



**UNIVERSIDAD NACIONAL
“PEDRO RUIZ GALLO”
FACULTAD DE INGENIERIA CIVIL,
SISTEMAS Y ARQUITECTURA**



ESCUELA PROFESIONAL DE INGENIERIA DE SISTEMAS

“Mejoramiento del proceso de generación de base de contactabilidad de clientes en Banco Pichincha”

TRABAJO DE SUFICIENCIA PROFESIONAL

**PRESENTADA PARA OPTAR EL TITULO
PROFESIONAL DE:**

INGENIERO DE SISTEMAS

**PRESENTADA POR EL BACHILLER:
Llanos Edquen, Nixon James**

**ASESOR:
Mg. Ing. Arteaga Lora, Roberto Carlos**

LAMBAYEQUE – PERU

2022



**UNIVERSIDAD NACIONAL
"PEDRO RUIZ GALLO"
FACULTAD DE INGENIERIA CIVIL,
SISTEMAS Y ARQUITECTURA**



ESCUELA PROFESIONAL DE INGENIERIA DE SISTEMAS

**"Mejoramiento del proceso de generación de base de contactabilidad de
clientes en Banco Pichincha"**

DR. ING. EDWARD RONALD HARO MALDONADO

PRESIDENTE DE JURADO

MG. ING. GAVINO MARCELO LOYAGA ORBEGOSO

SECRETARIO DE JURADO

MG. ING. MARÍA DE LOS ÁNGELES GUZMÁN VALLE

VOCAL

Línea de Investigación:

Sistemas de información



**UNIVERSIDAD NACIONAL
"PEDRO RUIZ GALLO"**
**FACULTAD DE INGENIERIA CIVIL,
SISTEMAS Y ARQUITECTURA**



ESCUELA PROFESIONAL DE INGENIERIA DE SISTEMAS

**"Mejoramiento del proceso de generación de base de contactabilidad de
clientes en Banco Pichincha"**

MG. ING. ARTEAGA LORA ROBERTO CARLOS.

ASESOR

LLANOS EDQUEN, NIXON JAMES

BACHILLER EN ING. SISTEMAS

Dedicatoria

A mi querida esposa Elvia y a mis hijos

*Por ser mi mayor motivación de mi vida para salir adelante
y nunca rendirme en mis objetivos y ser un ejemplo para ellos*

Dios siempre los bendiga y proteja su camino.

A mis padres,

*Por forjar a la persona que soy en la actualidad, sus valores y la forma de ver la
vida siendo siempre optimistas y perseverantes.*

Agradecimientos

Agradezco mi Esposa

Por ser mi compañía y alentarme a presentar el informe.

Gracias al gran sacrificio que ambos hicimos para llegar a este momento.

Agradezco a Dios

Por bendecirme con la vida que me ha dado y darme la oportunidad de vivir estos momentos, por guiar mi camino y nunca soltarme de su mano.

Agradezco a mi Escuela Profesional de Ingeniería de Sistemas de la Universidad Nacional Pedro Ruiz Gallo, su gran trabajo de formación de profesionales e impulsar el desarrollo de la región Lambayeque, de forma especial a mi Asesor Ing. Roberto Arteaga por ser

Agradezco a mi asesor y docentes, su gran interés por llegar a este objetivo

ÍNDICE

RESUMEN.....	7
ABSTRACT	8
I. DEFINICIÓN DEL PROBLEMA	9
Descripción del Proceso Actual.....	11
II. OBJETIVOS	20
OBJETIVO GENERAL	20
OBJETIVOS ESPECIFICOS	20
III. Limitaciones de la investigación	21
IV. JUSTIFICACIÓN DEL INFORME	22
V. FUNDAMENTO TEÓRICO	23
1. Ciclo Deming o ciclo PHVA	23
2. SQL Server	23
3. ETL	24
4. JOB en SQL.....	24
5. Procedimiento Almacenado (Store Procedure).....	25
6. Diagrama Flujo.....	25
7. Tablas en SQL	25
8. Indexación de Columnas de Tabla	26
VI. GLOSARIO DE TERMINOS.....	27
VII. DESARROLLO DE LA SOLUCIÓN	28
1. DEFINICION DE METODOLOGIA.....	28
2. APLICACIÓN DE METODOLOGIA.....	28
A. PLAN	28
B. HACER	28
C. VERIFICAR	51
D. ACTUAR.....	55
VIII. CONCLUSIONES	56

IX. RECOMENDACIONES	57
X. LISTA DE REFERENCIAS BIBLIOGRÁFICAS	58

ÍNDICE DE FIGURAS

Figura 1	Reporte de ejecución del proceso BUC.....	10
Figura 2	Ubicación del proceso de BUC, configurado en el Agente del Servidor 43	10
Figura 3	Programación de la ejecución del JOB a las 2:30 de forma diaria	11
Figura 4	Reporte exitoso de la ejecución del proceso BUC.....	11
Figura 5	Detalle de pasos que ejecuta el JOB en SQL Server 43	12
Figura 6	Generación de base compras	12
Figura 7	Configuración de ejecución de ETL importación de tablas	13
Figura 8	Detalle de paquete de Integration Services	13
Figura 9	Detalle de paso: Generación de DNI	14
Figura 10	Detalle de Paso: Generación de Teléfonos.....	14
Figura 11	Detalle de paso Generación de Emails	15
Figura 12	Representación del proceso actual BUC 1/2.....	16
Figura 13	Representación del proceso actual BUC 2/2.....	18
Figura 14	Caída del proceso en el paso 6	29
Figura 15	Caída del proceso en el paso 5	29
Figura 16	Caída del proceso en el paso 4	30
Figura 17	Caída del proceso en el paso 2	30
Figura 18	Caída del proceso en el paso 3	31
Figura 19	Configuración de archivo log	31
Figura 20	Detalle de espacio en Servidor 43	31
Figura 21	Reporte de ejecución de tareas de limpieza	32
Figura 22	Diagrama de flujo del escenario actual	33
Figura 23	Identificación de tablas redundantes en tres secuencias	34
Figura 24	Flujo propuesto con optimizaciones	35
Figura 25	Secuencia de pasos que ejecuta la ETL.....	37
Figura 26	Query de limpieza de tablas en el Servidor 43.....	37
Figura 27	Secuencias de transferencia de tablas del servidor 1 al 43.....	38

Figura 28	Secuencias de transferencia de tablas del servidor 15 al 43.....	38
Figura 29	Optimización de importación de tabla BUC_TELEFONO.....	39
Figura 30	Configuración de control de búsqueda.....	39
Figura 31	Identificación de PK para la importación diferencial BUC_TELEFONO..	40
Figura 32	Creación de índice para BUC_TELEFONO	40
Figura 33	Primera ejecución con mejora de Importación de BUC_TELEFONO	40
Figura 34	Segunda ejecución con mejora de Importación de BUC_TELEFONO ...	41
Figura 35	Verificación de importación de BUC_TELEFONO	41
Figura 36	Optimización de importación de tabla BUC_EMAIL	42
Figura 37	Configuración de control de búsqueda para BUC_EMAIL.....	42
Figura 38	Identificación de PK para la importación diferencial BUC_EMAIL	42
Figura 39	Creación de índice BUC_EMAIL.....	43
Figura 40	Primera ejecución con mejora de importación BUC_EMAIL	43
Figura 41	Segunda ejecución con mejora de importación BUC_EMAIL.....	43
Figura 42	Verificación de importación BUC_EMAIL.....	44
Figura 43	Optimización de importación BUC_CLIENTE	44
Figura 44	Identificación de PK para la importación diferencial BUC_CLIENTE.....	44
Figura 45	Creación de índice BUC_CLIENTE	45
Figura 46	Ejecución de mejora de importación BUC_CLIENTE.....	45
Figura 47	Optimización de importación TT_COMERCIAL	46
Figura 48	Creación de índice TT_COMERCIAL	46
Figura 49	Ejecución de mejora de importación TT_COMERCIAL.....	47
Figura 50	Optimización de importación MAE_CLIENTES.....	47
Figura 51	Creación de índice MAE_CLIENTES.....	47
Figura 52	Ejecuciones de mejora de importación MAE_CLIENTES	48
Figura 53	Diagrama simplificado de Importación de Tablas.....	48
Figura 54	Query de Limpieza de LOG de BF_BUC	49
Figura 55	Diagrama final con optimizaciones de importación de tablas	49

ÍNDICE DE TABLAS

Tabla 1	Descripción de la programación del JOB del proceso BUC	9
Tabla 2	Detalle de tiempo y volumetría en cada Paso	19
Tabla 3	Especificaciones técnicas del de Servidor 43	19
Tabla 4	Detalle de causas de caída en del proceso base de contactabilidad	28
Tabla 5	Cuadro comparativo de escenarios: actual vs propuesto	52
Tabla 6	Cuadros comparativos de tablas resultantes	53
Tabla 7	Cuadro comparativo de cantidad de datos en ambos escenarios.....	54

RESUMEN

Ante el crecimiento del volumen de datos en las organizaciones; estas tienden a tener problemas sobre los procesos que los generan, por ese motivo los procesos deben estar en constante revisión y mejora. Los datos de contactabilidad para las entidades financieras son críticos para la colocación de sus productos crediticios, por lo cual, el proceso para la generación de esta data es de vital importancia que sea eficiente. En los meses de mayo, junio y julio del 2021 se han reportado caídas de este proceso, todas debido a que el log de la base de datos en SQL sobrepasa el límite y además por la falta de espacio en el servidor donde se ejecutan. Estas caídas generan contratiempos para el despliegue de campañas de oferta de préstamos y por ende una disminución de las colocaciones de productos: lo descrito ocurrió en Banco Pichincha Perú, donde tengo la oportunidad de laborar en el área de Manejo de relaciones con el cliente (CRM), utilizando la metodología de mejora continua (Ciclo de Deming), se logró identificar oportunidades de mejora en los distintos niveles del proceso: a nivel de flujo, secciones e instrucciones, las mejoras implementadas ayudan a que el proceso sea más eficiente, tomando en cuenta la transferencia de información, la reducción en 49% del tiempo de ejecución y contar con una base de contactabilidad altamente disponible para las futuras campañas.

ABSTRACT

Faced with the growth of the volume of data in organizations; These tend to have problems about the processes that generate them, for that reason the processes must be constantly reviewed and improved. The contactability data for financial institutions is critical for the placement of their credit products, therefore, the process for generating this data is vitally important to be efficient. In the months of May, June and July 2021, crashes in this process have been reported, all due to the fact that the log of the database in SQL exceeds the limit and also due to the lack of space on the server where they are executed. These falls generate setbacks for the deployment of loan offer campaigns and therefore a decrease in product placements: what has been described happened in Banco Pichincha Peru, where I have the opportunity to work in the area of Customer Relationship Management (CRM), using the continuous improvement methodology (Deming Cycle), it was possible to identify opportunities for improvement at the different levels of the process: at the flow level, sections and instructions, the implemented improvements help make the process more efficient, taking into account account for information transfer, a 49% reduction in execution time and having a highly available contact base for future campaigns.

I. DEFINICIÓN DEL PROBLEMA

Actualmente, el área de CRM en el Banco Pichincha, esta área es la encargada del enriquecimiento con datos de contactabilidad para las campañas de ofertas de préstamos.

La base única de contactabilidad, es una información crítica para el despliegue de campañas, la cual, para ser oportuna, debe ser actualizada diariamente.

El proceso que genera la base única de contactabilidad a quien llamaremos por sus siglas “BUC”, se realiza en un servidor SQL denominado “Servidor 43” donde el área de CRM y otras áreas desarrollan sus procesos.

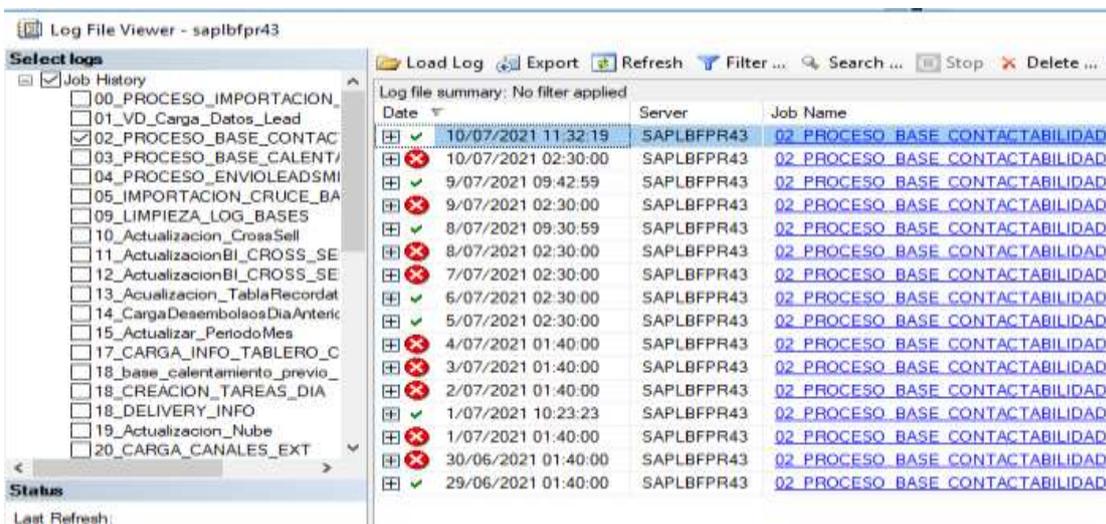
Tabla 1 Descripción de la programación del JOB (tarea) del proceso BUC

DESCRIPCIÓN	DETALLE
Nombre	02_PROCESO_BASE_CONTACTABILIDAD
Tiempo de ejecución	1 hora aproximadamente
Hora de ejecución	2:30 AM
Periodicidad	Diaria
Agente de ejecución	Agente SQL Server Servidor 43

Dicho proceso en los meses de Mayo, Junio y Julio del presente año, presentaron caídas de ejecución por falta de espacio y de recursos, se han revisado los reportes de estas ejecuciones y se aprecia que el tiempo de ejecución de aproximadamente 1 hora aumentó hasta casi 2 horas durante los meses indicados.

