



Estándar - Depuración Automática DB

Versión 1.0

Arquitectura y Estándares TI

31/12/2021



Tabla de contenido

Versiones.....	1
Convenciones.....	1
Objetivo.....	2
Descripción.....	2
Consideraciones	2
Esquema Mensual.....	2
Esquema Acumulado.....	3
Solución.....	3
Observaciones	6

Versiones

Descripción	Versión	Autor	Fecha
Publicación de documento original	1.0.0	Braian Espejo	31/12/2021

Convenciones

Las convenciones usadas en este documento son:

Convención	Significado
DB	Base de Datos
Filegroup	Son útiles para distribuir tablas con alto volumen de información en diferentes discos para separar los índices de los datos.
Backup	Una copia de seguridad es un duplicado de los datos que se hace para poder recuperarlos ante cualquier pérdida o incidente.
Script	Conjunto de comandos, que se le da a un motor SQL para decirle lo que debe hacer y en qué orden debe hacerlo.
Paquete SSIS	Forma parte de la plataforma de base de datos Microsoft SQL Server y una herramienta para construir aplicaciones de flujo de trabajo e integración de datos de alto rendimiento.

Objetivo

El documento tiene como objetivo estandarizar la depuración automática de las tablas en las DB creadas dentro del Banco de Crédito de Bolivia SA.

Descripción

La depuración de datos consiste en el proceso de eliminación de datos y la reorganización de las tablas en una DB. El proceso se utiliza principalmente cuando una o varias tablas contiene mucha información almacenada.

Si queremos asegurarnos de que los datos se utilicen de la manera más productiva posible, de tal forma que podamos añadir rapidez en las consultas y respuestas se DEBE utilizar los esquemas de depuración.

Éstas son las ventajas que ofrece utilizar los esquemas de depuración:

- **Mejora la eficiencia**

Mediante la depuración de datos, se elimina los datos no necesarios y/o desactualizados, podemos crear consultas más eficientes, después del proceso de depuración los datos han de estar limpios, actualizados y exactos.

La duplicación de datos es otro aspecto que puede ser erradicado. El impacto de la duplicación es directamente proporcional al tiempo que permanecen en la base de datos.

- **Mejora el proceso.**

Con unos datos limpios se puede ejecutar mejor las consultas costosas en hardware, lo cual contribuirá al mejor desempeño de una aplicación.

- **Agiliza las respuestas**

La eliminación de datos no necesarios y/o desactualizados ayuda a optimizar en tiempo las respuestas de una aplicación.

- **Aumenta rendimiento**

Tener una DB limpia y bien mantenida puede evitar que se utilice en todo momento datos innecesarios y/o desactualizados. Al trabajar con datos limpios maximizamos la eficiencia y la productividad de los recursos del servidor.

Consideraciones

Para que la DB pueda entrar en el esquema de depuración automática es requerido que la misma no contenga datos de más de 2 años, de lo contrario se debe realizar previamente una depuración manual, además se sugiere realizar previamente un backup de la DB antes de aplicar el esquema de depuración.

Esquema Mensual

En este esquema se creará una tabla de depuración con el mes en curso, en la tabla creada se guardarán los datos de la depuración de la tabla principal.

- En la tabla principal se DEBE guardar solo la información de 7 días, para tablas altamente transaccionales solo se DEBE tener la información del día, la cantidad de días es configurable en el script que se explicara líneas abajo.

- El nombre de las tablas de depuración será igual a la tabla principal donde solo se agregará el mes en curso (YYYYMM).

Donde: auth.HISTORIAL_CONSUMO, es la tabla principal

auth.HISTORIAL_CONSUMO_202009, es la tabla de depuración



- Si se utiliza el comando Filegroup en la creación de las tablas de depuración estas pueden ser almacenadas en otra unidad, esto será mostrado líneas abajo.
- Solo se DEBE guardar 3 meses de historial en tablas de depuración, la cantidad de meses también es configurable en el script.

Esquema Acumulado

En este esquema se realiza la acumulación de meses en la misma tabla no crea tablas de depuración adicionales, la depuración de datos no necesarios es de manera automática, la cantidad de días a acumular es configurable.

Solución

La solución de ejemplo la pueden solicitar al **Área de Arquitectura y Estándares TI**, también se encuentra en el SharePoint en la carpeta de **BDD**. La misma contiene la estructura de una tabla para controlar los parámetros de depuración, un script de ejemplo de depuración por tabla, un componente SSIS para la ejecución del script de depuración y un Job para la ejecución del componente SSIS.

