Application Migration Assessment Discovery and Design

Application Name

CCM PAC & RAD Data Mart

(UHGWM110-028731)

Table of Contents

[1 Overview​ 2](#_Toc1466224097)

[2 Application Profile 3](#_Toc596412984)

[3 Methodology 4](#_Toc1287891065)

[4 Architecture and Platform (Current & Future) 4](#_Toc694986879)

[4.1 Findings and Risks 6](#_Toc490138357)

[4.2 Recommendations 6](#_Toc75927899)

[5 Security and Compliance 6](#_Toc923519069)

[5.1 Findings and Risks 6](#_Toc677284984)

[5.2 Recommendations 7](#_Toc100618789)

[6 Data Management 7](#_Toc14569427)

[6.1 Findings and Risks 8](#_Toc868565594)

[6.2 Recommendations 8](#_Toc1459319223)

[7 TCO Assessment & Management 8](#_Toc1877921373)

[7.1 Findings and Risks 9](#_Toc703893897)

[7.2 Recommendations 9](#_Toc62498911)

[8 Business Continuity & Disaster Recovery 9](#_Toc626541998)

[8.1 Findings and Risks 9](#_Toc1107532167)

[8.2 Recommendations 9](#_Toc1934806505)

# Overview​

The purpose of the Application Migration Assessment is to evaluate the readiness and completeness of the migration assessment and design phase documentation in preparation for migration planning.

Key findings and recommendations​:

1. There are three MS SQL Servers running in ELR datacenter
2. 1 Dev/Test, 1 UAT and 1 Production Server
3. There are additional components used – Enterprise Airflow, and Optum Functions Platform which are not part of this migration scope
4. Once the databases are migrated to Azure, the database instances still need connectivity to the Enterprise Airflow and Optum Function environments
5. Data is ingested using ECG from multiple data sources – Pathway PODs & Datamart (SQL Server Databases), ICUE OHBI DWaaS (Snowflake)
6. Enterprise Airflow DAG jobs trigger the data ingestion & data processing functions running in Optum Functions Platform
7. Functions are written in Python, containerized as docker images. The images and other artifacts are stored in jfrog.
8. Only the 3 database servers are in scope for migration, other processes remain unchanged
9. There are plans to migrate to Azure App Services and Azure Data Factory in the future
10. Full backups are taken once a week
11. There is no requirement for HA or DR, as the data can be re-processed if required
12. PHI & PII data are stored in the databases.
13. Processed data is sent to Pathway Common Enhancements SQL databases and File shares for downstream consumption and processing.
14. ECG can be continued to be used for file transfers.

# Application Profile

* CCM PAC & RAD Datamart application support Member Eligibility and Risk Adjustment Data. This is a data repository and ETL process of member eligibility and risk adjustment data for clinical programs that are managed on Pathway Application.
* **Points of contact**: Manoj Raot, Janga Reddy Metai
* **LOB**: Optum Health & Financial Tech
* **App R Disposition:** Re-Platform

# Methodology

The following methods were used as part of this assessment.