Al recibir una alerta del área de gobierno de datos, sobre varias caídas del proceso, se tiene que revisar cual es el paso en donde se produjo la caída y tener que ejecutar nuevamente el proceso desde ese punto.

Figura 1 Reporte de ejecución del proceso BUC



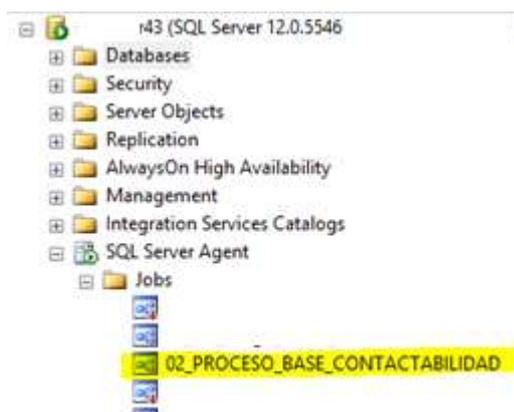
The screenshot shows a 'Log File Viewer' window for server 'saplbfp43'. The left pane shows a tree view of job history with '02 PROCESO BASE CONTACTABILIDAD' selected. The main pane displays a table of job execution logs.

Date	Server	Job Name
10/07/2021 11:32:15	SAPLBFP43	02 PROCESO BASE CONTACTABILIDAD
10/07/2021 02:30:00	SAPLBFP43	02 PROCESO BASE CONTACTABILIDAD
9/07/2021 09:42:59	SAPLBFP43	02 PROCESO BASE CONTACTABILIDAD
9/07/2021 02:30:00	SAPLBFP43	02 PROCESO BASE CONTACTABILIDAD
8/07/2021 09:30:59	SAPLBFP43	02 PROCESO BASE CONTACTABILIDAD
8/07/2021 02:30:00	SAPLBFP43	02 PROCESO BASE CONTACTABILIDAD
7/07/2021 02:30:00	SAPLBFP43	02 PROCESO BASE CONTACTABILIDAD
6/07/2021 02:30:00	SAPLBFP43	02 PROCESO BASE CONTACTABILIDAD
5/07/2021 02:30:00	SAPLBFP43	02 PROCESO BASE CONTACTABILIDAD
4/07/2021 01:40:00	SAPLBFP43	02 PROCESO BASE CONTACTABILIDAD
3/07/2021 01:40:00	SAPLBFP43	02 PROCESO BASE CONTACTABILIDAD
2/07/2021 01:40:00	SAPLBFP43	02 PROCESO BASE CONTACTABILIDAD
1/07/2021 10:23:23	SAPLBFP43	02 PROCESO BASE CONTACTABILIDAD
1/07/2021 01:40:00	SAPLBFP43	02 PROCESO BASE CONTACTABILIDAD
30/06/2021 01:40:00	SAPLBFP43	02 PROCESO BASE CONTACTABILIDAD
29/06/2021 01:40:00	SAPLBFP43	02 PROCESO BASE CONTACTABILIDAD

La ejecución del proceso BUC es en horario fuera de oficina, pero cuando se produce una caída, este debe ejecutar en horario de oficina, afectando aún más el tiempo de ejecución, debido a que el servidor está siendo usado para otros procesos y existe más congestión de usuarios.

El no tener lista la BUC, produce demoras en el despliegue de campañas y otros requerimientos.

Figura 2 Ubicación del proceso de BUC, configurado en el Agente del Servidor 43



Fuente: ScreenShot

Figura 3 Programación de la ejecución del JOB a las 2:30 de forma diaria



Fuente: ScreenShot

Figura 4 Reporte exitoso de la ejecución del proceso BUC

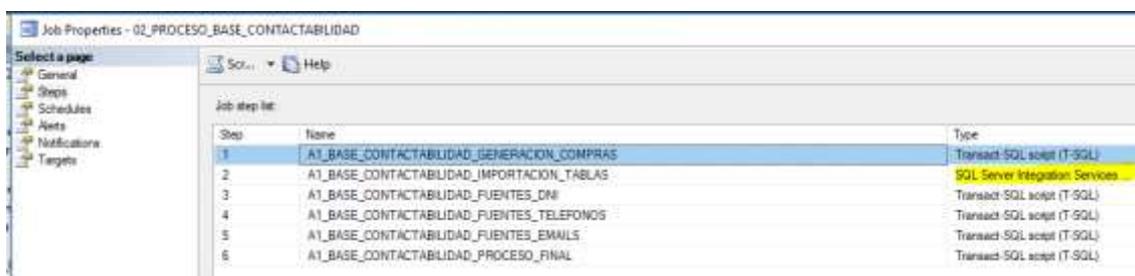
29/06/2021 01:40:00	00:33:53	43		
✓ 29/06/2021 02:10:41	00:03:12	43	A1_BASE_CONTACTABILIDAD_PROCESO_FINAL	5
✓ 29/06/2021 02:06:53	00:03:48	43	A1_BASE_CONTACTABILIDAD_FUENTES_EMAILS	5
✓ 29/06/2021 01:53:38	00:13:15	43	A1_BASE_CONTACTABILIDAD_FUENTES_TELEFONOS	4
✓ 29/06/2021 01:49:56	00:03:42	43	A1_BASE_CONTACTABILIDAD_FUENTES_DNI	3
✓ 29/06/2021 01:40:00	00:09:56	43	NUEVOBUC_DIARIO	2
✓ 29/06/2021 01:40:00	00:00:00	43	A_GENERACION_COMPRAS	1

Fuente: ScreenShot

Descripción del Proceso Actual

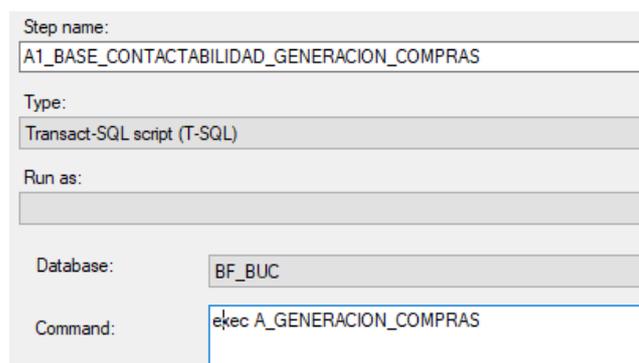
El proceso actual es ejecutado por el Agente SQL a través de un JOB en la Figura 5 , detalla el paso a paso.

Figura 5 Detalle de pasos que ejecuta el JOB en SQL Server 43



Paso 1: A1_BASE_CONTACTABILIDAD_GENERACION DE COMPRAS

Figura 6 Generación de base compras

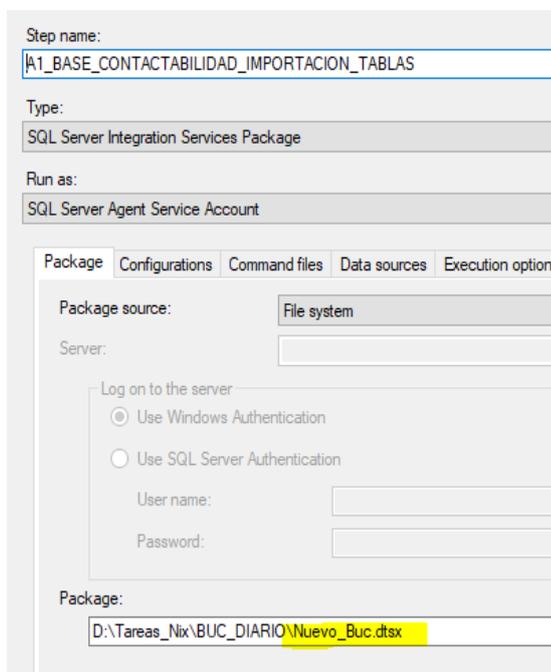


El paso 1 ejecuta un Store Procedure (SP) llamado A_GENERACION_COMPRAS el cual genera una base consolidada de las tablas de compras.

Ese procedimiento en la actualidad se encuentra **desactivado**, es solo ejecutado a demanda, cuando existe una compra de datos de contactabilidad.

Paso 2: A1_BASE_CONTACTABILIDAD_IMPORTACION_TABLAS

Figura 7 Configuración de ejecución de ETL importación de tablas



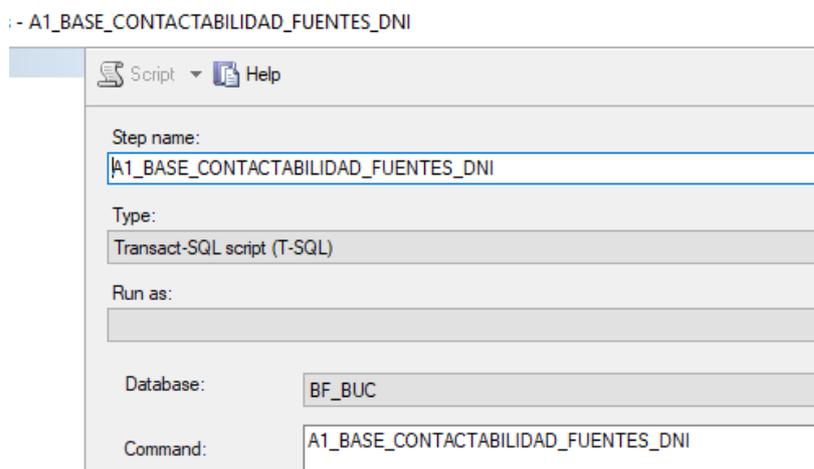
El paso 2 ejecuta un paquete de Integration Services, el cual se encarga de la limpieza inicial y de la importación de las diferentes tablas que serán inputs para la generación de la BUC

Figura 8 Detalle de paquete de Integration Services



Paso 3: A1_BASE_CONTACTABILIDAD_FUENTES_DNI

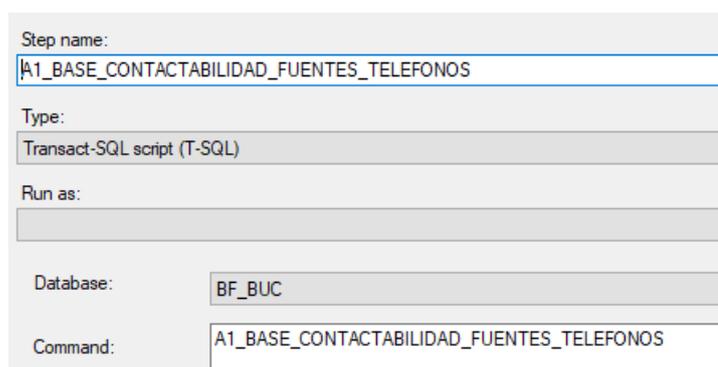
Figura 9 Detalle de paso: Generación de DNI



El paso 3, ejecuta un Store Procedure(SP) de la base de datos BF_BUC, este procedimiento, se encarga de la consolidación de números de documento de identidad, en este caso se están filtrando documentos DNI y RUC. Los demás tipos de documentos no se están considerando.

Paso 4: A1_BASE_CONTACTABILIDAD_FUENTES_TELEFONOS

Figura 10 Detalle de Paso: Generación de Teléfonos



El paso 4 ejecuta un Store Procedure(SP) de la base de datos BF_BUC, este procedimiento, se encarga de la consolidación de teléfonos, además de las validaciones y de cruzar con las tablas de BlackList.