Para tal implementar la solución se deben seguir los siguientes puntos:

- Dentro de la DB a fin de no manipular los scripts que se describirá más adelante, debe crearse la tabla **PARAMETRO_DEPURA** con los siguientes campos:

CAMPO	TIPO	VALORES	DESCRIPCION
PADE_ID_IN	(int, not null)	Autogenerado	PRIMARY KEY, not null
PADE_TABLA_VC	(varchar(250), null)	HISTORIAL_CONSUMO	Nombre de la tabla a depurar
PADE_HISTORIAL_DIA_SI	(smallint, null)	1	Representa la cantidad de días de los que se debe mantener información en la tabla principal.
PADE_HISTORIAL_MES_SI	(smallint, null)	3	Determina la cantidad de meses de los que se respaldará información.
PADE_ACUMULADO_BT	(bit, null)	0	Campo de control para utilizar el esquema Acumulado = 1 o Mensual = 0
PADE_ACUMULADO_MES_SI	(smallint, null)	3	Determina la cantidad de meses de los que se mantendrá información en la tabla transaccional, en el esquema Acumulado.
PADE_ORDEN_IN	(int, null)	3	Establece orden de ejecución de los scripts de depuración.
PADE_USUARIOS_CREACION_VC	(varchar(20), null)	S31874	Campo de control.
PADE_FECHA_CREACION_DT	(datetime, null)	2020-08-03 00:00:00.000	Campo de control.
PADE_USUARIO_MODIFICACION_VC	(varchar(20), null)	S31874	Campo de control.
PADE_FECHA_MODIFICACION_DT	(datetime, null)	2020-08-03 00:00:00.000	Campo de control.
PADE_ESTADO_BT	(bit, null)	1	Estado lógico del registro.

2. A continuación, debe elaborarse el siguiente script **DEPURA_TABLA_XXXX**, Donde XXXX es el nombre de la tabla principal

```

USE [NOMBRE_BD]
GO
/***** Object: StoredProcedure [auth].[DEPURA_TABLA_HISTORIAL_NOMBRE TABLA] Script
Date: 15/11/2021 08:49:34 a.m. *****/
SET ANSI_NULLS ON
GO
SET QUOTED_IDENTIFIER ON
GO

ALTER PROCEDURE [auth].[DEPURA_TABLA_HISTORIAL_NOMBRETABLA]
/*****
* Descripción: Depurar tablas transaccionales a tablas historial mes
* Tablas: [auth].[HISTORIAL_NOMBRETABLA]
* Fecha Creacion: 09/08/2020
* Fecha Modificacion: 09/08/2020
* Autor: Braian Javier Espejo Peralta (S22715)
* Versión: Alfa|Beta Modificado Braian Javier Espejo Peralta (S22715)
*****/
AS
BEGIN
    DECLARE @FECHA_MES_CH AS CHAR(8)
    DECLARE @FECHA_DIA_CH AS CHAR(8)
    DECLARE @HISTORIAL_MES_VC VARCHAR(MAX)
    DECLARE @NOMBRE_TABLA_VC VARCHAR(MAX)
    DECLARE @CREAR_TABLA_VC VARCHAR(MAX)

    DECLARE @PADE_HISTORIAL_DIA_SI SMALLINT
    DECLARE @PADE_HISTORIAL_MES_SI SMALLINT
    DECLARE @PADE_ACUMULADO_BT BIT
    DECLARE @PADE_ACUMULADO_MES_SI SMALLINT

    SELECT @PADE_HISTORIAL_DIA_SI = [PADE_HISTORIAL_DIA_SI]
        ,@PADE_HISTORIAL_MES_SI = [PADE_HISTORIAL_MES_SI]
        ,@PADE_ACUMULADO_BT = [PADE_ACUMULADO_BT]
        ,@PADE_ACUMULADO_MES_SI = [PADE_ACUMULADO_MES_SI]
    FROM [NOMBRE_BD].[PARAMETRO_DEPURA]
    WHERE [PADE_TABLA_VC] = '[NOMBRETABLA]'
        AND [PADE_ESTADO_BT] = 1

    SET @FECHA_DIA_CH = CONVERT(CHAR(8), DATEADD(DAY, - @PADE_HISTORIAL_DIA_SI,
        GETDATE()), 112)
    SET @NOMBRE_TABLA_VC = 'NOMBRETABLA_' + CONVERT(CHAR(6), @FECHA_DIA_CH, 112)
    SET @HISTORIAL_MES_VC = 'NOMBRETABLA_' + CONVERT(CHAR(6), DATEADD(MM, -
        @PADE_HISTORIAL_MES_SI, GETDATE()), 112)

    IF (@PADE_ACUMULADO_BT = 1)
    BEGIN
        --Elimina en la tabla todo lo anterior a 3 meses
        DELETE
        FROM [NOMBRETABLA]
        WHERE CONVERT(CHAR(8), [CAMPOFECHA], 112) <= @PADE_ACUMULADO_MES_SI
    END
    ELSE
    BEGIN
        --Si no existe, crea la tabla de Backup mensual
        IF OBJECT_ID(@NOMBRE_TABLA_VC, 'U') IS NULL
        BEGIN
            SET @CREAR_TABLA_VC =
                ' SELECT [CAMPO1]' +
                ' , [CAMPO2]' +
                ' , [CAMPOFECHA]' +
                ' INTO ' + @NOMBRE_TABLA_VC +
                --' ON [DEPURATION] ' + -- Realiza el guardado en el FileGroup [DEPURATION]
                ' FROM [NOMBRETABLA] ' +
                ' WHERE CONVERT(CHAR(8), [CAMPOFECHA], 112) <= '' ' +
                @FECHA_DIA_CH + ''''
        END
    END
    ELSE