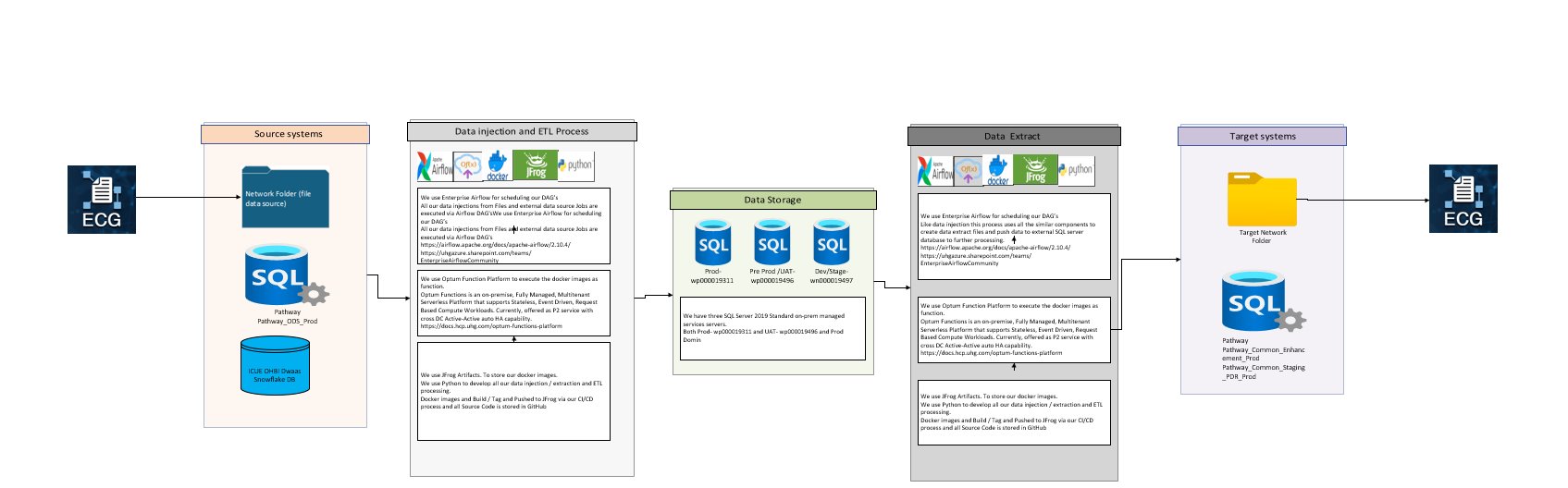
**Discovery sessions:**

Discovery sessions were held using Microsoft Teams Meeting and Chat to review existing documentation for the FACETS (Optum Care) Application.  Items reviewed during Discovery are as follows:

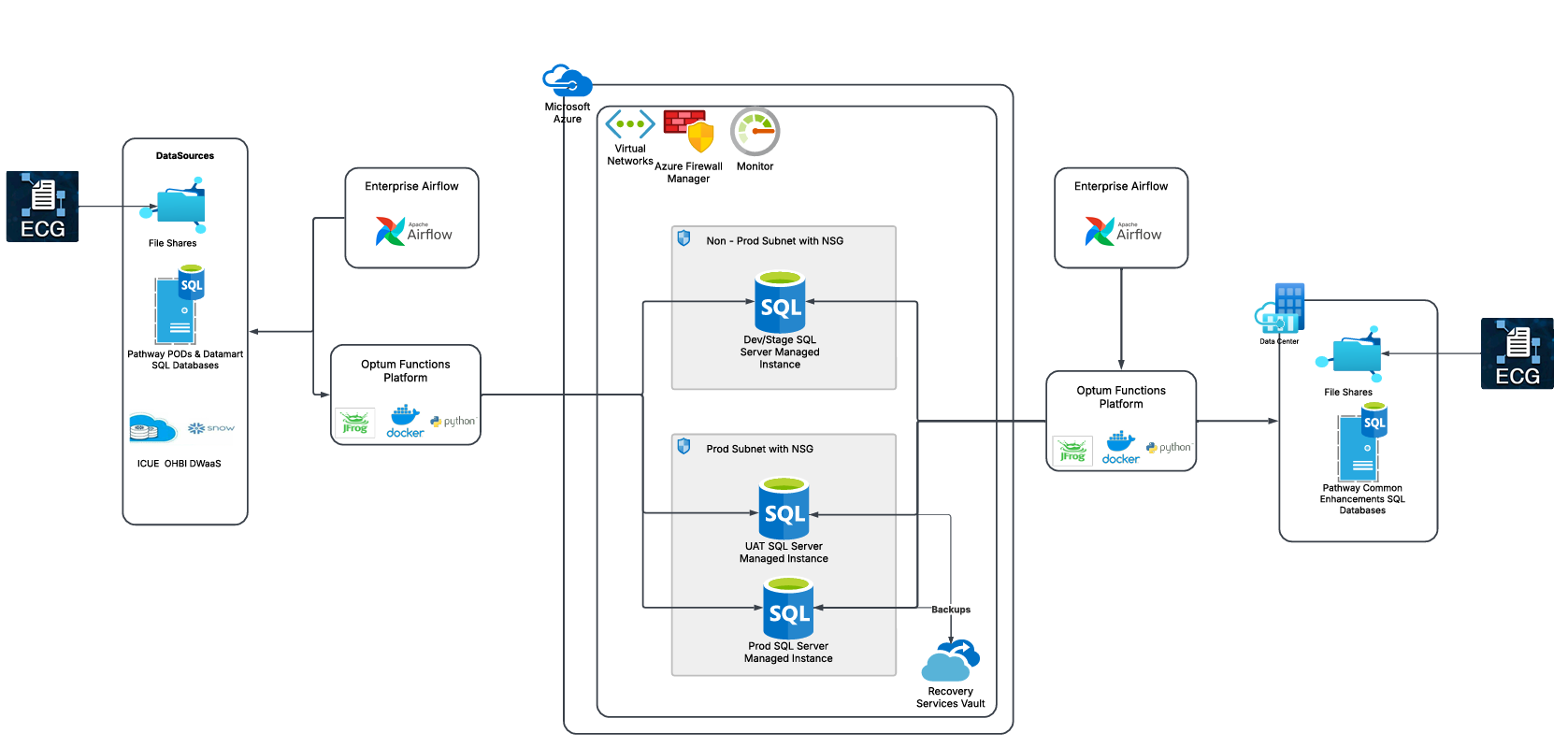
1. Current State Architecture
2. Sizing of Servers
3. Backup Requirement
4. Disaster recovery and BCP requirements
5. Application requirements
6. Database Migration Requirements
7. ETL & Batch Processing requirements
8. Security & compliance requirement
9. Monitoring and Observability
10. Cost Management

# Architecture and Platform (Current & Future)

Current State Architecture:



Future State Architecture:



CCM PAC & RAD Datamart application receives data from Pathway SQL Server databases and ICUE OHBI DWaaS Snowflake tables, after data processing and transformation the data is sent to Pathway Common Enhancements databases for downstream processing.

Tech Stack:

* Operating System – Windows Server 2019
* Programming Languages – Python
* Containerization – Docker Containers
* DevOps Tools - JFrog
* Database – MS SQL Server 2019
* Access – Active Directory (MSID)

## Findings and Risks

* There are only three database servers to be migrated
* Overall data volume is around 2 TBs
* UAT and Prod Server instances need to stay close to each other
* Access is managed by Active Directory (MSID)

Risks:

* Need connectivity back to Enterprise Airflow and Optum functions platform the data processing jobs to be triggered and processed.
* There is no HA or DR at this point, full backups are taken once a week.
* These are the software that are pending approval in Tech PADU
* SQL Server Management Studio

## Recommendations

* Leverage SQL Server managed instances in Azure, over running SQL Server databases in Azure VMs
* UAT and Prod servers need to stay close, either both the servers can be in the same subnet and network security group or allow proper network security group rules for them to communicate seamlessly.
* Choose the right database migration method from the available options – Azure data box, Azure Data migration Services, Azure Data Factory
* Ensure data security while in transit and at rest, encrypt data at rest using TDE, TLS1.2

# Security and Compliance

The application handles PII & PHI data, application does need to meet HIPPA and PII compliance.

## Findings and Risks

* Data is classified as sensitive data with PII & PHI data
* Network segregation, Firewall rules are in place for network layer security.
* This is an internally facing application but provides/exchange data with Pathway databases
* All hosts are under HCP, and are scanned for vulnerabilities
* Airflow and Optum functions platform are also managed services, scanned for vulnerabilities and remediated

Risks:

* There are no additional, new risks identified.

## Recommendations

* Leverage Azure provided security features such as Azure Firewall,  Azure KMS
* Maintain same level or improved network security controls, use VNETs, Subnets and Network Security Groups.
* Leverage privileged access control policies for data access
* Collect and aggregate logs using Splunk or Azure Sentinel
* Continue Application level static code analysis in Github/Jfrog repositories.
* Ensure data is encrypted both in transit and rest and follow this policy defined by EIS

# Data Management

## Findings and Risks

Data Types:

* PHI: True
* PCI: False
* PII: True
* Company Confidential:  False

Backup and Recovery:

* RTO: 8 hours
* RPO: 8 hours (Current)
* RTO: 8 hours (Target)
* RPO: 8 hours (Target)
* The target state RPO/RTO objectives are based on the proposed future state architecture
* Weekly full backups are taken
* All back-ups of SQL Server are performed by the Enterprise SQL Server Team.
* Data retention is for 3 years

## Recommendations

* In cloud, configure Database level backups and store backups in Recovery Service Vaults.
* Long term backups can be moved to Cold Storage for cost savings
* Configure two Recovery Service Vaults with GRS, store backups in GRS enabled Vault for data that requires DR/Cross region replication.
* Configure backup and data life cycle policies as required

# TCO Assessment & Management

## Findings and Risks

* The current on-premises TCO is **$78,000** /Year
* TCO for Azure Cloud Hosting with Azure Hybrid Benefits + 7% Managed services Charge back is **$56,894** /Year
* More savings are possible with a 3-year term commitment

## Recommendations

* Leverage Azure Hybrid Benefits for SQL Server and Windows licenses as applicable.
* Scale up and scale down Managed Instances as required in non-prod environment.

.

# Business Continuity & Disaster Recovery

## Findings and Risks

* There are no strict RPO/RTO requirements, the target is 8 hours.
* No Active DR is required, backup are recommended to be replicated to the DR region
* Cross region replication will be enabled for critical databases
* No Active HA required, databases can be restored or processing jobs can be rerun when required.

## Recommendations

* Enable DB level backups and cross-region replication for critical resources as required
* Ensure Azure Site Recovery is configured for critical databases
* Automate the Azure Site Recovery to replicate backups to the DR region.