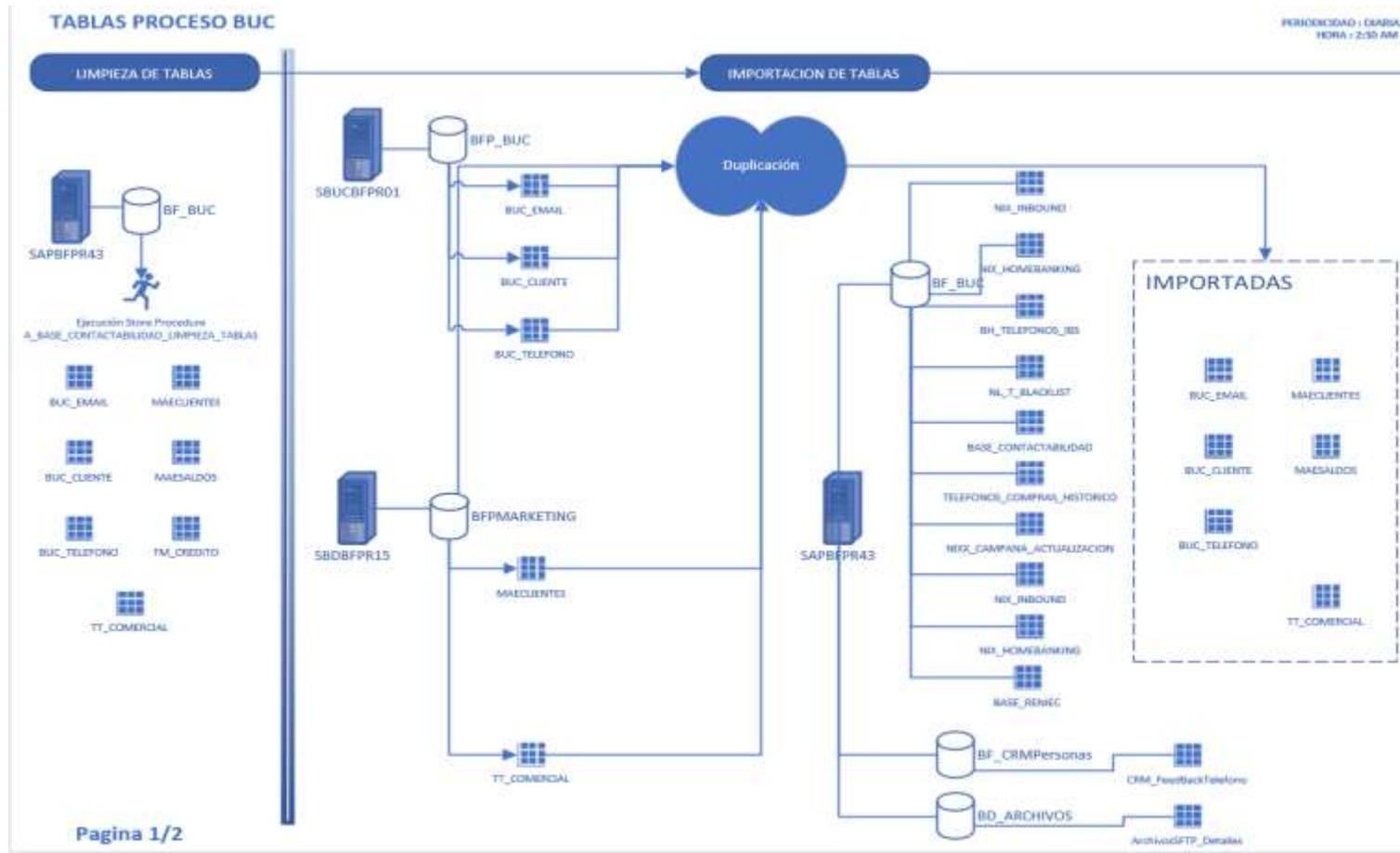
Paso 5: A1_BASE_CONTACTABILIDAD_FUENTES_EMAIL

Figura 11 Detalle de paso Generación de Emails

Step name:	A1_BASE_CONTACTABILIDAD_FUENTES_EMAILS
Type:	Transact-SQL script (T-SQL)
Run as:	
Database:	BF_BUC
Command:	A1_BASE_CONTACTABILIDAD_FUENTES_EMAILS

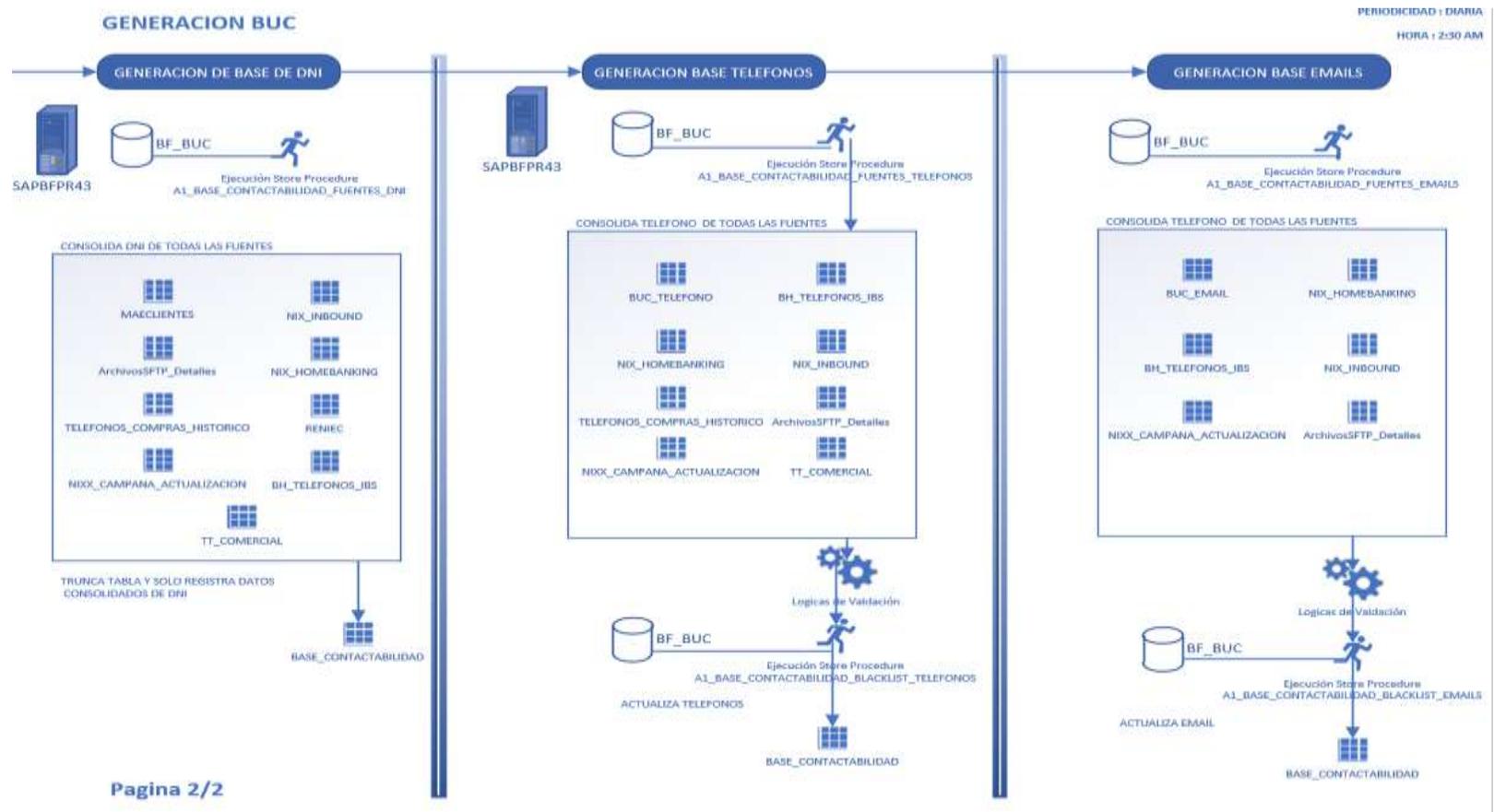
El paso 5 ejecuta un Store Procedure(SP) de la base de datos BF_BUC, este procedimiento , se encarga de la consolidación de emails, además de las validaciones y de cruzar con las tablas de BlackList. En resumen, el proceso de contactabilidad tiene el flujo representado en la figura 12 y 13

Figura 12 Representación del proceso actual BUC 1/2



Fuente: Elaboración propia

Figura 13 Representación del proceso actual BUC 2/2



Fuente: Elaboración propia

DESCRIPCIÓN DE TIEMPO Y VOLUMETRIA DE DATA DE CADA PASO EN EL PROCESO BUC

Tabla 2 Detalle de tiempo y volumetría en cada Paso

NRO	DESCRIPCION	VOLUMETRIA	TIEMPO
PASO 1	GENERACIÓN DE COMPRAS	1.5MM	15 minutos (a demanda)
PASO 2	IMPORTACIÓN DE TABLAS	8MM	20 minutos
PASO 3	GENERACIÓN DE DNI	7 MM	20 minutos
PASO 4	GENERACIÓN DE TELÉFONOS	17 MM	35 minutos
PASO 5	GENERACIÓN DE EMAILS	1,6 MM	15 minutos
TOTAL		35.1 MM	105 minutos

DESCRIPCIÓN DE SERVIDOR 43

El servidor 43 donde se ejecutan los procesos, es un servidor virtual que tiene las siguientes características de Software y Hardware:

Tabla 3 Especificaciones técnicas del de Servidor 43

DESCRIPCION	VALOR
Procesador	Intel (R) Xeon(R) Silver 4215 CPU @2.5GHz 2.49 GHz (2 procesadores)
RAM	16 GB
SO	Windows Server 2012 r2
BD	Microsoft SQL Server 2014
DISCO	80 GB libres

II. OBJETIVOS

OBJETIVO GENERAL

- Optimizar el proceso de generación de bases de contactabilidad de clientes del Banco Pichincha

OBJETIVOS ESPECIFICOS

- Reducir el tiempo de ejecución del proceso de generación de datos de contactabilidad de clientes en Banco Pichincha
- Reducir el consumo de recursos de infraestructura como procesador, memoria y espacio en disco, utilizados para la generación de datos de contactabilidad de clientes en Banco Pichincha
- Mejorar la calidad de datos de contacto de teléfonos en relación a la limpieza lógica, titularidad, procesos masivos e identificación de tipo de documento

III. Limitaciones de la investigación

Debido a la data sensible del proceso que se tomó como estudio, el presente informe detallará de manera general algunos cambios propuestos.

No se mostrará detalle de datos, estructuras, lógicas, rutas y configuraciones que pongan en riesgo la seguridad de la información del Banco Pichincha

IV. JUSTIFICACIÓN DEL INFORME

Gracias a la formación de la carrera de Ingeniería de Sistemas, cuento con la capacidad para identificar, analizar, proponer e implementar soluciones y mejoras en los procesos de la organización en donde me encuentre participando.

El proceso de Consolidación de Base de Contactabilidad, es uno de los más críticos para el despliegue de campañas, ya que permite generar los datos de contacto de los clientes que tienen una oferta.

Si no se contara con esta base, el banco tendría que tercerizar el 100% de sus campañas, esto produciría un aumento de costos por la colocación y una disminución de utilidad.

Se han reportado caídas del proceso de la generación de esta base en los últimos meses mayo, junio y julio del 2021, los reportes nos llevan a realizar una revisión de todo el proceso y plantear soluciones que aseguren que se ejecute forma eficiente y la base esté disponible oportunamente.

Por tanto, las mejoras planteadas e implementadas ayudarán a tener oportunamente la base y poder desplegar las campañas sin causar cuellos de botella por el área de CRM.

V. FUNDAMENTO TEÓRICO

1. Mejora Cotinua

Según: (Universidad de los Llanos, 2021)

El concepto de mejora continua (método Kaizen) significa que siempre estamos en proceso con la oportunidad de cambiar, desarrollar y mejorar. En este sentido, los esfuerzos de mejora continua son un ciclo continuo de identificar áreas de mejora, planificar formas de lograr la mejora, implementar mejoras, identificar, validar resultados y tomar medidas para corregir desviaciones o hacer recomendaciones. una nueva meta

2. Ciclo Deming o ciclo PHVA

Según: (ISOTools, 2021)

Una de las principales herramientas para la Mejora Continua en las organizaciones es el conocido Ciclo Deming o también nombrado ciclo PHVA (Planear – Hacer – Verificar – Actuar).

- **Planificar.** En primer lugar, se deben identificar las actividades que es probable que se mejoren y también se deben establecer los objetivos a alcanzar.
- Para ello, se recopila toda la información necesaria para analizar la causa raíz de esta situación adversa con el fin de elaborar un plan de acción, con acciones preventivas y correctivas que se estimen necesarias para su recuperación. Esta fase incluye la formación del personal para poder aplicar y comprender las medidas identificadas.
- **Hacer.** Incluye la realización de todas las acciones que se planificaron en la fase anterior. Durante esta fase, se deben realizar pruebas piloto para confirmar el funcionamiento de las acciones antes de emprender modificaciones a gran escala.
- **Verificar.** La efectividad de las acciones tomadas debe ser evaluada. Si las actividades no logran el efecto deseado, será necesario modificarlas para que se logren las metas establecidas.
- **Actuar.** Una vez que se completa el proceso, los resultados obtenidos deben estudiarse y compararse con los datos ya recopilados antes de implementar acciones de mejora. Además, cada acción tomada debe ser registrada.

3. SQL Server

Según: (GABILLAUD, 2015)

- SQL Server es un RDBMS (sistema de administración de bases de datos relacionales) que está completamente integrado en Windows, lo que permite múltiples niveles de simplificación de la administración y brinda capacidades máximas.

Según: (GABILLAUD, 2008)

Microsoft Transact-SQL es un lenguaje de consulta mejorado sobre SQL, basado en él. SQL (Structured Query Language) es un lenguaje estándar, creado por IBM en la década de 1970 para administrar sistemas RGBMS.

Tres tipos de comandos componen este lenguaje:

- Lenguaje de definición de datos (Data Description Language - DDL), que permite la creación, modificación y eliminación de objetos SQL (TABLAS, ÍNDICE, VISTAS, PROCEDIMIENTOS, etc.)
- Manipulación de datos El lenguaje (DML) proporciona instrucciones para crear, actualizar, eliminar y recuperar datos de la memoria
- Lenguaje de control de datos (DCL) para administrar el acceso a datos, transacciones y configuración de sesiones y bases de datos.

4. ETL

Según: (HASELDEN, 2006)

- ETL son las siglas (en inglés) de (Extraer, Transformar y Cargar) que en español sería Extraer, Transformar y Cargar; que describe los procesos que tienen lugar en un entorno de almacenamiento de datos para extraer datos de los sistemas de transacción de origen, transformar, limpiar, unificar y ajustar los datos, y finalmente cargarlos en bloques u otros destinos analíticos

5. JOB en SQL

Según: (VIEIRA, 2007)

Los (JOBS) son una colección de una o más actividades. Una actividad es una unidad de acción lógica, como la copia de seguridad de una base de datos o la ejecución de un script T-SQL para completar una acción específica, como la reconstrucción de un índice completo.