```

```

BEGIN
-- Deshabilita el identity key e inserta datos para el backup.
SET @CREAR_TABLA_VC =
' SET IDENTITY_INSERT ' + @NOMBRE_TABLA_VC + ' ON ' +
' INSERT INTO ' + @NOMBRE_TABLA_VC + ' ( ' +
'      [CAMPO1]' +
'      , [CAMPO2]' +
'      , [CAMPOFECHA]' +
' SELECT [CAMPO1]' +
'      , [CAMPO2]' +
'      , [CAMPOFECHA]' +
' FROM [NOMBRETABLA]' +
' WHERE CONVERT(CHAR(8), [CAMPOFECHA], 112) <= '' ' +
@FECHA_DIA_CH + '''' +
' SET IDENTITY_INSERT ' + @NOMBRE_TABLA_VC + ' OFF '

END

EXEC (@CREAR_TABLA_VC)

IF @@ERROR = 0
BEGIN
DELETE
FROM [NOMBRETABLA]
WHERE CONVERT(CHAR(8), [CAMPOFECHA], 112) <= @FECHA_DIA_CH

END

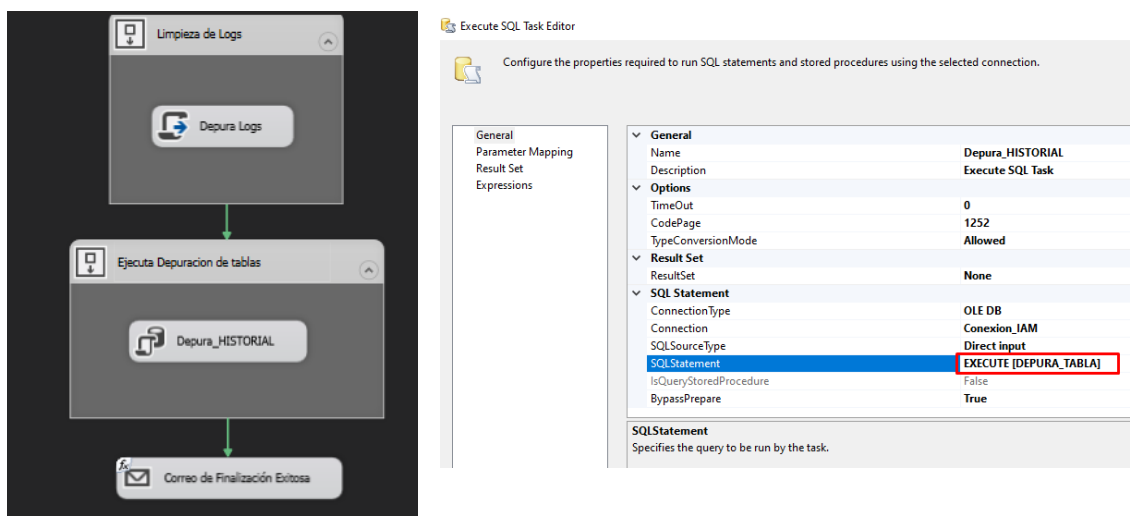
IF OBJECT_ID(@HISTORIAL_MES_VC) IS NOT NULL
BEGIN
EXEC ('DROP TABLE ' + @HISTORIAL_MES_VC)

END

END
END

```

3. La configuración en el componente SSIS se realizara de la siguiente manera



4. Un Job debe programarse y llamar al paquete descrito anteriormente.

Observaciones

En caso de tener alguna observación o agregar alguna consideración, enviarlo a los siguientes correos:

Bolivia - Arquitectura y Estándares TI

BolArqEstaTI@bancred.com.bo

Braian Espejo Peralta

BEspejo@bcp.com.bo

Jose Zurita Plata

JZurita@bcp.com.bo

Diego Tarquino Tapia

DTarquino@bcp.com.bo