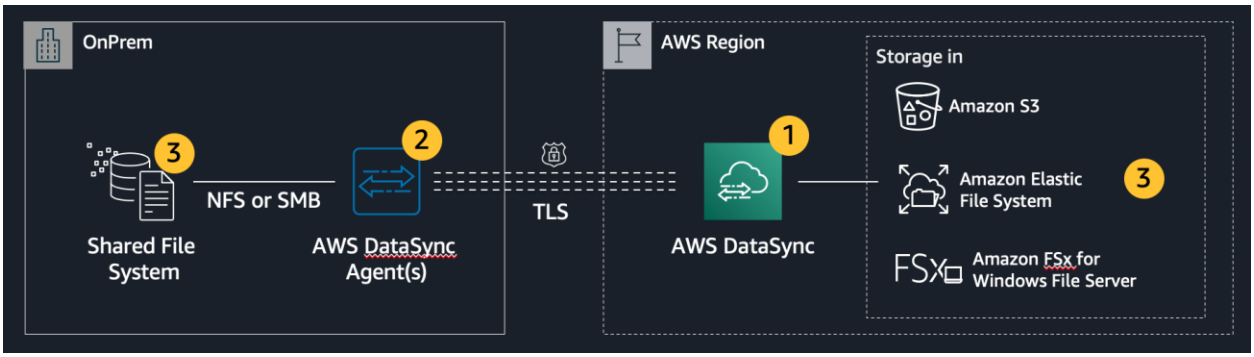

Figure 5 – Main DataSync components.

The main DataSync components are:

1. A DataSync service that orchestrates the data transfer between source and destination.
2. The DataSync agent is associated to customer's AWS account through the console or API. The agent is used to access the customer's NFS server, SMB file share, or self-managed object storage to read data from it or write data to it. Deploying an agent isn't required to transfer data between AWS storage services within the same AWS account.

   The customer deploys a DataSync agent to their on-premises hypervisor or in Amazon EC2. To copy data to or from an on-premises file server, a customer downloads the agent's virtual machine image from the console and deploys it to their on-premises VMware ESXi, Linux Kernel-based Virtual Machine (KVM), or Microsoft Hyper-V hypervisor. When a DataSync agent is used, the agent must be deployed so that it can access the customer's file server using the NFS, SMB protocol, or the Amazon S3 API. To set up transfers between their S3 on AWS Outposts buckets and S3 buckets in AWS Regions, customers deploy the agent on their Outpost. To set up transfers between a Snowcone device and AWS storage, customers use the DataSync agent AMI that comes pre-installed on the device.
3. The locations are the source and destination of customer data, these can be NFS and SMB file shares, Amazon FSx for Windows File Server, Amazon EFS file systems, or Amazon S3.

A task is a set for two locations (source and destination) and a set of options that customers use to control the behavior of a task. If customers don't specify options when they create a task, DataSync populates the options with service default settings. An execution task is an individual run of a task, which includes information such as start

time, end time, bytes written, and status. After a task starts, customers can monitor its progress, add or adjust bandwidth throttling for it, or cancel it before it completes.

Based on the previously explained components, when customers are configuring a task, it's possible to define as source location, for example, an S3 bucket, a network file system server on-premises as the target location. Customers can review Working with AWS DataSync Transfer Tasks for more information.

Document revisions

Date	Description
October 2023	Initial draft.
March 2024	First publication