6. Procedimiento Almacenado (Store Procedure)

Según: (QUINTO, 2021)

Un procedimiento almacenado es un conjunto de instrucciones T-SQL que SQL Server compila en un solo plan de ejecución. El llamado "procedimiento almacenado" se almacena en la base de datos y se puede ejecutar en cualquier momento.

7. Diagrama Flujo

Según: (Iván TORRES, 2021)

Un diagrama de flujo es una representación gráfica de un proceso. Se trata de mostrar los pasos que toma un proceso de principio a fin, utilizando una serie de elementos visuales para ayudar a representar cada paso que toma el proceso.

8. Tablas en SQL

Según: (Microsoft, 2021)

Una tabla es un objeto de base de datos que contiene todos sus datos. Una tabla organiza los datos en filas y columnas, de forma similar a una hoja de cálculo. Cada fila representa un registro individual y cada columna representa un campo dentro del registro. Por ejemplo, una tabla que contiene datos sobre los automóviles de una empresa, puede tener una fila por automóvil y varias columnas con detalles sobre el automóvil, como: Por ejemplo, número de placa,

chofer asignado, seguro, fecha de mantenimiento, etc., y número de teléfono chofer.

9. Indexación de Columnas de Tabla

Según: (IBM, 2018)

Utilice índices para mejorar el rendimiento de la base de datos.

Un índice bien diseñado puede reducir o eliminar la necesidad de realizar exploraciones de tablas completas durante la ejecución de consultas SQL y acelerar la recuperación de datos.

VI. GLOSARIO DE TERMINOS

BUC	Base única de contactabilidad, es la denominación que tiene la tabla resultante del caso en estudio. Esta base consolida toda la información de contacto de los clientes y es usada para enriquecer a las campañas para poder desplegarlas.
Campaña	Es un conjunto de ofertas de productos crediticios que genera en el área de RIESGOS o PRODUCTO de Banco Pichincha, de manera periódica, en el caso de estudio es mensual.
Refactorizar	Es una técnica que consiste en simplificar el código sin alterar la funcionalidad.
BlackList	Es una tabla que consolida el consentimiento de los clientes para el uso de sus datos personales.
PK	Siglas en ingles Primary Key, es el atributo identificador único en una tabla.
Datos de contactabilidad	Para el caso de estudio, serán 3 los tipos de datos: Teléfonos, Email y Nros de Documentos
Query	Un 'query' es un lenguaje estándar que le permite recuperar datos de una o más tablas, actualizar contenido y eliminar contenido de una manera rápida y potente. <u>comúnmente</u> conocido como el lenguaje SQL
Horario Fuera de Oficina	Es el horario que comprende desde 18:30 hasta las 9:00 horas del día siguiente, este horario es oportuno para realizar tareas de procesamiento ya que hay menor cantidad de usuarios en las plataformas que podrían verse afectados con la ejecución de procesos de gran volumen de información

VII. DESARROLLO DE LA SOLUCIÓN

1. DEFINICION DE METODOLOGIA

Se utilizará el Ciclo de Deming, que es una metodología que permite resolver problemas utilizando al aprendizaje de forma continua, esta herramienta tiene cuatro pasos PHVA Plan, Hacer, Verificar y Actuar.

2. APLICACIÓN DE METODOLOGIA

A. PLAN

Objetivos:

- ✓ Identificar las causas de las caídas del proceso.
- ✓ Optimizaciones de flujo de ejecución
- ✓ Optimizaciones de ETL de importación de tablas
- ✓ Optimizaciones de Scripts SQL
- ✓ Optimizaciones adicionales.
- ✓ Realizar comparativos de las mejoras versus el proceso actual basados en tiempos, volumen de información y recursos que consume

B. HACER

- ✓ **IDENTIFICAR LAS CAUSAS DE LAS CAÍDAS DEL PROCESO:**

Según el reporte de ejecución, identificamos que las caídas son en distintos pasos, pero por la misma razón, el archivo log del backup está sobre pasando el límite máximo de tamaño.

Tabla 4 Detalle de causas de caída en del proceso base de contactabilidad

Fecha	Nro Paso	Descripción de caída
04/07/2021	Paso 6	Caída por Log Full
07/07/2021	Paso 5	Caída por Log Full
08/07/2021	Paso 4	Caída por Log Full
09/07/2021	Paso 3	Caída por Log Full
14/07/2021	Paso 2	Caída por Log Full

Figura 14 Caída del proceso en el paso 6

4/07/2021 01:40:00		SAPLBFPR43	02 PROCESO BASE CONTACTA
4/07/2021 02:0...	6	SAPLBFPR43	02 PROCESO BASE CONTACTA
4/07/2021 02:0...	5	SAPLBFPR43	02 PROCESO BASE CONTACTA
4/07/2021 01:5...	4	SAPLBFPR43	02 PROCESO BASE CONTACTA
4/07/2021 01:4...	3	SAPLBFPR43	02 PROCESO BASE CONTACTA
4/07/2021 01:4...	2	SAPLBFPR43	02 PROCESO BASE CONTACTA
4/07/2021 01:4...	1	SAPLBFPR43	02 PROCESO BASE CONTACTA

Selected row details:

Sql Severity 17
 Sql Message ID 9002
 Operator Emailed
 Operator Net sent
 Operator Paged
 Retries Attempted 0

Message
 Executed as user: FINANCIERO\USERCRM. The transaction log for database 'BF_CRMPersonas' is full due to 'LOG_BACKUP'. [SQLSTATE 42000] (Error 9002). The step failed.

Figura 15 Caída del proceso en el paso 5

7/07/2021 02:30:00		SAPLBFPR43	02 PROCESO BASE CONTACTA
7/07/2021 02:5...	5	SAPLBFPR43	02 PROCESO BASE CONTACTA
7/07/2021 02:4...	4	SAPLBFPR43	02 PROCESO BASE CONTACTA
7/07/2021 02:3...	3	SAPLBFPR43	02 PROCESO BASE CONTACTA
7/07/2021 02:3...	2	SAPLBFPR43	02 PROCESO BASE CONTACTA
7/07/2021 02:3...	1	SAPLBFPR43	02 PROCESO BASE CONTACTA
6/07/2021 02:30:00		SAPLBFPR43	02 PROCESO BASE CONTACTA
5/07/2021 02:30:00		SAPLBFPR43	02 PROCESO BASE CONTACTA
4/07/2021 01:40:00		SAPLBFPR43	02 PROCESO BASE CONTACTA

Selected row details:

Log Job History (02_PROCESO_BASE_CONTACTABILIDAD)

Step ID 5
 Server SAPLBFPR43
 Job Name 02_PROCESO_BASE_CONTACTABILIDAD
 Step Name A1_BASE_CONTACTABILIDAD_FUENTES_EMAILS
 Duration 00:03:53
 Sql Severity 17
 Sql Message ID 9002
 Operator Emailed
 Operator Net sent
 Operator Paged
 Retries Attempted 0

Message
 Executed as user: FINANCIERO\USERCRM. The transaction log for database 'BF_BUC' is full due to 'LOG_BACKUP'. [SQLSTATE 42000] (Error 9002). The step failed.

Figura 16 Caída del proceso en el paso 4

Icon	Date	Step ID	Server	Job Name
✖	8/07/2021 02:30:00		SAPLBFPR43	02_PROCESO_BASE CONTACTABILIDAD
✖	8/07/2021 02:42:47	4	SAPLBFPR43	02_PROCESO_BASE CONTACTABILIDAD
✔	8/07/2021 02:30:00	3	SAPLBFPR43	02_PROCESO_BASE CONTACTABILIDAD
✔	8/07/2021 02:30:00	2	SAPLBFPR43	02_PROCESO_BASE CONTACTABILIDAD
✔	8/07/2021 02:30:00	1	SAPLBFPR43	02_PROCESO_BASE CONTACTABILIDAD

Selected row details:

Date: 8/07/2021 02:42:47
 Log: Job History (02_PROCESO_BASE_CONTACTABILIDAD)

Step ID: 4
 Server: SAPLBFPR43
 Job Name: 02_PROCESO_BASE_CONTACTABILIDAD
 Step Name: A1_BASE_CONTACTABILIDAD_FUENTES_TELEFONOS
 Duration: 00:12:57
 Sql Severity: 17
 Sql Message ID: 9002
 Operator Emailed:
 Operator Net sent:
 Operator Paged:
 Retries Attempted: 0

Message
 Executed as user: FINANCIERO\USERCRM. primera depuracion [SQLSTATE 01000] (Message 0)
 primera eliminacion blacklist [SQLSTATE 01000] (Message 0) The transaction log for database 'BF_BUC' is full due to 'LOG_BACKUP'. [SQLSTATE 42000] (Error 9002). The step failed.

Figura 17 Caída del proceso en el paso 2

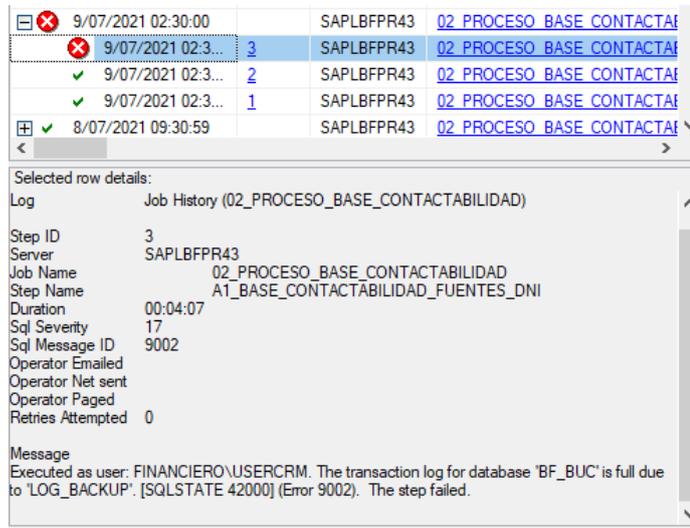
Icon	Date	Step ID	Server	Job Name
✖	14/07/2021 02:30:00		SAPLBFPR43	02_PROCESO_BASE CONTACTABILIDAD
✖	14/07/2021 02:38:18	2	SAPLBFPR43	02_PROCESO_BASE CONTACTABILIDAD
✔	14/07/2021 02:30:00	1	SAPLBFPR43	02_PROCESO_BASE CONTACTABILIDAD
✔	13/07/2021 09:13:16		SAPLBFPR43	02_PROCESO_BASE CONTACTABILIDAD

Selected row details:

Server: SAPLBFPR43
 Job Name: 02_PROCESO_BASE_CONTACTABILIDAD
 Step Name: A1_BASE_CONTACTABILIDAD_NUEVOBUC_DIARIO
 Duration: 00:08:18
 Sql Severity: 0
 Sql Message ID: 0
 Operator Emailed:
 Operator Net sent:
 Operator Paged:
 Retries Attempted: 0

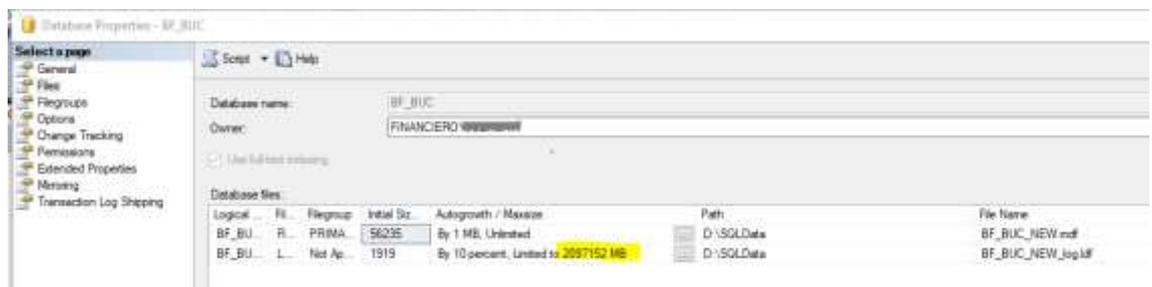
Message
 Executed as user: FINANCIERO\usercrm. Microsoft (R) SQL Server Execute Package Utility
 Version 12.0.6433.1 for 32-bit Copyright (C) Microsoft Corporation. All rights reserved. Started:
 02:30:00 a.m. Error: 2021-07-14 02:38:18.14 Code: 0xC0202009 Source: IMPORTACION
 MAECLIENTES NL_MAECLIENTES [2] Description: SSIS Error Code DTS_E_OLEDBERROR.
 An OLE DB error has occurred. Error code: 0x80004005. An OLE DB record is available. Source:
 "Microsoft SQL Server Native Client 11.0" Hresult: 0x80004005 Description: "The transaction log
 for database 'BF_BUC' is full due to 'LOG_BACKUP'." End Error Error: 2021-07-14 02:38:18.16
 Code: 0xC0209029 Source: IMPORTACION MAECLIENTES NL_MAECLIENTES [2]
 Description: SSIS Error Code DTS_E_INDUCEDTRANSFORMFAILUREONERROR. The
 "NL_MAECLIENTES.Inputs[OLE DB Destination Input]" failed because error code 0xC020907B
 occurred, and the error row disposition on "NL_MAECLIENTES.Inputs[OLE DB Destination Input]"
 specifies failure on error. An error occurred on the specified object of the specified component.
 There may be error messages posted before this with more information about the failure. End Error
 Error: 2021-07-14 02:38:18.16 Code: 0xC0047022 Source: IMPORTACION MAECLIENTES

Figura 18 Caída del proceso en el paso 3



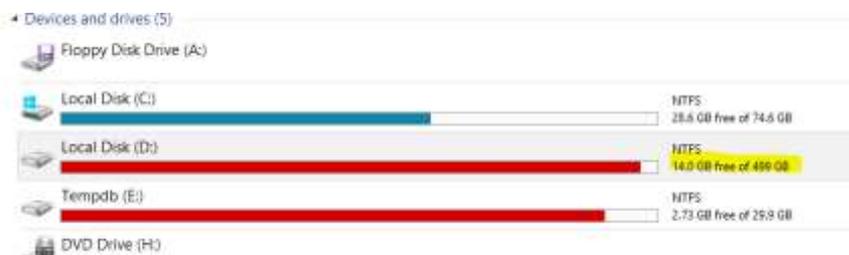
La configuración actual del límite máximo de tamaño del archivo log es 20GB:

Figura 19 Configuración de archivo log



El actual espacio disponible es de solo 14GB.

Figura 20 Detalle de espacio en Servidor 43



Se investigó si otros procesos se han caído y se notó que no es así debido a la ejecución del siguiente proceso se ejecuta a las 5:30 am de limpieza de logs, se libera el espacio de disco y es por eso por lo que aún se tiene espacio disponible para ejecutar otros procesos.

Figura 21 Reporte de ejecución de tareas de limpieza

Log file summary: No filter applied

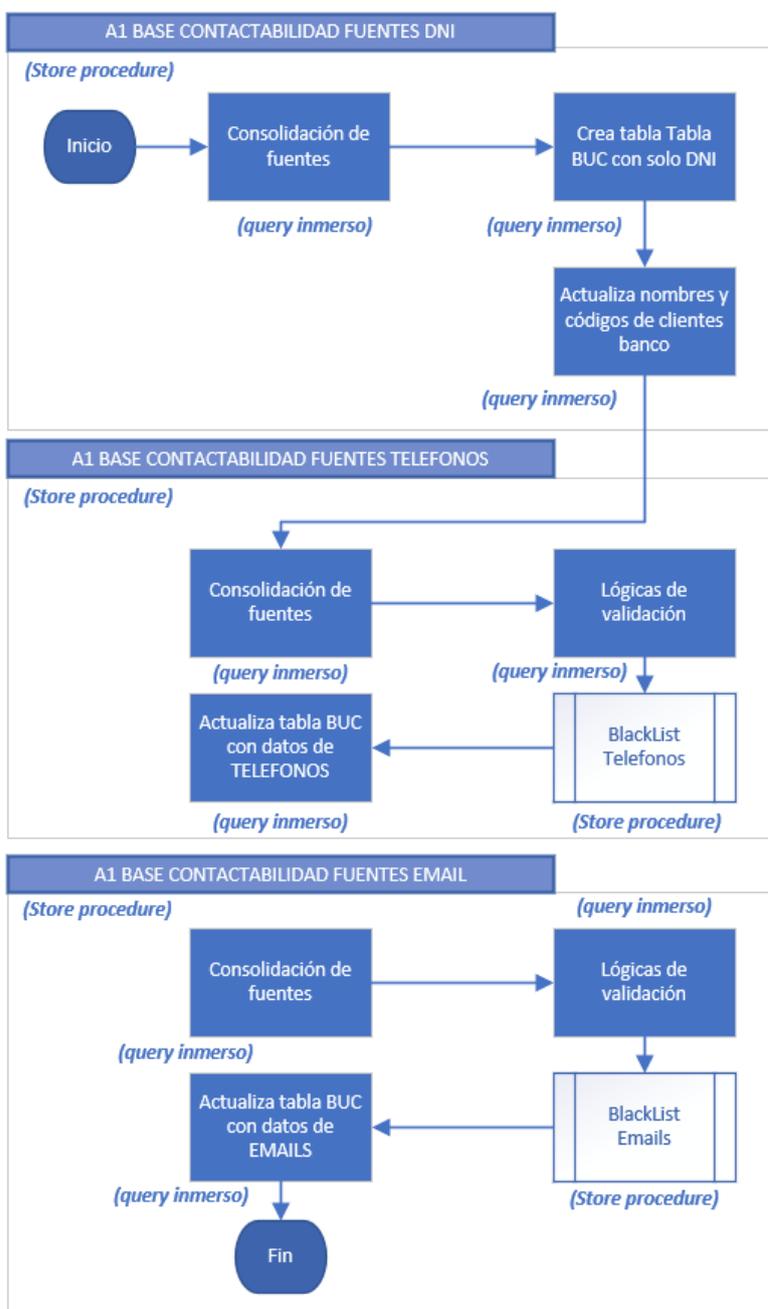
Date	Step ID	Server	Job Name
14/07/2021 05:30:00		SAPLBFPR43	09 LIMPIEZA LOG BASES
13/07/2021 05:30:00		SAPLBFPR43	09 LIMPIEZA LOG BASES
12/07/2021 05:30:00		SAPLBFPR43	09 LIMPIEZA LOG BASES
11/07/2021 05:30:00		SAPLBFPR43	09 LIMPIEZA LOG BASES
10/07/2021 05:30:00		SAPLBFPR43	09 LIMPIEZA LOG BASES
9/07/2021 05:30:00		SAPLBFPR43	09 LIMPIEZA LOG BASES
8/07/2021 05:30:00		SAPLBFPR43	09 LIMPIEZA LOG BASES
7/07/2021 05:30:00		SAPLBFPR43	09 LIMPIEZA LOG BASES
6/07/2021 05:30:00		SAPLBFPR43	09 LIMPIEZA LOG BASES
5/07/2021 05:30:00		SAPLBFPR43	09 LIMPIEZA LOG BASES
4/07/2021 05:30:00		SAPLBFPR43	09 LIMPIEZA LOG BASES

✓ **OPTIMIZACIONES EN FLUJO DE EJECUCIÓN**

Descripción:

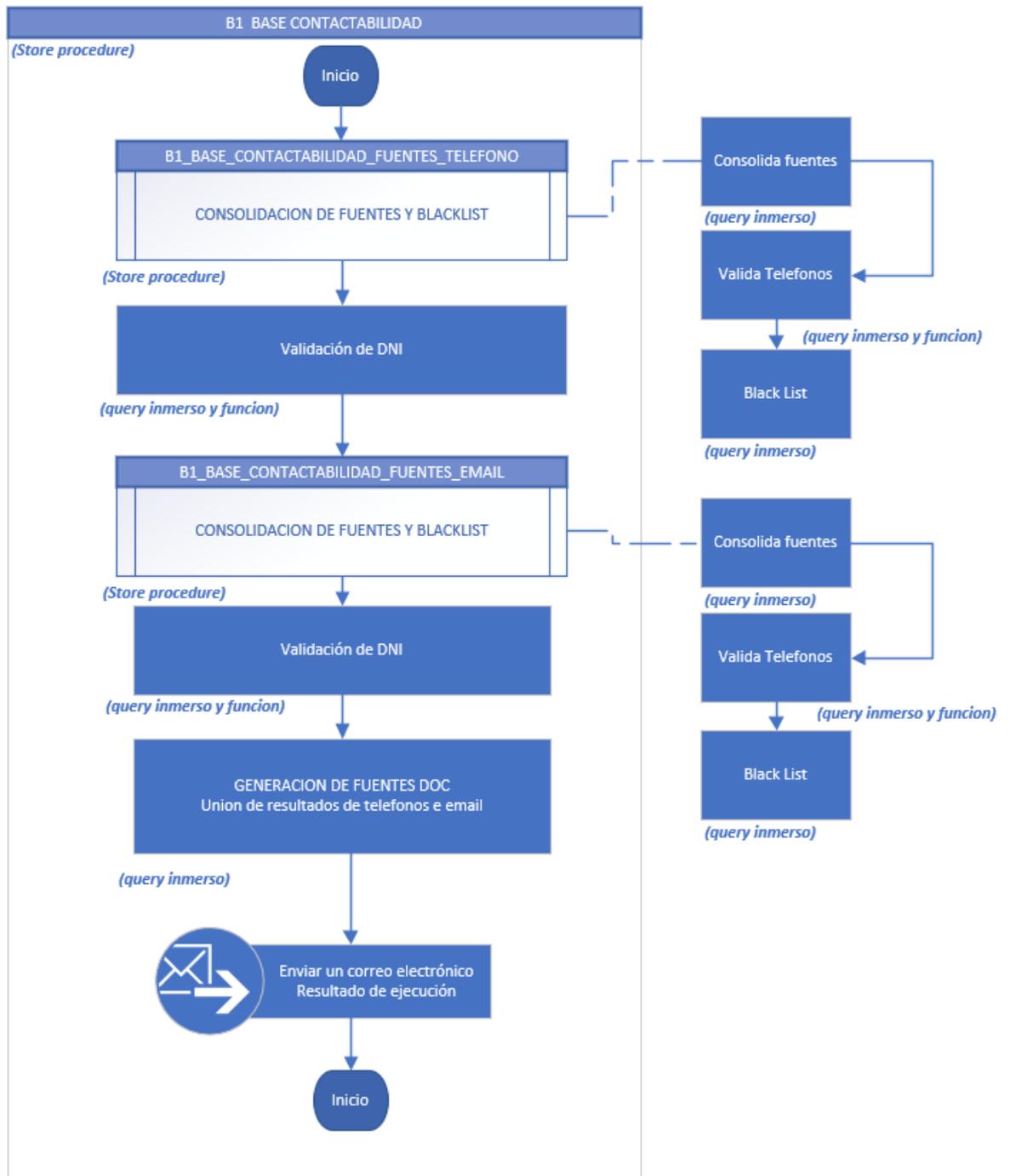
En la Figura 22 realice una simplificación del esquema de ejecución de la generación de la BUC actual.

Figura 22 Diagrama de flujo del escenario actual



Como consecuencia del análisis del proceso se plantea el esquema de ejecución descrito en la Figura 24

Figura 24 Flujo propuesto con optimizaciones

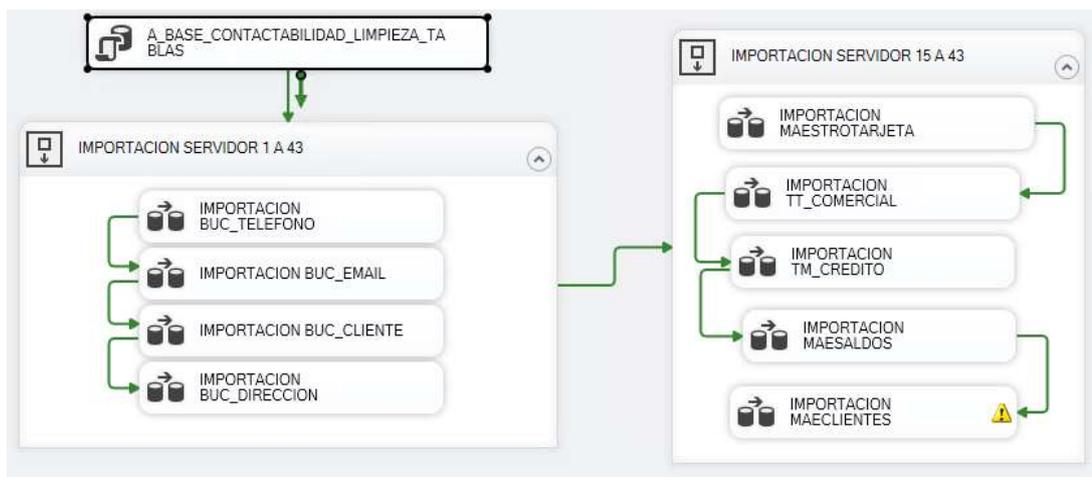


✓ **OPTIMIZACION DE ETL DE IMPORTACION DE TABLAS**

Descripción:

En la Figura 25 es una representación grafica de la secuencia de importación de las tablas que serán inputs para la generación de la BUC

Figura 25 Secuencia de pasos que ejecuta la ETL



Detalle de limpieza de tablas

Figura 26 Query de limpieza de tablas en el Servidor 43

```
BEGIN
truncate table DBO.BUC_CLIENTE
truncate table DBO.BUC_EMAIL
truncate table DBO.BUC_DIRECCION
truncate table DBO.BUC_TELEFONO
truncate table DBO.MAESALDOS
truncate table DBO.MAECLIENTES
truncate table DBO.MAESTROTARJETA
truncate table dbo.TT_COMERCIAL
truncate table dbo.TM_CREDITO
END
```

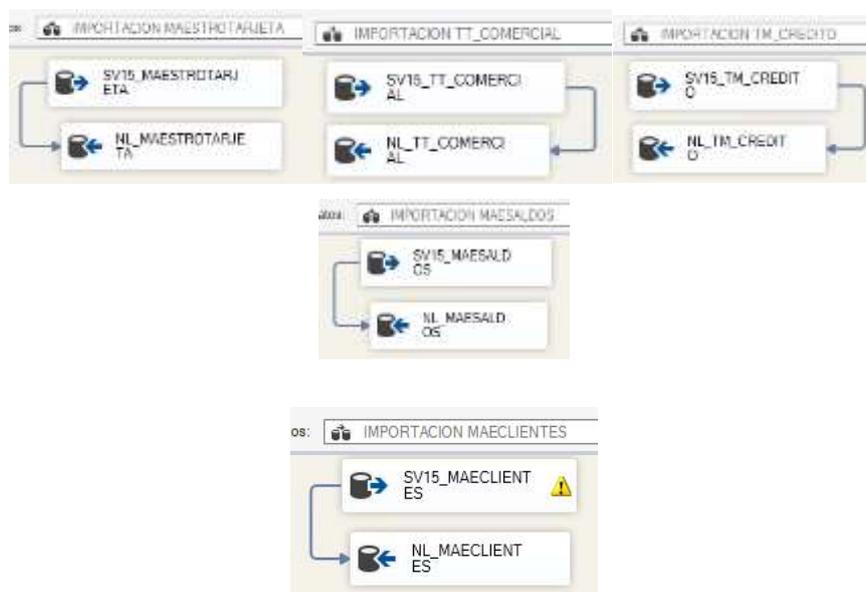
Detalle de flujo de datos de importación servidor 1 a servidor 43

Figura 27 Secuencias de transferencia de tablas del servidor 1 a 43



Detalle de flujo de datos de importación servidor 15 a servidor 43

Figura 28 Secuencias de transferencia de tablas del servidor 15 al 43



Optimizaciones:

Luego de realizar la revisión de la ETL de la importación de tablas, se identificaron las siguientes mejoras:

El realizar la limpieza y la importación de tablas, podría ser reemplazado por un flujo que incluya un control comparativo de bases y este permitirá que solo viaje la información que no se tiene en el servidor 43, es decir ya no realizará la limpieza de tablas, sino solo transportar la información faltante en el 43.

Complementando la mejora se deben crear los índices de cada tabla para mejorar el rendimiento de los comparativos tanto en el servidor de origen como el servidor de destino.

Esto permitirá que se reduzca el consumo de recursos y además asegura que siempre se cuente con información de manera en las tablas del servidor 43.

En la situación actual, si existiera algún fallo de transferencia de data, se truncaría el proceso total y debido a la limpieza de tablas inicial no permitiría tener la información ni de forma parcial.

Mejora de tarea importación tabla BUC TELEFONO

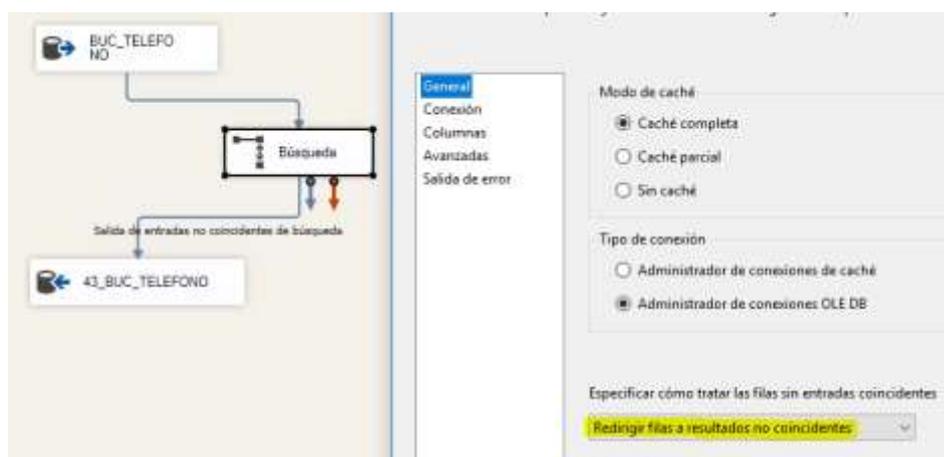
Para agregar un control comparativo agregamos en la lista de herramientas el control búsqueda.

Figura 29 Optimización de importación de tabla BUC_TELEFONO



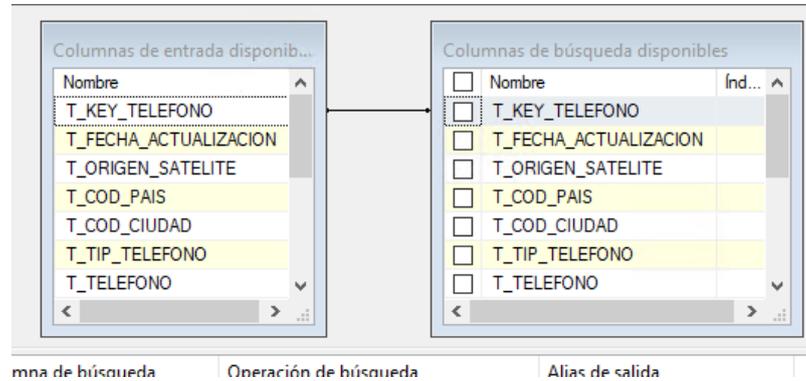
Indicamos que la salida del control solo debe ser aquellos que no arrojan coincidencias.

Figura 30 Configuración de control de búsqueda



Se identifico que el campo T_KEY_TELEFONO es único

Figura 31 Identificación de PK para la importación diferencial BUC_TELEFONO



Se creo índice de tabla BUC_TELEFONO en base al T_KEY_TELEFONO para aumentar el rendimiento de la búsqueda

Figura 32 Creación de índice para BUC_TELEFONO

```
create index ind_BUC_TELEFONO_T_KEY_TELEFONO on BUC_TELEFONO(t_key_telefono)
```

Se ejecuto la tarea y se visualiza el siguiente resultado:

Figura 33 Primera ejecución con mejora de Importación de BUC_TELEFONO



Al implementar la mejora de la tarea, la transferencia de la tabla BUC_TELEFONO, solo realiza la transferencia de 0.7MM registros, de los 9,3MM reduciendo en un 92% de la data a transferir.

En una segunda ejecución, la transferencia fue nula, así nos aseguramos de que solo se transfieran filas únicas.

Figura 34 Segunda ejecución con mejora de Importación de BUC_TELEFONO



Consulta de verificación en Servidor 43, de que solo se han transferido filas únicas

Figura 35 Verificación de importación de BUC_TELEFONO

```
SELECT TOP 100 T_KEY_TELEFONO ,COUNT(*) FROM BUC_TELEFONO
GROUP BY T_KEY_TELEFONO
ORDER BY 2 DESC
```

	T_KEY_TELEFONO	(No column name)
1	722	1
2	720	1
3	713	1
4	711	1
5	706	1
6	697	1
7	695	1
8	688	1
9	686	1

Mejora de tarea: importación tabla BUC_EMAIL

La aplicación es similar al anterior proceso, se agrega un control de búsqueda

Figura 36 Optimización de importación de tabla BUC_EMAIL

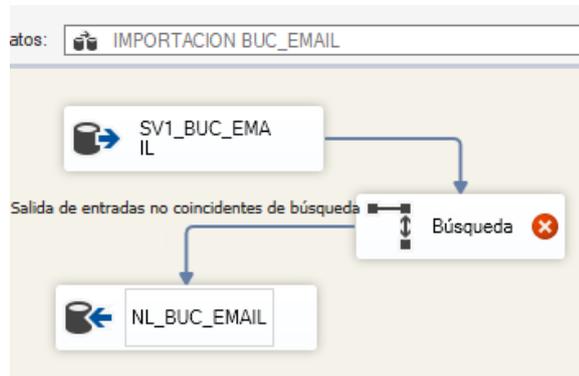


Figura 37 Configuración de control de búsqueda para BUC_EMAIL

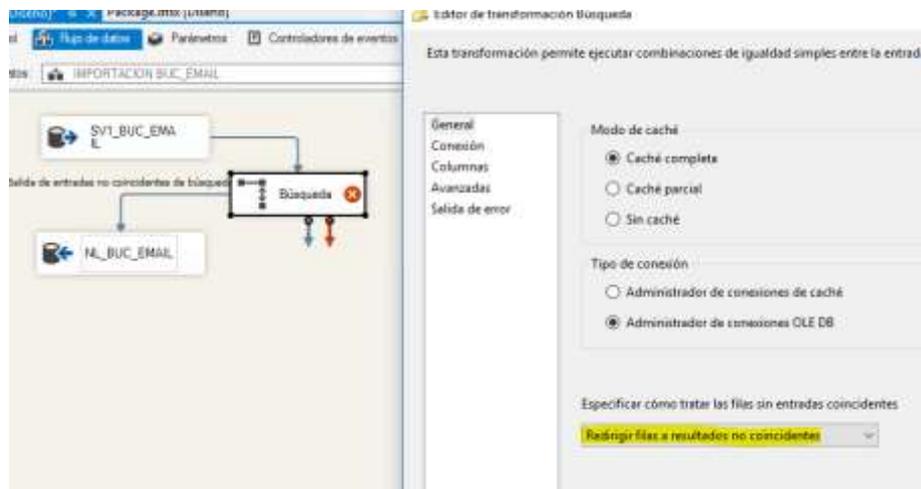
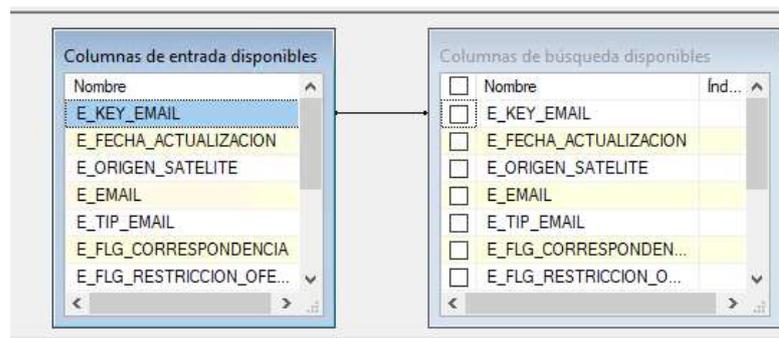


Figura 38 Identificación de PK para la importación diferencial BUC_EMAIL



Se crea el índice correspondiente

Figura 39 Creación de índice BUC_EMAIL

```
CREATE INDEX IND_BUC_EMAIL_E_KEY_EMAIL ON BUC_EMAIL(E_KEY_EMAIL)
```

Se ejecuta la tarea y se obtiene el siguiente resultado:

Figura 40 Primera ejecución con mejora de importación BUC_EMAIL



Figura 41 Segunda ejecución con mejora de importación BUC_EMAIL

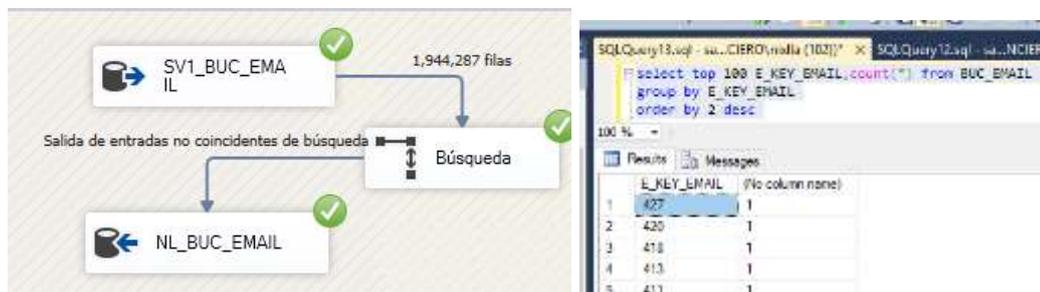
La imagen muestra una ventana de SSIS con la siguiente estructura de árbol de tareas:

- Nuevo_Buc
 - IMPORTACION SERVIDOR 1 A 43
 - Tarea IMPORTACION BUC_EMAIL
 - Inicio, 00:16:13
 - Se ha iniciado la validación
 - [SSIS.Pipeline] Información: Se está iniciando la fase de validación

Al implementar la mejora de la tarea, la transferencia de la tabla BUC_EMAIL, solo realiza la transferencia de 1,340 registros, de los 1,9MM reduciendo en un 100% de la data a transferir tan solo demorando solo 16 segundos.

Se ejecuta por segunda ocasión para verificar la transferencia de filas que sean únicas

Figura 42 Verificación de importación BUC_EMAIL



La transferencia es nula en la segunda ejecución, siendo un éxito la mejora.

Mejora de tarea: importación tabla BUC_CLIENTE

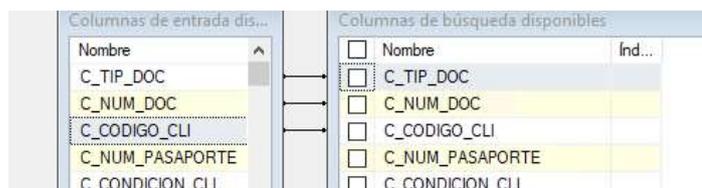
La aplicación es similar al anterior proceso, se agrega un control de búsqueda

Figura 43 Optimización de importación BUC_CLIENTE



En este caso se realizará comparativo de 3 columnas, C_TIP_DOC (tipo de documento), C_NUM_DOC (número de documento), C_CODIGO_CLI (código cliente banco)

Figura 44 Identificación de PK para la importación diferencial BUC_CLIENTE



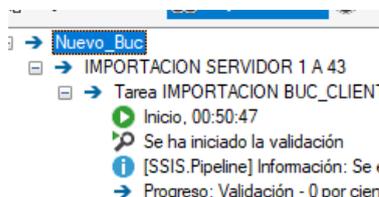
Se crea el índice respectivo en base a las columnas que se realizaran las búsquedas

Figura 45 Creación de índice BUC_CLIENTE

```
CREATE INDEX IND_BUC_CLIENTE_TIP_NUM_COD ON BUC_CLIENTE( C_TIP_DOC,C_NUM_DOC,C_CODIGO_CLI)
```

Se ejecuta la tarea y se obtiene el siguiente resultado:

Figura 46 Ejecución de mejora de importación BUC_CLIENTE



Eliminación de tarea: importación BUC DIRECCION

Debido a que esta tarea no se incluye en la generación de datos de contactabilidad no se debe considerar en el proceso.

Eliminación de tarea: importación MAESTROTARJETA

Debido a que esta tarea no se incluye en la generación de datos de contactabilidad no se debe considerar en el proceso.

Eliminación de tarea: importación TM CREDITO

Debido a que esta tarea no se incluye en la generación de datos de contactabilidad no se debe considerar en el proceso.

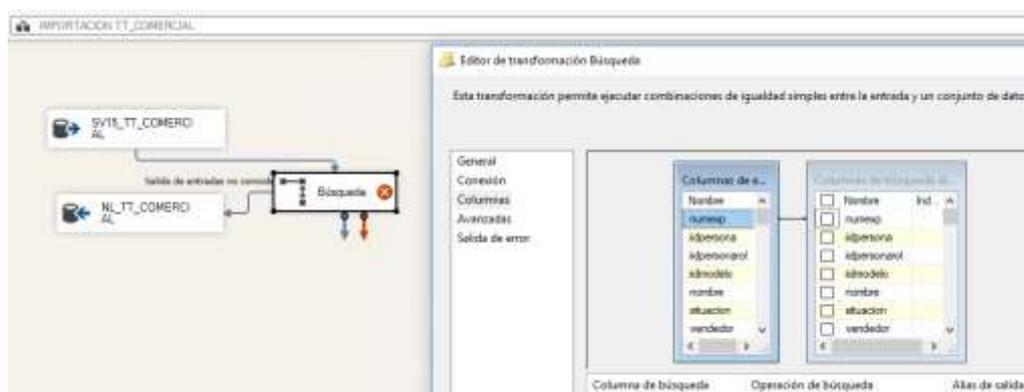
Eliminación de tarea: importación MAESALDOS

Debido a que esta tarea no se incluye en la generación de datos de contactabilidad no se debe considerar en el proceso.

Mejora de tarea: importación tabla TT_COMERCIAL

Se implemento el control de búsqueda realizando el comparativo de numexp(número de expediente) el cual es único.

Figura 47 Optimización de importación TT_COMERCIAL



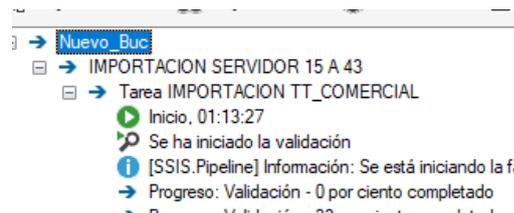
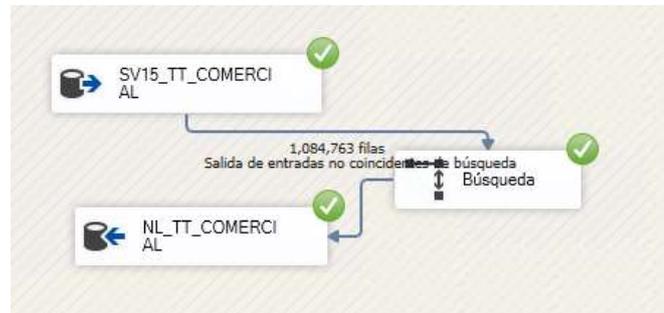
El respectivo índice:

Figura 48 Creación de índice TT_COMERCIAL

```
CREATE INDEX IND_TT_COMERCIAL_NUMEXP ON TT_COMERCIAL(numexp)
```

Se ejecuta la tarea y se obtiene el siguiente resultado:

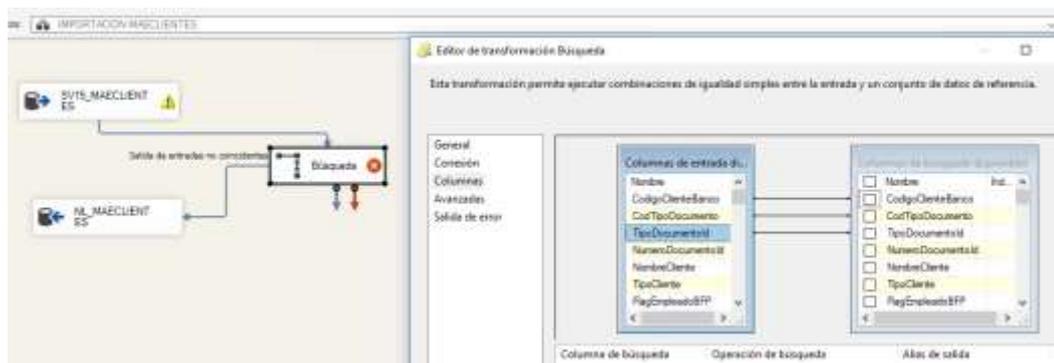
Figura 49 Ejecución de mejora de importación TT_COMERCIAL



Mejora de tarea: importación tabla MAECLIENTES

Se implemento el control de búsqueda realizando el comparativo de CódigoClienteBanco(código cliente banco), CodTipoDocumento(código de tipo de documento) y NumeroDocumentoId(número de documento) el cual es único.

Figura 50 Optimización de importación MAE_CLIENTES



El respectivo índice:

Figura 51 Creación de índice MAE_CLIENTES

```
CREATE INDEX IND_MAECLIENTES_CodTiplumDoc ON MAECLIENTES(CodigoClienteBanco,CodTipoDocumento,NumeroDocumentoId )
```

Se ejecuta la tarea y se obtiene el siguiente resultado:

Figura 52 Ejecuciones de mejora de importación MAE_CLIENTES

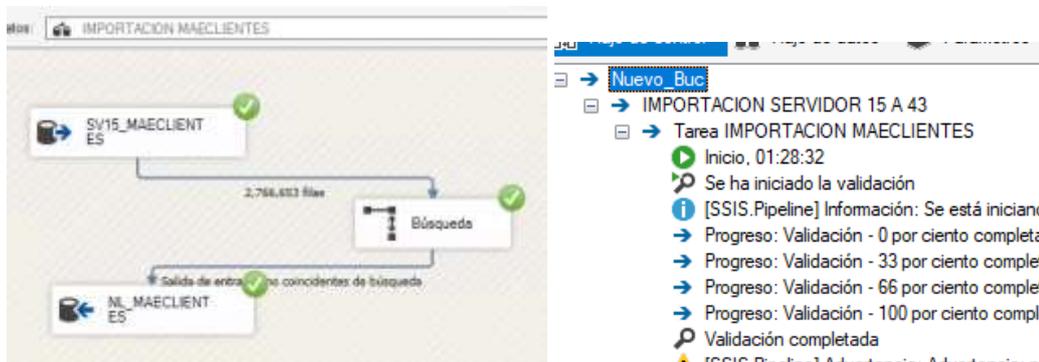


Figura 53 Diagrama simplificado de Importación de Tablas



Debido a los reportes de caídas en los distintos pasos, es necesario agregar la instrucción de limpieza de logs en cada sección.

Agregar instrucción de limpieza de logs de BD SERVIDOR 43

El agregar la instrucción después de cada transferencia, permitirá tener el espacio suficiente para que se ejecute el proceso sin problemas

El siguiente query será ejecutado en el servidor 43 después de cada bloque de importación de tablas

Figura 54 Query de Limpieza de LOG de BF_BUC

```
ALTER DATABASE BF_BUC
SET RECOVERY SIMPLE;
DBCC SHRINKFILE (BF_BUC_NEW_log, 1);
ALTER DATABASE BF_BUC
SET RECOVERY FULL;
```

Diagrama Final Simplificado de Proceso de importacion de tablas

Figura 55 Diagrama final con optimizaciones de importación de tablas



✓ **OPTIMIZACIONES EN SCRIPTS SQL**

Analizando los scripts se visualizan importantes deficiencias en el orden, por ejemplo, bloques de scripts que se repiten, validaciones de datos están inmersas y no consolidadas en bloques de entendimiento lo cual hace difícil el mantenimiento de este.

Analizando la data de teléfonos se encontró una importante casuística con respecto a teléfonos, que en años han pasado a distintos dueños; pero que en la data han permanecido aún vinculados a clientes que actualmente no los poseen. También, teléfonos repetidos en gran número de veces, debido a que son registrados por procesos masivos, como convenios o cuentas CTS y otros.

Solo se están considerando clientes que tienen DNI y RUC, reduciendo el potencial hacia clientes que tienen otros documentos de identidad.

✓ **OPTIMIZACIONES ADICIONALES**

Se planteo una mejora en el espacio del disco con el área de Infraestructura de TI, pero debido a la limitación de recursos, no sería viable hasta el año 2022 con la renovación tecnológica.

Se solicito a las áreas realizar una migración de data histórica a entornos de almacenamiento a largo plazo, dicha migración será una actividad planificada por cada área y se tendrá una mejora en el espacio, esta será una solución parcial al problema. De igual forma es una actividad donde se depende de la realización de otras áreas lo cual por el momento no soluciona de forma inmediata a la disponibilidad de la base de contactabilidad ante las caídas frecuentes.

C. VERIFICAR

Para salir con un nuevo flujo es necesario realizar pilotos, para lo cual se ejecutó el proceso actual y el proceso propuesto para medir varios factores como tiempo, volumetría de datos y utilización de recursos.

COMPARATIVOS DE ESCENARIOS

- Mediante la comparación de las ejecuciones de los 2 escenarios se puede verificar que el escenario propuesto, tiene un tiempo de ejecución 49% menor al del escenario actual.
- Se visualiza significativamente la reducción de data transportada por la ETL importación de tablas de 8 MM a 0.001MM debido a que solo transporta la data que no se tiene en las fuentes de destino, reduciendo el tiempo en un 85%.
- Al cambiar de orden en la ejecución de generación de teléfonos y la de emails para luego generar la de fuentes de DOC, hace que impacte la cantidad de DOC a procesar reduciéndolos de 7MM a 3.2MM, además que solo es un SELECT UNION debido a que en los procesos de generación de teléfonos y email estos DOC ya están validados reduciendo el tiempo de 20 min a 1 min siendo una mejora en un 95%.

Tabla 5 Cuadro comparativo de escenarios: actual vs propuesto

ESCENARIO ACTUAL(EA)			ESCENARIO PROPUESTO(EP)		
PROCESO	DATA	DURACIÓN	PROCESO	DATA	DURACIÓN
GENERACION DE COMPRAS	1.5MM	15 min	GENERACION DE COMPRAS	1.5MM	15 min
EJECUCION ETL IMPORTACIÓN DE TABLAS	8MM	20 min	EJECUCION ETL IMPORTACIÓN DE TABLAS	0,001MM	3 min
GENERACION FUENTES DNI	7 MM	20 min	GENERACION FUENTES TELEFONOS	17MM	20 min
GENERACION FUENTES TELEFONOS	17 MM	35 min	GENERACION DE FUENTES DE EMAIL	1,6MM	15 min
GENERACION DE FUENTES DE EMAIL	1,6 MM	15 min	GENERACION FUENTES DOC	3,2 MM	1 min
TOTAL	35.1 MM	105 min	TOTAL	23,3 MM	54 min

Donde: MM significa millones

COMPARATIVOS DE TABLAS RESULTANTES DE EA Y EP

- Se realizó la comparación de tablas resultantes, en la cual, en el EA, solo se tenía 1 tabla resultante (BASE_CONTACTABILIDAD) la cual en cantidad de filas es de 2,6MM y en cambio en la del EP se tiene 2,9MM, esto quiere decir que aplicando las nuevas lógicas de recuperación de datos, inclusión otros tipo de documentos, limpieza, y ajustes de titularidad se han recuperado 300k, siendo un incremento del 10% de clientes.
- En el EP, se tiene además 3 tablas adicionales que nos permitirán dar tracking de cada registro de contactabilidad.
- En el EA se tiene un total de consumo de espacio 468MB y el EP es de 1,127MB siendo un incremento de espacio en un 58%, esto debido a las 3 tablas adicionales que aportan valor al seguimiento de la data.

Tabla 6 Cuadros comparativos de tablas resultantes

ESCENARIO ACTUAL

Name	rows	data	index_size	columns
BASE_CONTACTABILIDAD	2,662,228	468MB	67MB	39

ESCENARIO PROPUESTO

Name	rows	data	index_size	columns
CONTACTABILIDAD_B	2,960,805	316MB	58MB	24

Name	rows	data	index_size	columns
FUENTES_DOC	3,219,071	116MB	81MB	8

Name	rows	data	index_size	columns
FUENTES_TELEFONO	6,301,573	556MB	230MB	15

Name	Rows	data	index_size	columns
FUENTES_EMAIL	1,658,847	139MB	<u>75MB</u>	10

COMPARATIVOS DE ESCENARIOS DE CANTIDAD DE DATOS RESULTANTES

- Se realizo la comparación de sobre los tipos de datos resultantes, el incremento de teléfonos es el más significativo, siendo de 1,3MM este representa un aumento de contactabilidad para futuras campañas.

Tabla 7 Cuadro comparativo de cantidad de datos en ambos escenarios

DATO	ESCENARIO ACTUAL	ESCENARIO PROPUESTO	INCREMENTO
DOC	2.6MM	3.2MM	0.6MM
TELEFONO	5MM	6.3MM	1.3MM
EMAIL	1.59MM	1.65MM	0.06MM
TOTAL			1.96MM

D. ACTUAR

- a) Después de realizar la ejecución y los comparativos de los escenarios, se mostraron en una presentación los resultados a la Jefatura de CRM, donde se tomó de decisión de actualizar el proceso con la propuesta implementada. (15/07.2021)
- b) Se envió a Gobierno de Datos la documentación, así como los archivos necesarios para realizar los cambios propuestos. (16/07.2021)
- c) Se informó al todo que equipo de CRM para que direccionen las nuevas campañas hacia las nuevas tablas de contactabilidad (16/07/2021)
- d) Gobierno de Datos confirmo el cambio (20/07/2021)

VIII. CONCLUSIONES

- a) Luego de realizar los comparativos del escenario actual y propuesto, se concluye que este último es más eficiente en tiempo de ejecución en un 49% menor, en procesamiento de data se redujo en un 33% y un aumento de 10% de datos de contactabilidad.
- b) Se identificaron de mejoras, en las cuales realizó un análisis a 3 niveles: a nivel de flujo, de sección e instrucciones, realizando mejoras en cada nivel.
- c) Los componentes de la solución son 2, las queries y la ETL, en ambos, se realizó una refactorización agregando secuencias que generan mejor eficiencia el traslado de data y en lógica para recuperación de información de contactabilidad.
- d) La identificación de procesos redundantes fue la razón para cambiar el orden de ejecución de las secciones, debido a esto se creó un nuevo flujo permitiendo reducir significativamente el proceso a un 49% en tiempo de ejecución.
- e) Se agregaron lógicas donde se incluyó todos los tipos de documento, se desestimaron teléfonos por cambio de titularidad y creaciones masivas por cuentas CTS y convenios además se realizó una limpieza previamente antes de descartarlos así aumentando hasta en 300k de nuevo clientes que servirán para las nuevas compañías.

