Hybrid and Multi-Cloud Data Replication Scenarios and Solutions

# Scenario 1: Financial Data Analytics Across AWS and Azure

Context:

A global financial firm operates analytics workloads in Azure, but its core banking application remains on AWS. The organization wants to replicate transactional data from AWS RDS (PostgreSQL) to Azure Synapse for real-time dashboards and regulatory compliance reporting.

Solution:

* Data Type: Structured, sensitive (financial transactions)
* Frequency: Near real-time (every few minutes)
* Direction: One-way (AWS to Azure)
* Tools:
* AWS Database Migration Service (DMS) to stream changes to Amazon S3
* Azure Data Factory to extract from S3 and load into Azure Synapse
* Encryption using KMS (AWS) and SSE; TLS for data in transit
* Architecture: AWS RDS → AWS DMS → Amazon S3 → Azure Data Factory → Azure Synapse
* Security and Latency Considerations:
* Use VPC endpoints and IP allowlisting
* Optionally establish private connectivity via Direct Connect and ExpressRoute
* Monitor replication lags with CloudWatch and Azure Monitor

# Scenario 2: E-Commerce with Disaster Recovery (Azure to AWS)

Context:

An e-commerce platform hosts its primary infrastructure in Azure (Cosmos DB and Blob Storage) and wants to maintain a disaster recovery environment in AWS to ensure service continuity during outages.

Solution:

* Data Type: Structured (orders, inventory), unstructured (product images)
* Frequency: Real-time for structured, daily batch for unstructured
* Direction: One-way (Azure to AWS)
* Tools:
* Azure Event Grid to detect changes and trigger replication pipelines
* Azure Data Factory for transformation and batch sync
* AWS RDS and Amazon S3 as DR targets
* Architecture: Cosmos DB / Blob Storage → Event Grid or Data Factory → AWS RDS + S3
* Security and Latency Considerations:
* Combine Azure ExpressRoute and AWS Direct Connect for low-latency replication
* Implement TLS/IPsec for data in transit and AES-256 for data at rest
* Validate failover and disaster recovery process regularly

# Scenario 3: Healthcare Data from On-Prem to Google Cloud

Context:

A hospital network stores electronic health records (EHRs) and imaging data on-premises but wants to replicate this data to Google Cloud for long-term storage and machine learning workloads. Data must remain HIPAA-compliant.

Solution:

* Data Type: Unstructured (medical images), structured (EHRs); highly sensitive
* Frequency: Daily or weekly batch replication
* Direction: One-way (on-premises to cloud)
* Tools:
* Google Transfer Appliance for large initial transfers
* Google Storage Transfer Service for scheduled syncs
* Use BigQuery and Cloud Healthcare API for advanced use cases
* Architecture: On-prem File Server → Transfer Appliance / STS → Google Cloud Storage → BigQuery
* Security and Latency Considerations:
* Use VPC Service Controls and CMEK for encryption
* All data encrypted in transit and at rest
* Maintain audit trails using Cloud Logging and Cloud Audit Logs

# Scenario 4: SaaS Company Expanding with Bidirectional Replication

Context:

A SaaS provider runs their core services in AWS (us-east-1) and is expanding into Azure (UK South) to meet regional performance and data residency requirements for European customers. Customer data and preferences must be synced in both directions.

Solution:

* Data Type: Semi-structured JSON containing user profiles and preferences (with PII)
* Frequency: Real-time
* Direction: Bidirectional (AWS ↔ Azure)
* Tools:
* Confluent Cloud or Apache Kafka for stream replication
* Azure Event Hubs and AWS MSK as cloud-native ingestion points
* Schema Registry to validate structure and enforce versioning
* Architecture: Application → Kafka Producer → Cross-cloud Topic Sync → Consumers (AWS and Azure)
* Security and Latency Considerations:
* Field-level encryption for PII
* GDPR-compliant data handling with regional storage controls
* End-to-end observability with distributed tracing and audit logs