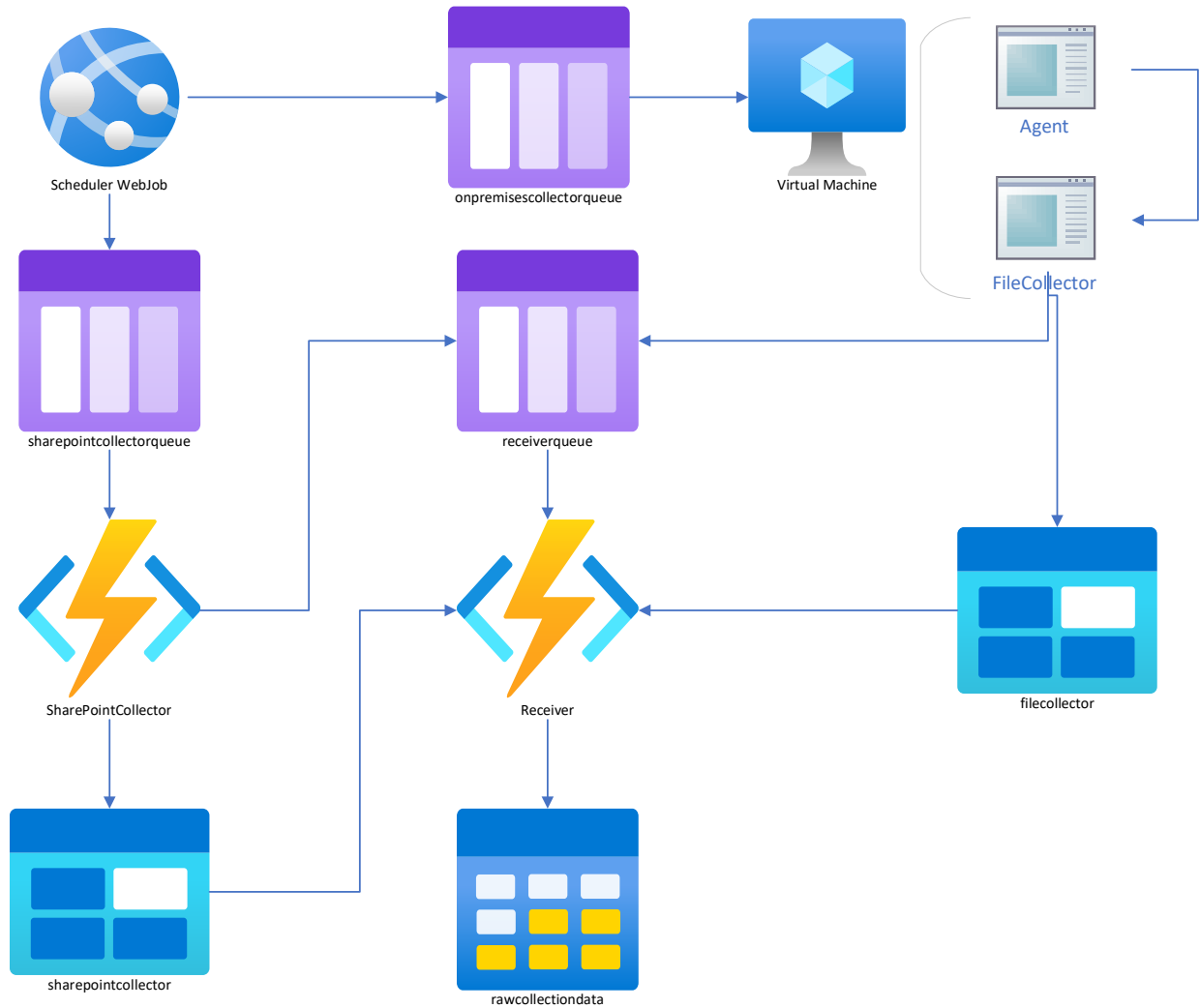# Sphere Messaging POC Architecture

Overview

This document describes the functionality included in the Sphere Messaging POC.

Diagram

Components

Scheduler

- The Scheduler is implemented as an Azure WebJob hosted on an Azure App Service. It can run continuously or via a cron schedule. When executed it places a message indicating that a collection process needs to execute on the onpremisescollectorqueue Azure Storage Queue and the sharepointcollectorqueue Azure Storage Queue.

Sphere Messaging POC Architecture

Agent

- The Agent is implemented as a console application hosted on a virtual machine, but could also be deployed as a Windows service. Its purpose it to monitor the onpremisescollectorqueue Azure Storage Queue for new work and invoke the on premises collector process specified in the message.

FileCollector

- The FileCollector is a sample on premises data collector implemented as a console application. It is invoked by the Agent when requested and gathers directory and file information from the specified path. Collected data is uploaded to the Azure Blob Storage container named filecollector in the form of a JSON file. The FileCollector places a message indicating there is a file to be processed on the receiverqueue when it completes uploading the file to Azure Blob Storage.

SharePointCollector

- The SharePointCollector is implemented as an Azure Function. It is automatically triggered when a new message arrives on the sharepointcollectorqueue Azure Storage Queue. An application registration has been created in Azure AD that allows the SharePointCollector read access to an organizations SharePoint site collection. The SharePointCollector gathers folder and file information from the specified document library. Collected data is uploaded to the Azure Blob Storage container named sharepointcollector in the form of a JSON file. The SharePointCollector places a message indicating there is a file to be processed on the receiverqueue when it completes uploading the file to Azure Blob Storage.

Receiver

- The Receiver is implemented as an Azure Function. It is automatically triggered when a new message arrives on the receiverqueue Azure Storage Queue. The Receiver downloads the JSON file specified in the message from Azure Blob Storage, parses the information, and loads it into a No-SQL table named rawcollectiondata in Azure Table Storage.