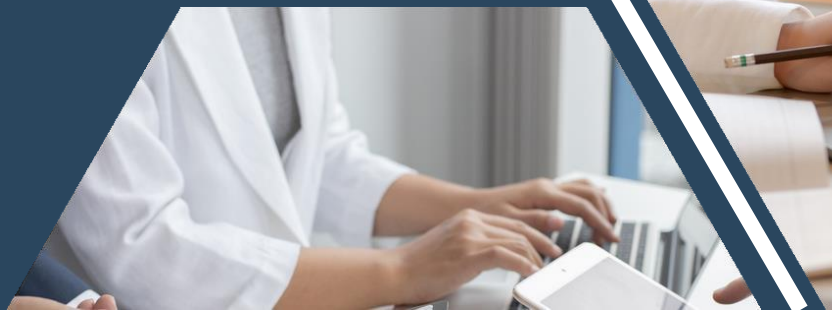
dt|digitaltrack®
If The Data Matters The Partner Matters

PRIMARYSTORAGE

Prepared by :
DigitalTrack Solutions Pvt Ltd

Prepared for :
Aarthi Scans and Labs

AARTHI
SCANS & LABS



CASE STUDY: ENHANCING BUSINESS CONTINUITY FOR AARTHI SCANS THROUGH AWS SOLUTIONS

OVERVIEW

- **Company:** Aarthi Scans
- **Industry:** Healthcare

ABOUT CUSTOMER:

Aarthi Scans and Labs offers its services to more than 7000 patients a day across India. We have 50 full-fledged diagnostic centers with CT, MRI, Ultrasound scan, Xray, Mammogram, OPG and Lab facilities. We have more than 150 collection centers providing lab services.

Executive Summary

Aarthi Scans faced challenges such as frequent disruptions to business continuity, difficulties in managing records, and limitations in scalability due to their reliance on on-premises infrastructure. DigitalTrack, an AWS partner, addressed these issues by migrating Aarthi Scans' systems to the AWS cloud. This included automating ETL processes with AWS Glue, integrating advanced analytics with Databricks, and securing data storage with Amazon S3 and KMS. The migration enabled improved business continuity, efficient record management, enhanced scalability, optimized ETL operations, and advanced data analytics capabilities. As a result, Aarthi Scans achieved greater operational efficiency, cost reduction, and better decision-making through robust data management and analytics solutions.

CUSTOMER BUSINESS CHALLENGES: Aarthi Scans faced significant challenges in maintaining business application continuity due to various factors including floods, security threats, record deterioration, storage space limitations, physical handling issues, lost documents, record retrieval difficulties, dispersed records, and power outages. These issues were compounded by their reliance on on-premises infrastructure, leading to global outages as all servers were hosted in a single physical location. To address these challenges, Aarthi Scans decided to migrate to the cloud .

CHALLENGES

1. **Business Application Continuity:** Frequent interruptions due to environmental and physical threats **.**

2. **Record Management** : Difficulty in storing, handling, and retrieving records.

3. **Scalability:** Inability to scale operations efficiently on a private cloud.

4. **Power Outages:** Dependence on a single physical location for servers.

5. **ETL Jobs Management** : Complicated ETL processes using multiple open- source applications on separate servers with large data volumes.

Why Choose AWS for Aarthi Scans?

AWS was chosen for Aarthi Scans due to its ability to ensure business continuity and disaster recovery, offering high availability and redundancy across multiple regions, thus reducing disruptions and downtime. The cloud's scalability and flexibility enable seamless scaling of resources to meet growing demands, while the pay-as-you-go model and tiered storage options, such as Amazon S3 and S3 Glacier, provide significant cost savings. AWS's advanced data analytics services like Glue and Athena facilitate efficient data processing and insights, and robust security features ensure compliance and data protection. With a simplified data management process and a global network infrastructure, AWS allows Aarthi Scans to efficiently manage their operations and focus on delivering quality healthcare services without the complexities of traditional IT infrastructure.

Why customer chosen DigitalTrack?

Digital Track is an AWS Advance Consulting Partner – Service Path and we had certified Solution Architect Professionals to support customer base account as well they had Dedicated Service Delivery Management & Project Management team.

SOLUTIONS PROPOSED: DigitalTrack, an AWS partner, proposed and implemented the following solutions to address this business.

Challenges:

ADDRESSING ETL CHALLENGES: Previously, Aarthi Scans used on-premises servers with multiple open-source applications for ETL (Extract, Transform, Load) processes, managing over 50 TB of data. DigitalTrack proposed the following AWS solutions:

1. **AWS Glue:** Automated the extraction, transformation, and loading of data across various sources, helping Aarthi Scans categorize and organize their data by field.

2. **Data Bricks:** Provided advanced data analytics capabilities to understand trends and make informed business decisions.

3. **AWS Glue Data Brew:** A visual data preparation tool that allowed data analysts to clean and normalize data without writing code.

