

Centers for Medicare and Medicaid Services (CMS)

## Beneficiary Information on the Cloud (BIC)

Publishing Beneficiary Information to a CMS Enterprise  
Data Sharing Platform

Approach Paper and  
Proof of Concept (POC)

Draft Version 0.1

February 12, 2020



# Beneficiary Information on the Cloud (BIC)

Publishing Beneficiary Information to a CMS Enterprise Data Sharing Platform

## Table of Contents

1	Introduction.....	3
2	Overview of BIC .....	3
3	CMS Enterprise Data Lake .....	3
4	Approach/Assumptions .....	4
5	Organization of Data in S3 .....	5
6	Data Publishing Architecture.....	6
7	Beneficiary Information Available.....	7

## List of Figures

Figure 1: BIC Data Propagation Architecture.....	6
--	---

## List of Tables

Table 1: Beneficiary Data in CMS Enterprise Data Lake .....	7
Table 2: Record of Changes .....	14

# Beneficiary Information on the Cloud (BIC)

Publishing Beneficiary Information to a CMS Enterprise Data Sharing Platform

## 1 Introduction

This document outlines an approach for publishing beneficiary demographics, entitlement, enrollment, premium, health status and other coverage information to a CMS Enterprise Data Sharing Platform (Data Lake) that could eventually grow to a comprehensive source of Medicare Beneficiary, Provider, Claims and Quality information. The approach described here is based on a Proof of Concept (POC) completed with full size de-identified beneficiary data from the Beneficiary Information on the Cloud (BIC) Data Services and Application Programming Interface (API) platform.

## 2 Overview of BIC

Beneficiary Information on the Cloud (BIC) is a strategic CMS initiative to minimize and eventually eliminate redundancy, latency and inconsistency in beneficiary information used by systems and business partners supporting the Medicare Program. This was the first step towards modernizing and migrating to the cloud of the Medicare Enrollment and Premium Billing Systems (MEPBS).

BIC utilizes the Amazon Web Services (AWS) cloud to provide a highly available and scalable source of Medicare Beneficiary Information for Medicare systems and business partners. Changes to beneficiary information during MEPBS transaction processing on the CMS IBM Mainframe are propagated to the cloud based data repository in real time thereby eliminating latency. BIC is accessible to authorized partners via an Application Programming Interface (API) for real time queries and an Amazon RedShift based data warehouse for bulk consumption.

BIC has been in production since April 2019 and has several consumers in production and integration testing phases. The BIC API is available via AWS PrivateLink for consumers on the AWS platform and Amazon API Gateway for other consumers. The beneficiary information in the Amazon RedShift data warehouse is accessible via Virtual Private Cloud (VPC) Peering.

## 3 CMS Enterprise Data Lake

A logical extension to BIC is making beneficiary information available in a CMS Enterprise Data Lake platform. The advantages of a Data Lake would be easy access to valuable information for authorized entities and individuals while also minimizing administrative burden. Consumers will access the Data Lake under a “Bring your own compute” paradigm and help keep operational costs low for source systems.

The CMS Enterprise Data Lake platform could co-host beneficiary data with other complementary Medicare Information such as Providers, Claims and Quality to provide a comprehensive source of Medicare Information for research, analysis and Extract Transform and Load (ETL) by authorized partners. The specific types of Medicare Information will be published and maintained by the respective systems of record. CMS will need to establish an

# Beneficiary Information on the Cloud (BIC)

## Publishing Beneficiary Information to a CMS Enterprise Data Sharing Platform

enterprise data governance process to manage the data lake and control access to authorized consumers.

