

YES PREP DATA MIGRATION RFQ

SCOPE OF WORK

Please Note: Proposals, one (1) original and one (1) digital copies (USB or external hard drive), will be received by the District at the address shown below until **2:00PM Thursday July 29, 2021**. in time for delivery before the closing date or can emailed ricardo.rivera@yesprep.org

Attn: Rivera, Ricardo, Director of Business Intelligence
YES Prep Public Schools
5515 South Loop E FWY STE B,
Houston, TX, 77033
Ph: +1 (713) 967-8467
Email: ricardo.rivera@yesprep.org

CONTENTS

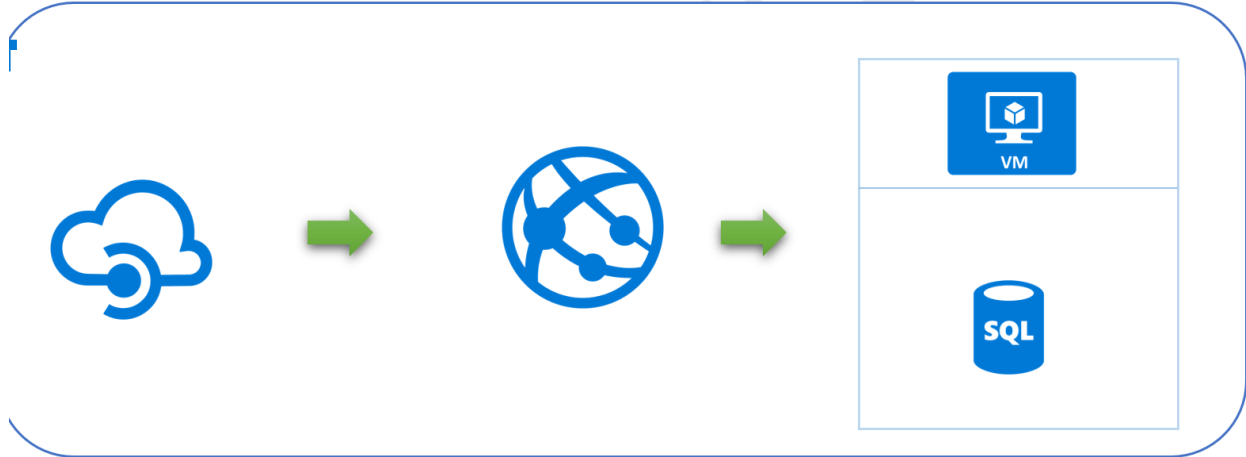
Executive Summary	2
CUrrent Information Architecture	2
Pre-Migration Cloud Infrastructure	2
POST-Migration Cloud Infrastructure	2
Timeline	4
High Level Objective	4
Risks and Assumptions	4
Project Execution Strategy	5
Phase 1 Planning	5
Pre-Migration Planning	5
Project Initiation	6
Landscape Analysis	6
Phase 2	6
Solution Design	6
Phase 3	7
Implementation	7
Phase 4	7
Decommission & Monitor	7
Cloud Infrastructure	8
Development Infrastructure	8
Data Security	8
Passive Data Security	9
Active Data Security	9

EXECUTIVE SUMMARY

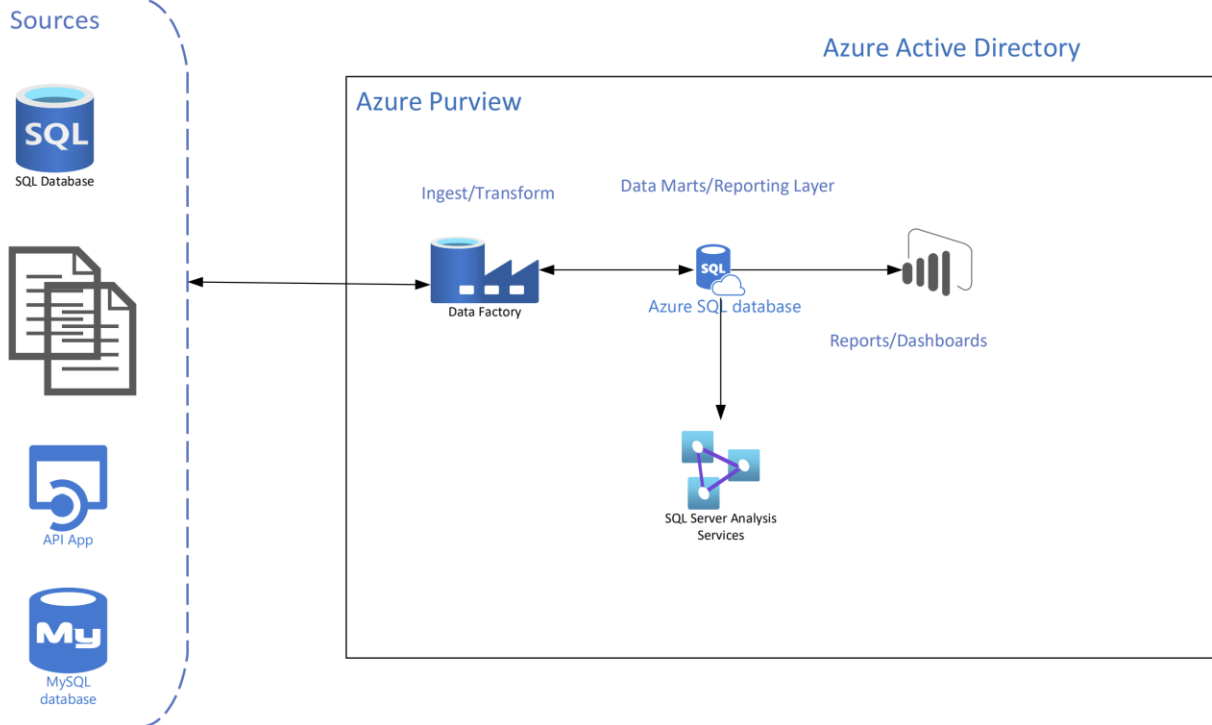
This document provides an outline for vendors who will be directly or indirectly involved with submitting an RFQ. These Migration efforts are prompted by the need for upgrades in technical infrastructure. All vendors must provide all information requested including the profiles of the people involved in the project, a technical approach, security requirements, operability, and accessibility. In addition, all vendors must include an IT portfolio, Education Experience, Cloud Implementation Expertise, an implementation plan, and references.

CURRENT INFORMATION ARCHITECTURE

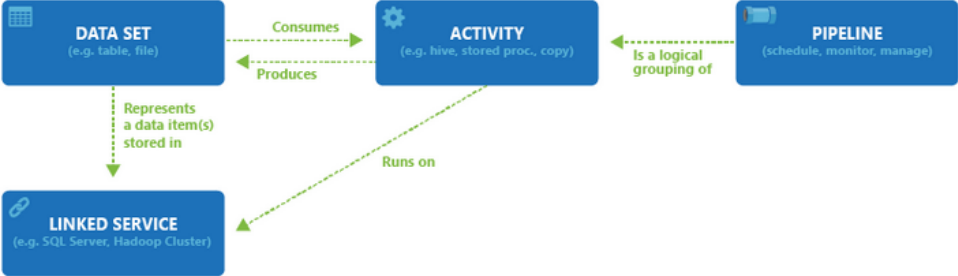
PRE-MIGRATION CLOUD INFRASTRUCTURE



POST-MIGRATION CLOUD INFRASTRUCTURE



The following diagram shows the relationships among pipeline, activity, dataset, and linked service in Data Factory:



TIMELINE

1. **Staging** environment & test migration run (July – Sept.)
 - a. Vendor Scope: (65%) (New Partner, Catapult, MS premier, Azure design)
 - i. Design Architecture
 - ii. Build environment
 - iii. Move to Azure SQL Managed Instance
 - iv. Hours for Post Support
 - b. YES Scope (35%)
 - i. Work with partner for due diligence prep
 - ii. Sign off
 - iii. Costs of 2 Azure SQL Managed Instance to run
 1. Add object storage for copy/dump
 - iv.
2. **QA Test Sign off**
 - a. Multiple sign off (Nov)
3. Production Data Refresh
 - a. Connections move
 - b. **Go Live** (Dec)
4. Decom Common DB

HIGH LEVEL OBJECTIVE

Before YES Prep data is moved, it will require that the current environment is staged in SQL Azure. This will allow us to implement the necessary changes that will migrate our current SSIS views into Azure Data Factory. For example, our current system moves data into Common using one method, the target system will use a different method to move the data to and from Common.

1. Deploy the SSIS packages from Azure VM to ADF.
2. Source and target connections of SSIS packages should be modified to point to Azure SQL DBs.
3. An additional SQL instance is required to deploy the SSIS packages or if we don't want to create a new instance, we can use the SQL instance hosting the Azure VM's SQL DB.
4. Configure Views or Ge-replication on Azure SQL DB to have exact copy of data for reporting purposes (Power BI).
5. Decommission Legacy SQL Server
6. Decommission legacy VM

Each data migration has obstacles: specifically, risks, constraints, challenges, dependencies, and assumptions. While planning the stages of the migration, we have to identify as many of the obstacles as possible.

RISKS AND ASSUMPTIONS

Assumptions	Constraints	Risk
Sufficient resources are available for all aspects of data migration.	Staff Schedules dictates what must be completed and when.	Unexpected delays occur.

Expertise is available for all aspects of migration.	Budget might limit access to resources that can be dedicated to the project.	Some SSIS packages may not be able to migrate to ADF
All environments are available during necessary steps of migration.	Personnel might be limited or unavailable.	Power BI may not see the new server

PROJECT EXECUTION STRATEGY

The following section describe the desired execution strategy for the Data Migration Project. The high-level processes recommended for each phase of the Data Migration Project Lifecycle. While many migration projects follow these phases, the processes may vary depending on the size, scope and complexity of each ETL that YES Prep will need to migrate. As a result, this information will serve as a guideline for developing, evaluating, and implementing our migration efforts.

Phase 1: Planning	Phase 2: Analysis	Phase 3: Implementation	Phase 4: Decommission and Monitor
BRD Requirements			
Assess Current Environment	Design Data Environment	Cleanse Data	Knowledge Transfer
Develop Data Migration Plan	Design Migration Procedures	Convert Transform Data (as needed)	Migration Results
Define and Assign Team Roles and Responsibilities	Validate Data Quality	Migrate Data (test then late deployment)	
Pre-Migration Planning		Review pre and post results	

PHASE 1 PLANNING

The result of the planning phase needs to be:

- A dedicated PM Plan
- A Migration Plan as part of the PM Plan
- A BRD

PRE-MIGRATION PLANNING

In this phase the following tasks will be accomplished:

1. Assessment of current environment
2. Complete project estimates
3. Communicate team progress
4. Verify Cyber Security Charter status
5. Identify project resources

6. Create structured work flow
7. Outline configuration management, data migration and governance policies

The assessment of the current environment requires the compilation of all identified work products to create a blueprint of the current architecture. Documenting and analyzing the current environment from a functional and technical perspective is important for a full understanding of the data given that there is little to no documentation in YES Prep, and the staff cannot address data governance at the technical level given the lack of resources.

The following must be gathered:

- Logical and physical data
- Database definition for existing relational databases
- Data dictionaries documenting each data element (labels and definitions as well as properties)
- Relevant business rules and processes in the current environment and the target system
- Data mapping from source system to data warehouse
- Names of systems interfacing with the source systems indicating whether the application sends data to or extracts data from the source system (for context only). This information will help gauge the time constraints for potential downtime of the source system. If there are many interfacing systems, the coordination task will be more complex than with fewer interfaces
- Data profiling analysis for source system (if available)

PROJECT INITIATION

In this phase the following tasks will be accomplished:

1. Define communication plan
2. Define project policies
3. Confirm third party agreements
4. Define cloud and software requirements

LANDSCAPE ANALYSIS

In this phase the following tasks will be accomplished:

5. Create data dictionary
6. Create risk management process
7. Compile data quality and management report
8. Perform security impact assessment

PHASE 2

SOLUTION DESIGN

In this phase we will seek to develop the core functionality that will be used to produce the desired results of the plan. The following are the steps involved in this phase:

1. Create detailed data map specifications

This task is one of the most important as it is necessary to ensure that both the original and destination datasets match correctly and completely. In this activity we will identify all the data fields,

structures and components that exist in the original database and create the specifications for the matching structure in the destination.

2. Create interface design specifications

The creation of interface design specifications will be minimal in this project as the majority the functionality will be based on scripts or existing tools.

3. Create data quality management specifications

This task is focused on the processes that will be utilized guarantee the accurate and efficient migration of the original data to its desired destination.

PHASE 3

IMPLEMENTATION

- Document migration logic
 - a. This task is dedicated to the documentation of the processes used in migration process and includes user manuals & code documentation where possible.
- Define environment monitoring/reporting methods
 - a. In this activity we create the infrastructure and tools necessary to trace and verify the migration process as we as the decommissioning activities. This can the form of logs and console output. A final report on the summary of the all processes will also be compiled.
- Define migration fallback policy
 - a. Since we actually be creating copies of the original data, the default fallback option will be to continue using the original data. Sample datasets may also be needed to ensure that migration and transformation functions perform as desired.
- Execute migration strategy
 - a. This is by all means the most important task. Once all the checks and balances have been passed, execution of the strategy can begin. This step involves running several previously tested scripts and saved configuration files that will initiate the migration process. During this time, it may be necessary to take the primary data offline to ensure that no further changes can be made to the original dataset once the migration process has begun. This ensures a close to 100% data integrity index. Monitoring, logging and reporting tools will also be in place to detect errors during migration.

PHASE 4

DECOMMISSION & MONITOR

This is the final phase of the migration process and involves the following tasks.

1. Complete system retirement

The following steps are involved in completely retiring the original dataset.

- a. Making a full backup of the database(s) with includes both data and structure including database intrinsic components like views, stored procedures, functions, triggers, collations, indices, users, etc.
- b. Redirect all applications and endpoint to the newly migrated dataset

- c. Save the original dataset files in cold storage
2. Validate system retirement

Once the migration process has been completed, the new dataset must be validated as functional before project itself can be marked as a success. This involves a time consuming but necessary task of verifying that the new dataset works seamlessly with all it connected local and external applications. This requires thorough test of these application for subtle errors like date-time formatting, currency and numeric formatting, time zone issues, integer and decimal rounding etc. This process can be accelerated using unit testing, but in general may sometimes take weeks to fully discover and/or verify.

CLOUD INFRASTRUCTURE

The primary IT infrastructure for this project is currently hosted on Microsoft Azure. The following are a list of the cloud services are or will be utilized in the development of this statement of work:

Cloud Infrastructure Services	
Azure Virtual Machines	This service currently used to host the Data Visualization web application.
Azure Functions (if needed)	This service acts as the primary communication hub for all activity related to the project. It allows secure access the business logic of the platform and is easily documented and extendible.
Azure App Services (if needed)	This service provides a central platform to create and store business logic.
Azure WebJobs(if needed)	This service is used to perform long running tasks that would be too long to perform while using a web browser.
ADF	The Azure Data Factory (ADF) is a service designed to allow developers to integrate disparate data sources. ... It provides access to on-premises data in SQL Server and cloud data in Azure Storage (Blob and Tables) and Azure SQL Database
SQL Azure	Azure SQL Database is an intelligent, scalable, relational database service

DEVELOPMENT INFRASTRUCTURE

The following development tools/languages will be utilized in this project:

- C#
- SQL

DATA SECURITY

Each data migration project needs to operate within the security controls in place at YES Prep. As a result, we must ensure the appropriate processes determined by the CMG are followed to protect all data at each source and

during the migration of the data between sources. This guide also addresses the following operational security controls of:

- Production controls
- Contingency planning
- Software maintenance controls
- Incident reporting
- Data encryption

In addition, the Data security will be implemented in a two (2) methods.

PASSIVE DATA SECURITY

Security Best-practices as defined by BSIMM

ACTIVE DATA SECURITY

- Microsoft Azure Sentinel
- Azure Security

End of Scope of Work