INTEGRATING DATABRICKS WITH AWS GLUE:

1. **Databricks:** Unified analytics platform powered by Apache Spark, simplifying big data processing and enabling advanced analytics and machine learning.

2. **AWS Glue Crawlers:** Automated discovery and registration of schemas for datasets in the AWS Glue Data Catalog.

3. **AWS Glue Data Catalog:** A central repository to store and make structural metadata for all data assets searchable and query able.

4. **AWS Glue Spark Jobs** : Ran ETL operations to prepare data for analysis.

5. **Amazon Athena:** Enabled interactive querying of data in Amazon S3 using standard SQL, simplifying data analysis.

6. **AWS S3 Glacier:** Used for archival storage, managing and retrieving infrequently accessed data efficiently.

7. **Amazon S3 with KMS:** Provided secure storage and management of patient data.

DATA FLOW STEPS:

1. **Data Migration and Processing:**
    - **From Local Databases to AWS EC2:** Scripts running on local databases were migrated to AWS EC2.
    - **Script Execution and Notifications:** After executing scripts on AWS EC2, AWS Lambda functions, triggered by AWS Event Bridge, sent email alerts to notify stakeholders of the process status and any issues.
2. **Data Catalog and Transformation:**
    - **AWS Glue Crawlers:** Connect to data sources (PostgreSQL, MySQL, SQL Server) to catalog the data.
    - **Transformation:** Data is cataloged in AWS Glue Data Catalog and transformed by AWS Glue Spark Jobs. Transformed data is categorized into different layers in Amazon S3: Bronze (raw data), Silver (refined data), and Gold (aggregated data).
3. **Data Storage and Analysis:**
    - **Amazon S3 Storage:** Raw data is initially stored in the Bronze layer, refined data in the Silver layer, and aggregated data in the Gold layer.
    - **Amazon Athena:** Used for ad-hoc queries on data stored in Amazon S3.

APPLICATIONS DEPLOYED:

- **Patient Appointment Booking**
- **Patient Entry Management**
- **Patient Reports Management**
- **Field Team ERP**
- **Asset Management**
- **Billing Management**

OUTCOMES AND BENEFITS

1. IMPROVED BUSINESS CONTINUITY:

- **Outcome:** Ensured continuous operations despite environmental or physical threats.
- **Details:** Reduced Recovery Time Objective (RTO) and Recovery Point Objective (RPO) by implementing AWS cloud infrastructure, allowing for quick recovery and minimal data loss during disruptions.

2. EFFICIENT RECORD MANAGEMENT:

- **Outcome:** Improved record management and retrieval.
- **Details:** Automated AMI backups every 30 minutes with a retention policy of seven images, reducing downtime and ensuring minimal impact on operations.

3. ENHANCED SCALABILITY:

- **Outcome:** Dynamic scaling of operations.
- **Details:** Leveraged AWS cloud infrastructure to scale storage capacity seamlessly, accommodating growing data demands.

4. OPTIMIZED ETL PROCESSES:

- **Outcome:** Streamlined data transformation and loading.
- **Details:** Managed 2 TB of data efficiently with AWS Glue, resulting in faster data processing and reduced operational overhead.

5. DATA COMPRESSION AND COST REDUCTION:

- **Outcome:** Reduced data size and capital investment.
- **Details:** Compressed data to 20 TB using AWS S3 storage classes, achieving approximately 60% cost reduction while maintaining data accessibility.

6. ADVANCED DATA ANALYSIS:

- **Outcome:** Enhanced data analysis and visualization.

Details: Enabled detailed comparisons and graphical data representation, facilitating better decision- making and business insights.

7. IMPROVED STORAGE EFFICIENCY:

- **Outcome:** Enhanced storage practices.
- **Details:** Archived infrequently accessed data using S3 Glacier, reducing storage costs by approximately 70% compared to on-premises solutions.

8. DATA DEDUPLICATION AND VERSIONING:

- **Outcome:** Reduced data redundancy and managed data versions effectively.
- **Details:** Implemented S3 Versioning to keep multiple versions of data objects, allowing for recovery from accidental deletions or overwrites. AWS Glue's data deduplication features helped eliminate redundant data entries, ensuring clean and accurate datasets.

9. ENHANCED DATA SECURITY:

- **Outcome:** Improved data security.
- **Details:** Utilized AWS Key Management Service (KMS) for encrypting sensitive data, ensuring robust protection against unauthorized access.

CONCLUSION: By migrating to AWS and leveraging services such as S3, Glue, and Data Lakes, Aarthi Scans significantly improved their data management and streamlined operations. This transition not only reduced capital investment but also provided robust solutions to scalability and data analysis needs. AWS Glue's automatic ETL code generation and automated data schema recognition simplified data processing tasks, enabling efficient data cleaning, deduplication, and versioning. The cloud-based approach facilitated better business continuity, enhanced record management, and allowed Aarthi Scans to achieve advanced data analytics capabilities, leading to more informed decision-making and operational efficiency.

ARCHITECTURE DIAGRAM: