**COBOL Program EDBNAECE Overview**

*NAS Evergreening System - Data Quality and Maintenance Processing*

This document describes the COBOL program EDBNAECE, which is part of the "NAS Evergreening" system for data quality and maintenance processing within a mainframe batch environment.

## Program Purpose and Business Context

The program serves as a critical component in the NAS Evergreening data processing pipeline, functioning as a batch summarizer and statistical reporter. It processes audit logs and DIN records to maintain data quality and support business operations.

### Core Functions:

* Processes audit logs and DIN records (Data Identification Numbers)
* Counts and aggregates key business metrics
* Updates statistical files and generates comprehensive reports
* Monitors data quality and processing health
* Creates reapply files for downstream correction workflows

### Operating Modes:

* **Read Mode ('R')** - Performs analysis and reporting without creating output files
* **Update Mode ('U')** - Performs analysis, reporting, and creates reapply files for reprocessing

The program plays a vital role in ensuring data integrity by tracking metrics such as data quality changes, error counts, and processing volumes. It supports broader business goals of maintaining accurate data without manual intervention and enables systematic correction workflows.

## Input Processing and File Structures

### Audit Input File (AUDITI)

This variable-length file contains comprehensive audit records from the NAS Evergreening process. The program sequentially reads these records to accumulate critical business metrics.

#### Record Structure:

* Error and processing status flags
* Counts of associated PINs (Primary Identification Numbers) and LINs (Linked Identification Numbers)
* Detailed timestamps and primary DIN identifier
* Source data including entity names and address information
* Quality classification codes: 'H' (High), 'L' (Low), 'T' (Text-based)
* Error codes for diagnostic information
* Arrays supporting up to 500 PINs, 200 LINs, and 2200 DIN reapply entries per record

#### Processing Usage:

The program accumulates PIN counts, LIN counts, error frequencies, and quality transition statistics that inform data health monitoring and business intelligence reporting.

### DIN Input File (DININP)

This fixed-length file contains DIN records with associated subject identification data, serving dual purposes in the processing workflow.

#### Record Components:

* Primary DIN identifier
* Subject numbers and sequence numbers
* Hierarchical relationship data within the data structure

#### Processing Functions:

* Provides input for DIN counting operations
* Serves as source for generating reapply records in update mode
* Supports duplicate detection through sequential processing logic

### Statistics File (STATSO)

This input-output file maintains accumulated statistics from previous program executions and serves as both data source and target.

#### Maintained Metrics:

* Reapply DIN counts
* Total PIN counts and deleted LIN counts
* Cumulative error totals
* Processing time metrics (CPU and elapsed time)

#### Processing Cycle:

At initialization, existing statistics establish baseline values. Upon completion, updated totals reflecting current processing sessions maintain historical continuity for trend analysis.

### Parameter Interface

#### Required Parameters:

* **Call Mode:** 'U' (Update operations) or 'R' (Read-only operations)
* **Logging Level:** Controls output verbosity
* **Log Length:** Specifications for external logging module interface

Invalid parameter values trigger immediate program termination with appropriate error messaging.

## Output Generation and Reporting

### Reapply Output File (REAPPLY)

Generated exclusively during update mode operations, containing DIN records identified for reprocessing or correction.

#### Record Structure:

* Source process identifiers and run date stamps
* Primary DIN value and subject count indicators
* Structured arrays with detailed subject number and sequence information
* Capacity for up to 9 subject occurrences per DIN

#### Key Processing Features:

* Implements duplicate detection logic comparing each DIN against previously processed values
* Ensures only unique DINs are written to prevent downstream processing inefficiencies
* Maintains data integrity in subsequent operations through deduplication mechanisms

### Detailed Statistics File (STATSD)

This human-readable reporting file provides comprehensive monitoring information for business users and operational teams.

#### Report Format:

* Descriptive headers (up to 40 characters) with corresponding numeric values
* Zero-suppressed numeric formatting for enhanced readability
* Consistent line structure throughout the report

#### Content Examples:

* "NUMBER OF ERRORED RECORDS" with corresponding count values
* "TOTAL NUMBER OF INPUT RECORDS" providing overall processing volumes
* Quality improvement transitions and detailed processing breakdowns

### Statistics File Updates

The original statistics file receives comprehensive updates reflecting current processing session results, maintaining historical trends for operational analysis.

#### Updated Fields Include:

* Final reapply DIN counts and accumulated PIN totals
* LIN deletion counts and error summaries
* Detailed timing information (CPU utilization and elapsed processing time)
* Historical processing trends for capacity planning and performance monitoring

## Processing Flow and Business Logic

### Initialization Phase

Program startup establishes the foundation for successful processing through systematic validation and resource allocation.

#### Key Activities:

* Comprehensive parameter validation with immediate termination for invalid inputs
* External module integration for logging (ET530) and time tracking (E1902) services
* File opening operations with appropriate error handling
* Baseline statistics reading from existing statistics file
* Sequential input file positioning and time recording initiation

### Main Processing Operations

The core processing implements dual-loop architecture to handle audit and DIN file processing independently while maintaining operational efficiency.

#### Audit Processing Loop:

* Accumulates business metrics from PIN counts, LIN counts, and error indicators
* Maintains running totals reflecting overall processing health
* Tracks data quality trends and processing effectiveness
* Captures quality change transitions between different data quality states

#### DIN Processing Loop:

* Increments DIN counters for volume tracking
* Evaluates records for reapply file inclusion during update mode
* Implements duplicate detection logic comparing current against previously processed entries
* Ensures efficient downstream processing through redundancy elimination

#### Business Intelligence Capture:

* Quality change tracking provides insights on data improvement trends
* Error accumulation enables operational monitoring for system health indicators
* Processing metrics support capacity planning and performance optimization

### Termination and Reporting

Program completion ensures data integrity and provides comprehensive audit trails for operational monitoring.

#### Completion Activities:

* Statistics file updates with accumulated totals from current processing session
* Detailed statistics report generation with formatted business-friendly output
* File closure operations ensuring data integrity and resource cleanup
* Final logging operations providing audit trails and performance metrics
* Graceful termination with appropriate return codes for batch workflow integration

#### Operational Characteristics:

The program maintains operational efficiency through file-based processing architecture, utilizing external modules for logging and timing services while avoiding database dependencies in the mainframe environment.