IX. RECOMENDACIONES

- a) Solicitar al área de TI, Implementar alertas en el proceso para monitoreo de espacio disponible y otros recursos como procesador, memoria y concurrencia de usuarios del Servidor 43.
- b) Debido a que el servidor 43, se usa para distintas áreas, se debe realizar una migración de data histórica cada 6 meses, con ello aumentando el espacio para ejecutar los procesos con normalidad.
- c) Seguir documentando los procesos aplicando el Ciclo de Deming cada 6 meses para encontrar oportunidades de mejora.
- d) Gestionar la adquisición de más espacio de almacenamiento
- e) En cuanto existan más presupuesto, se debería realizar la migración hacia la nube que permitan escalar las características según las necesidades del negocio.

X. LISTA DE REFERENCIAS BIBLIOGRÁFICAS

- GABILLAUD, J. (2008). *SQL Server 2008 SQL, Transact SQL*. Barcelona: Ediciones ENI. Obtenido de <https://books.google.com.pe/books?id=vsYTINIEwd4C&pg=PA351&dq=transact+sql&hl=es-419&sa=X&ved=2ahUKEwii8seCh8z0AhXOqpuCHSqCAbwQ6AF6BAgFEAI#v=onepage&q=transact%20sql&f=false>
- GABILLAUD, J. (2015). *SQL SERVER 2014 Administración de la una base de datos transaccional*. Barcelona: EDICIONES ENI. Obtenido de https://books.google.com.pe/books?id=7Ju_kBwnBgUC&printsec=frontcover&dq=sql+server&hl=es-419&sa=X&redir_esc=y&pli=1#v=onepage&q=sql%20server&f=false
- HASELDEN, K. (2006). *Microsoft SQL Server 2005 Integration Services*. Estados Unidos: Sams Publishing. Obtenido de <https://books.google.com.pe/books?id=jcniBu32iu0C&pg=PA4&dq=ETL+sql&hl=es-419&sa=X&ved=2ahUKEwivi5CUicz0AhVXI7kGHAM9AIAQ6AF6BAgKEAI#v=onepage&q=ETL%20sql&f=false>
- IBM. (2018). *Documentación IBM*. Obtenido de Indexación de columnas de tabla: https://www.ibm.com/docs/es/netcoolomnibus/8.1?topic=SSSHTQ_8.1.0/com.ibm.netcool_OMNibus.doc_8.1.0/omnibus/wip/admin/task/omn_adm_workingwithindexes.html
- ISOTools. (2021). *Software Gestión Procesos*. Obtenido de <https://www.isotools.org/soluciones/procesos/mejora-continua/>
- Iván TORRES. (2021). *IVE Consultores*. Obtenido de Diagrama de Flujo: <https://iveconsultores.com/diagrama-de-flujo/>
- Microsoft. (2021). *Documentación de SQL*. Obtenido de <https://docs.microsoft.com/es-es/sql/relational-databases/tables/tables?view=sql-server-ver15>
- QUINTO, C. (2021). *Universidad San Martín de Porres*. Obtenido de Laboratorio de GIS y Tecnologías Emergentes: <https://www.usmp.edu.pe/publicaciones/boletin/fia/info41/procedimiento.html>
- Universidad de los Llanos. (2021). *Sistema Integrado de Gestión*. Obtenido de <https://sig.unillanos.edu.co/index.php/mejora-continua>
- VIEIRA, R. (2007). *Professional SQL Server 2005 Programming*. Indianapolis, Indiana: Wiley Publishing, Inc. doi:<https://books.google.com.pe/books?id=qrk7botldbYC&pg=PA704&dq=jobs+sql&hl=es-419&sa=X&ved=2ahUKEwjJq8iZjMz0AhVI2qQKHSErB4cQ6AF6BAgCEAI#v=onepage&q=jobs%20sql&f=false>

MEMORIA DESCRIPTIVA

DNI:	44699655
EDAD	35 AÑOS
CELULAR	935289563
CORREO	nixonllanos@gmail.com



EXPERIENCIA PROFESIONAL

1. Banco Pichincha del Perú

Cargo: **ESPECIALISTA DE INTELIGENCIA COMERCIAL**

Periodo: 26 de mayo 2017 hasta 20 de abril

Cargo: **ESPECIALISTA DE CRM Y GESTOR DE CAMPAÑAS**

Periodo: 20 de abril 2019 hasta la actualidad

Funciones:

- Generación de Modelos Predictivos con tecnología de Machine Learning con metodología Canvas sobre el proceso de scoring de clientes con ofertas del banco Pichincha.
- Selección de modelo predictivo que se ajusta a las variables en estudio, preprocesamiento de datos y análisis descriptivo.
- Calibración periódica de modelos para mejorar la precisión de la predicción.
- Revisión de resultados de la predicción al término de cada periodo .
- Diseñé, implementé, desplegué y doy mejora continua de la aplicación CRM hecha con tecnología moderna y una experiencia de usuario intuitiva reduciendo los tiempos en los procesos core, generando información de calidad y de interés para la empresa.

Módulos Desarrollados

- Banca Personas
 - Módulo gestión de ofertas PLD, Reenganche, Hipotecario y Convenios

- Módulo de consulta masiva de productos vigentes
 - Módulo de reporte de seguimiento de avance de Campañas.
 - Módulo de firma digital para aceptación de metas
 - Módulo de registro de BlackList de clientes.
 - Módulo de derivación de caja a plataforma
 - Modulo de reporte de derivaciones
 - Integración con Excel Online para el embebido de archivos en web CRM
 - Módulo de Tablero Comercial
 - Módulo de reasignación de ofertas
 - Módulo de creación de nuevos usuarios y reinicio de contraseñas
 - Módulo de consulta de campañas de convenios
 - Proceso automático de consolidación de datos de contactabilidad
 - Proceso automático de actualización de ofertas concretadas.
 - Proceso automático con WebScraping para la obtención de registros de ventas de todos los productos del banco desde una web externa.
-
- Banca MicroEmpresa
 - Módulo gestión de ofertas clientes preferente, potenciales con ofertas y desertores.
 - Módulo de agenda comercial, para planificación de gestiones.

- Integración con Google Maps para obtención la ruta crítica en auto, bicicleta o a pie.
 - Módulo de reporte de seguimiento de avance de Campañas.
 - Proceso automático de envío de base desde la nube a servidor local BP.
-
- Banca Empresa
 - Módulo de consulta de campañas de cliente de banca empresa.
 - Módulo de agenda comercial, para planificación de gestiones.
-
- Participo del proceso de gestión de Campañas, realizando el procesamiento y carga de las mismas a los sistemas de gestión CRM.
 - Coordino con los gerentes de productos, agencias y zonales de banca personas, micro y grandes empresas para lograr recoger los requerimientos de cada área y lograr un alto Impacto al desarrollar las soluciones.
 - Participo continuamente durante el desarrollo de la campaña realizando seguimiento, cambios y dando solución a cada incidente que se presente.
 - Soy un miembro activo de la Comunidad de Ingenieros de Software de Banco Pichincha.

5. REFERENCIAS LABORALES

Jared Goñez Montoya

JEFE DE ANALYTICS Y CRM | BANCO PICHINCHA

<https://www.linkedin.com/in/ivanmostacero/>

Celular: 940 401 202

Ivan Mostacero Plasencia

GERENTE DE RED DE AGENCIAS | BANCO PICHINCHA

<https://www.linkedin.com/in/ivanmostacero/>

Celular: 993 513 128

Rocio Alarcón

JEFE DE MODELOS DE GESTION DE AGENCIAS | BANCO PICHINCHA

<https://www.linkedin.com/in/roc%C3%A0oalarcon/>

Celular: 997 955 187

Oscar Izquierdo Monzón

GERENTE DE AGENCIA | BANCO PICHINCHA

<https://www.linkedin.com/in/oscar-rub%C3%A9n-izquierdo-monz%C3%B3n-21a091a4/>

Celular: 997 955 187

Juan Francisco Barthelmes Galloso

GERENTE ZONAL | BANCO PICHINCHA

<https://www.linkedin.com/in/juan-francisco-barthelmes-galoso-6b72ab73/>

Celular: 959 698 089

6. CAPACITACIÓN ACADÉMICA COMPLEMENTARIA

1. Massachusetts Institute of Technology

- Certificado: Data Science y Big Data. Decisiones Basadas en Datos
- Mayo – Julio 2021

2. Data Mining Consulting

- Curso: Credit Scoring
- Junio 2021

- Curso: Machine Learning Inmersion
- Marzo 2021

- Curso: Machine Statistics for Data Science
- Marzo 2021

- Curso: Machine PFA Python for Analytics
- Marzo 2021

- Curso: Machine FPP Fundamentos de Programación Python
- Enero 2021

- Curso: SQL Server Avanzado ETL
- Julio 2020

3 CertiProf

- Certificado: OKRC Certified OKRCP
- Mayo 2021
-
- Certificado: SMPC Scrum Master Professional Certificate
- Abril 2021

4. ITCERT

- Insignia Digital: Scrum Master self paced course
- Diciembre 2021
-
- Insignia Digital: Agile Coach Training self paced course
- Agosto 2021

5. CIBERTEC

- Curso: Power BI para Excel
- Junio 2019



UNIVERSIDAD NACIONAL PEDRO RUIZ GALLO
FACULTAD DE INGENIERÍA CIVIL, DE SISTEMAS Y ARQUITECTURA
UNIDAD DE INVESTIGACIÓN



"Año del Fortalecimiento de la Soberanía Nacional"

CONSTANCIA DE APROBACIÓN DE ORIGINALIDAD DE TESIS

Según Res. N° 659-2020-R

Yo, ARTEAGA LORA ROBERTO CARLOS, asesor de tesis del bachiller:

LLANOS EDQUEN, NIXON JAMES

TITULADA:

INFORME DE SUFICIENCIA PROFESIONAL: "MEJORAMIENTO DEL PROCESO DE GENERACIÓN DE BASE DE CONTACTABILIDAD DE CLIENTES EN BANCO PICHINCHA"

Luego de la revisión exhaustiva del documento constato, que el mismo tiene un índice de similitud de **15%** verificable en el reporte de similitud del programa TURNITIN.

El suscrito analizó dicho reporte y concluyó que, cada una de las coincidencias detectadas **NO CONSTITUYEN PLAGIO**. A mi leal saber y entender, la tesis cumple con todas las normas para el uso de citas y referencias establecidas por la Universidad Nacional Pedro Ruiz Gallo.

Se expide la presente según lo dispuesto en la Resolución N° 659-2020-R, de fecha 8 de setiembre de 2020, que aprueba la directiva para la evaluación de originalidad de los documentos académicos, de investigación formativa y para la obtención de Grados y Títulos de la UNPRG:

Lambayeque, 5 de setiembre de 2022

Atentamente,

MA. ING. ROBERTO CARLOS ARTEAGA LORA
DNI 16755764

Se adjunta:
Recibo digital de Turnitin
Revisión de informe en Turnitin



Recibo digital

Este recibo confirma que su trabajo ha sido recibido por Turnitin. A continuación podrá ver la información del recibo con respecto a su entrega.

La primera página de tus entregas se muestra abajo.

Autor de la entrega: Nixon James Llanos Edquen
Título del ejercicio: Informe borrador de tesis
Título de la entrega: MEJORAMIENTO DEL PROCESO DE GENERACIÓN DE BASE DE...
Nombre del archivo: Informe_Profesional_-_Nllanos_v6.docx
Tamaño del archivo: 2.22M
Total páginas: 60
Total de palabras: 6,602
Total de caracteres: 35,585
Fecha de entrega: 01-sept.-2022 10:52p. m. (UTC-0500)
Identificador de la entre... 1891101115



MEJORAMIENTO DEL PROCESO DE GENERACIÓN DE BASE DE CONTACTABILIDAD DE CLIENTES EN BANCO PICHINCHA (FINAL)

INFORME DE ORIGINALIDAD

15%	15%	2%	4%
INDICE DE SIMILITUD	FUENTES DE INTERNET	PUBLICACIONES	TRABAJOS DEL ESTUDIANTE

FUENTES PRIMARIAS

1	repositorio.unprg.edu.pe:8080 Fuente de Internet	8%
2	www.scribd.com Fuente de Internet	1%
3	ri.ues.edu.sv Fuente de Internet	1%
4	repositorio.lamolina.edu.pe Fuente de Internet	1%
5	Submitted to euroinnova Trabajo del estudiante	1%
6	hdl.handle.net Fuente de Internet	1%
7	www.fotografiemonique.nl Fuente de Internet	<1%
8	repositorio.utn.edu.ec Fuente de Internet	<1%



MA. ING. ROBERTO CARLOS ARTEAGA LORA
DNI 16755764