### 4 Approach/Assumptions

- Data Storage
  - The Data Lake will be implemented with Amazon S3 for storage and Amazon Athena as the Query Engine/Interface.
  - Types of beneficiary data that would be available in the Data Lake are listed in Section 7: Beneficiary Information Available. Other types of beneficiary data may be added in the future based on business needs and approval from respective business owners.
  - Data will be stored in encrypted S3 buckets that are not publicly accessible
  - Data in S3 will be stored in the Apache Parquet format.
  - The cost of S3 storage for beneficiary data will be incurred by BIC. Similarly, other systems publishing data into the Data Lake will pay for storage of their respective data.
  - S3 buckets will be setup as 'Requester Pays'. Data access costs will be paid by consumers.
  - Beneficiary data will be partitioned in S3 using the Beneficiary Partition Key generated at the time of accretion into MEPBS. The Beneficiary Partition Key is an integer number ranging from 0 to 999 and allows beneficiary data to be distributed across 1000 partitions.
  - Additional levels of partitioning (for example a computed sub partition key) may be considered after performance analysis
  - Initially, beneficiary data will be refreshed on a weekly basis. The refresh frequency could be adjusted (to more or less frequent) based on evolving operational and consumer requirements.
- Potential Use Cases:
  - CMS Enterprise level analytics
  - Reference data for CMS systems
  - ETL source for CMS systems
  - Source of beneficiary entitlement/enrollment information for a secure environment serving authorized internal/external researchers and analysts
- Security/Governance
  - Beneficiary tables will be organized under data groups based on business process and sensitivity of information.
  - Access to tables in S3 can be controlled at the data group level or as individual tables.

# Beneficiary Information on the Cloud (BIC)

## Publishing Beneficiary Information to a CMS Enterprise Data Sharing Platform

- Access to data will be restricted using policies on S3 buckets and folders containing data groups and tables.
- CMS will define enterprise level data governance activities to administer/maintain the Data Lake.

### 5 Organization of Data in S3

Beneficiary Data in S3 will be organized under data groups, tables and beneficiary partitions as shown below:

- Data Group 1
  - Table 1
    - Beneficiary Partition 0
    - Beneficiary Partition 1
    - ...
    - Beneficiary Partition 999
  - Table 2
    - Beneficiary Partition 0
    - Beneficiary Partition 1
    - ...
    - Beneficiary Partition 999
- Data Group 2
  - Table 3
    - Beneficiary Partition 0
    - Beneficiary Partition 1
    - ...
    - Beneficiary Partition 999
  - Table 4
    - Beneficiary Partition 0
    - Beneficiary Partition 1
    - ...
    - Beneficiary Partition 999
- Data Group 3
- ...
- Data Group *n*

# Beneficiary Information on the Cloud (BIC)

Publishing Beneficiary Information to a CMS Enterprise Data Sharing Platform

## 6 Data Publishing Architecture

Beneficiary Information in the CMS Enterprise Data Lake will be updated from the BIC Beneficiary Data Warehouse on a weekly basis. The update process will be built on the existing BIC data propagation architecture that replicates beneficiary data from the operational data store on an IBM Mainframe in the CMS Baltimore Data Center to data stores on AWS. The data stores on AWS include Amazon DynamoDb for real time APIs and Amazon Redshift for the data warehouse. Tables in Amazon RedShift will be unloaded in Parquet format to S3 utilizing the partitioning scheme detailed in Section 4: Approach/Assumptions.

**Figure 1: BIC Data Propagation Architecture** illustrates the propagation of beneficiary data from the IBM Mainframe to AWS. The steps populating the Data Lake are illustrated with a red box.

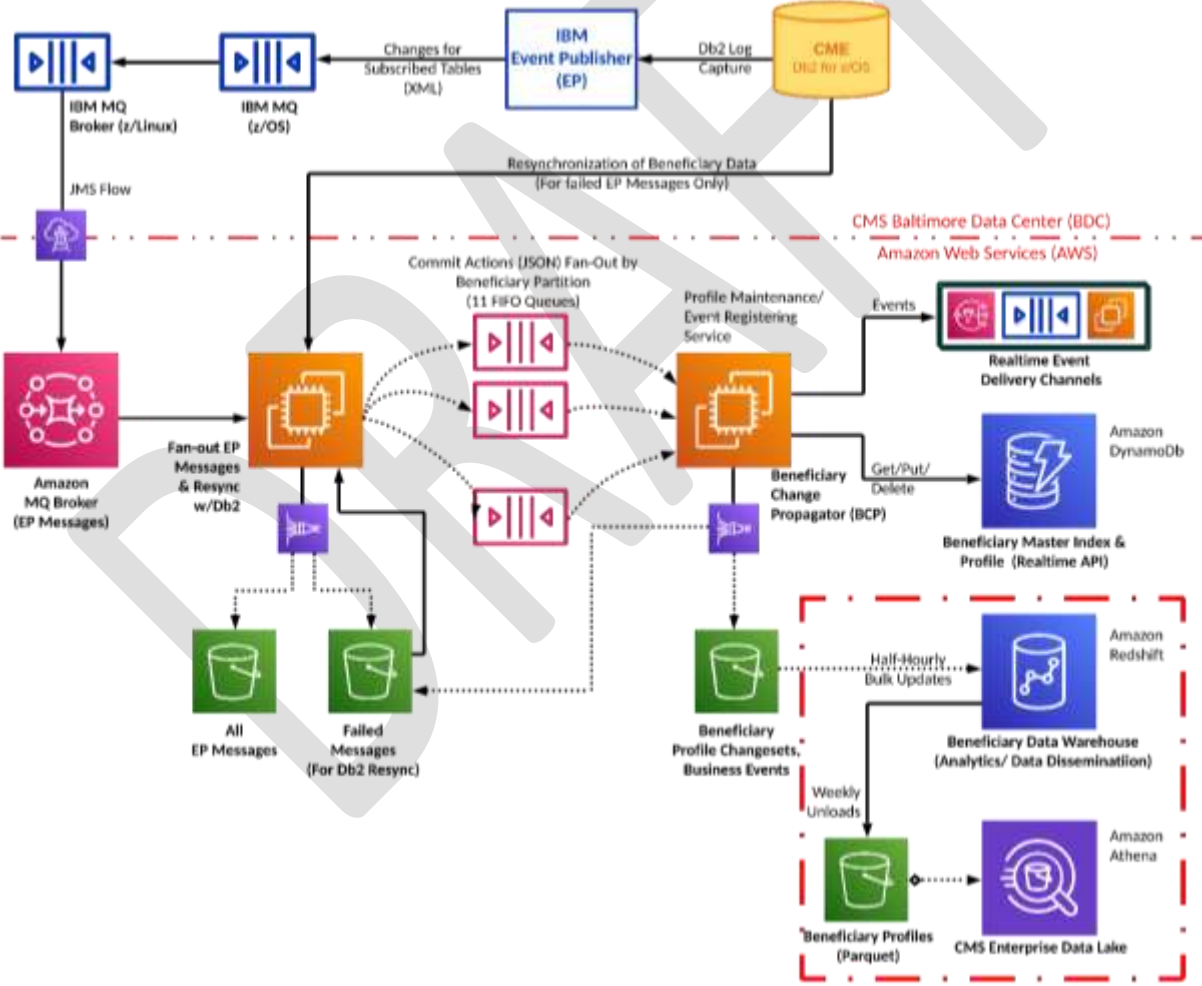


Figure 1: BIC Data Propagation Architecture

# Beneficiary Information on the Cloud (BIC)

Publishing Beneficiary Information to a CMS Enterprise Data Sharing Platform

## 7 Beneficiary Information Available

Beneficiary data that would be published to the CMS Enterprise Data Lake are listed in **Table 1: Beneficiary Data in CMS Enterprise Data Lake**.

The complete BIC data dictionary (including Attribute level details) can be accessed by CMS Enterprise User Administration (EUA) users at <https://developer.bic.cms.gov>.

*Table 1: Beneficiary Data in CMS Enterprise Data Lake*

Table Name	Description
beneficiary	Beneficiary's base information.
beneficiary_cross_reference	Beneficiary's cross-reference Health Insurance Claim Numbers (HICNs) information.
beneficiary_demographic_history	Beneficiary's gender, date of birth, and date of death information. The table was created with the implementation of MARx R&M in April 2011; the initial load of data was the current values of the beneficiary's' demographic data (BENE_BRTH_DT, BENE_DTH_DT, and SEX_CD).
beneficiary_drug_profile2	Beneficiary's principal or other International Classification of Diseases (ICD) codes.
beneficiary_dual_eligible	Medicaid dual eligible beneficiaries that are submitted in the Medicare Modernization Act (MMA) state file and matched to the beneficiary in the Common Medicare Environment (CME).
beneficiary_election_period_usage	Beneficiary's election period usages including the type of election period and corresponding usage dates.

# Beneficiary Information on the Cloud (BIC)

Publishing Beneficiary Information to a CMS Enterprise Data Sharing Platform

Table Name	Description
beneficiary_maenrollment_segment	Segment level enrollment periods for a beneficiary enrolled in Medicare Advantage (MA) or Medicare Advantage Prescription Drug Plan (MAPD).
beneficiary_mapd_enrollment_address	Beneficiary's residence address information submitted by a plan.
beneficiary_name	Name changes for a beneficiary. Common reasons for change are marriage or divorce.
beneficiary_part_b_premium_reduction	Premium reduction amount for a beneficiary's Part B premium.
beneficiary_part_d_cumulative_uncovered_months	System calculated, point in time, number of months a beneficiary did not have creditable coverage and was eligible for, but not covered by, a Part D plan.
beneficiary_part_d_irmaa_history	Beneficiary's Social Security Administration (SSA) Part D Income Related Monthly Adjustment Amount (IRMAA) payment status history.
beneficiary_part_d_irmaa_period	Beneficiary's billable Medicare Part D Income Related Monthly Adjustment Amount (IRMAA) period information.
beneficiary_part_d_other_insurance	Beneficiary's other prescription insurance that is reported by the plan.
beneficiary_premium_part_c_d	Beneficiary's Part D premium basic rate (net of rebates) amount that is part of the Part D premium calculation.
beneficiary_premium_withhold_election_status	Status of the beneficiary's premium withhold option. The option identifies if the premium will be withheld from the beneficiary's payment/check, paid directly by the beneficiary, or if there is no dollar amount to withhold.



# Beneficiary Information on the Cloud (BIC)

Publishing Beneficiary Information to a CMS Enterprise Data Sharing Platform

Table Name	Description
beneficiary_ssa_part_d_irmaa_benefit_payment	Beneficiary's Part D Income Related Monthly Adjustment Amount (IRMAA) premium or arrearage collection responsibility.
beneficiary_status	Current snapshot of core information for a beneficiary. Includes identifiers, name, and system status (active or archived).
beneficiary_third_party_history_part_a	Beneficiary's Part A third party periods. The third party billing system collects Medicare Part A premium payments from third party agencies that pay bills on behalf of qualified Medicare beneficiaries.
beneficiary_third_party_history_part_b	Beneficiary's Part B third party periods. The third party billing system collects Medicare Part B premium payments from third party agencies that pay bills on behalf of qualified Medicare beneficiaries.
combined_dual_medicare_medicaid	Medicare/Medicaid dual eligibility status history for a beneficiary.
common_working_file_location	Common Working File (CWF) location where beneficiary utilization records are maintained.
comprehensive_addiction_recovery_event	Beneficiary's Comprehensive Addiction and Recovery Act (CARA) event. Part D plan sponsors may establish a drug management program for beneficiaries at risk for prescription drug abuse or misuse. If casework reveals a problem, the beneficiary's access to coverage of frequently abused controlled drugs may be restricted to a selected prescriber and/or network pharmacy. Additionally, CMS limits the use of the Special Enrollment Period (SEP) for dually or other Low Income Subsidy (LIS) beneficiaries who are identified as at-risk or potentially at-risk for prescription drug abuse under a drug management program.

# Beneficiary Information on the Cloud (BIC)

Publishing Beneficiary Information to a CMS Enterprise Data Sharing Platform

Table Name	Description
contract	Administrative information about an organization that provides coverage to eligible Medicare beneficiaries. This information includes the contract number, contract name, and contract address information. The contract number begins with 'H' for a Medicare Advantage (MA) plan and Medicare Advantage Prescription Drug (MAPD) plan. The contract number begins with 'S' for a Medicare Prescription Drug Plan (PDP).
dual_month_statistics	Count of enrollments and disenrollments for each eligibility month from a Medicare Modernization Act (MMA) state file. It also qualifies the counts by matched or non-matched to the Medicare Beneficiary Database (MBD).
eligibility_part_d	Part D (prescription drug) eligibility history for a beneficiary.
enrollment_mapd	Beneficiary's Medicare Advantage (MA), Medicare Advantage (MAPD), and Prescription Drug Plan (PDP) enrollment history.
entitlement_disability	Disability entitlement details for a beneficiary.
entitlement_part_a	Part A entitlement history for a beneficiary.
entitlement_part_b	Part B entitlement history for a beneficiary.
entitlement_reason	Reason for beneficiary's Medicare entitlement.
esrd_coverage	Beneficiary's periods of Medicare coverage based on End Stage Renal Disease (ESRD).

# Beneficiary Information on the Cloud (BIC)

Publishing Beneficiary Information to a CMS Enterprise Data Sharing Platform

Table Name	Description
esrd_dialysis_clinical	Beneficiary's clinical End Stage Renal Disease (ESRD) dialysis start and stop dates. The clinical dates reflect the actual begin and end dates for the dialysis health service.
esrd_self_training	First instance of the End Stage Renal Disease (ESRD) self-care training.
esrd_transplant	End Stage Renal Disease (ESRD) kidney transplant occurrence date and failure date.
hospice_coverage	Beneficiary's current and historical periods of hospice elections.
incarceration_period	Incarceration periods for beneficiary.
low_income_subsidy_deemed	Beneficiary that has been deemed through the CMS deeming process to receive 100% low income subsidy for Part D.
low_income_subsidy_deemed_copayment	Part D deemed beneficiary's copayment information. A beneficiary's copayment may change during the effective period.
low_income_subsidy_denied_ssa	Beneficiary that has been denied Low Income Subsidy (LIS). Disapproval date as well as the denial reason codes are included.
low_income_subsidy_effective	Effective Part D Low Income Subsidy (LIS) status history for a beneficiary.
low_income_subsidy_ssa	Beneficiary that has applied and been approved for low income subsidy. The effective period for the subsidy is identified as well as additional information that relates to the subsidy determination for this period.
mailing_address_override	Request for overriding Medicare card suppression for a beneficiary with an address that failed address validation.

# Beneficiary Information on the Cloud (BIC)

Publishing Beneficiary Information to a CMS Enterprise Data Sharing Platform

Table Name	Description
mailing_address_scrubbed	Beneficiary's mailing address that has been scrubbed using addressing validation software.
medicare_beneficiary_identifier	Medicare Beneficiary Identifiers (MBIs) assigned to a beneficiary.
medicare_part_d_irmaa_rate	Reference table for the Medicare Part D Income Related Monthly Adjustment Amount (IRMAA) rates/tiers.
notification_instruction2	Data dissemination (DDX) notification instructions for a beneficiary.
not_lawful_presence_period	Beneficiary's not lawful presence periods.
opt_out_financial_alignment_demo	Beneficiary's choice to opt out of the enrollment into a Financial Alignment Demonstration (FAD) plan.
opt_out_part_d_automatic_assignment	Beneficiary's election not to be automatically enrolled by CMS in a Part D plan.
other_drug_coverage_cob	Beneficiary's period of enrollment in other drug insurance separate from Medicare Part D.
part_d_prescription_card	Beneficiary's Medicare Part D prescription identification information.
pbp_contact	Contract level contact information for the following contract contacts: MA appeals/grievances contact and Part D appeals contact. Record Identification Code (RIC) Q.

# Beneficiary Information on the Cloud (BIC)

Publishing Beneficiary Information to a CMS Enterprise Data Sharing Platform

Table Name	Description
pbp_status	Record Identification Code (RIC) AB - Plan Benefit Package (PBP) record. This record contains a description of a PBP offered under the contract. The PBP links to the contract through the contract number. There may be one, many, or no PBP records associated with a contract.
retiree_drug_benefit	Beneficiary's drug benefits from their employer retirement plan.
retiree_drug_plan	Beneficiary's employer drug plan information including enrollment periods.
segment_status	Segment (service area) where a plan is offered under a contract. [Record Identification Code (RIC) AS].
social_security_number	History of Social Security Numbers (SSN) assigned to a beneficiary.
ssa_benefit_payment_status	Social Security Administration (SSA) benefit payment status for a beneficiary. The benefit payment code is also known as the Ledger Account File (LAF) code.
state_county_scrubbed	Beneficiary state and county history that has been scrubbed using address validation software.
state_county_ssa	State and county history for a beneficiary as received from the Social Security Administration (SSA).

# Beneficiary Information on the Cloud (BIC)

Publishing Beneficiary Information to a CMS Enterprise Data Sharing Platform

*Table 2: Record of Changes*

Version Number	Date	Author/Owner	Description of Change
0.1	02/12/2020	Lakshmi Manambedu	Draft version.

DRAFT