

 Institución Universitaria	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-27

ESTUDIO DEL SERVICIO DE INTEGRACIÓN DE DATOS DEL SQL SERVER (SSIS) Y CASO DE ESTUDIO SOBRE MIGRACIÓN ETL.

Yesika Marcela Montes Rodas

Ingeniería de Sistemas

Director(es) del trabajo de grado

MANUEL SEGUNDO BLANCO PALENCIA

INSTITUTO TECNOLÓGICO METROPOLITANO

2018

	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

RESUMEN

La necesidad de integración de datos en el mundo actual es cada vez mayor. La cantidad de información que se genera a diario es cada vez más grande y diversa. Esta necesidad está dada porque la forma de representar de la información en las diferentes fuentes es también muy diversa; de allí que al integrarse deba ser preparada para garantizar en los procesos de tomas de decisiones y de cualquier análisis o consolidación que se realice la confiabilidad de los resultados.

La selección de una herramienta que de soporte a las necesidades del proceso de integración constituye un elemento muy importante en el éxito del mismo. En este trabajo se realiza un estudio de la herramienta de integración de datos Microsoft SQL Server Integration Services (SSIS), para determinar sus posibilidades de satisfacer la creación de las complejas soluciones que se presentan en el ámbito de la creación y mantenimiento de los almacenes de datos.

En el presente trabajo se mostrará un poco el manejo que se le da a la herramienta, su arquitectura y su forma de organizar los datos para lograr la integración de los mismos. Se mostrará el paquete, como núcleo principal de la herramienta de integración, con sus componentes que lo integran y sus interrelaciones.

Para finalizar se expondrá un ejemplo de cómo funciona la herramienta ante una migración real, que realiza una tarea necesaria en la creación de un almacén de datos: Extraer y Publicar las solicitudes pendientes por aprovisionar de los productos asociados a IDC.

 Institución Universitaria	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

RECONOCIMIENTOS

Gracias a Dios por haberme dado fortaleza para salir adelante en mi carrera.

Gracias a TigoUne por la oportunidad al escogerme para la realización de mis prácticas profesionales y al ITM por haberme seleccionado para llegar a esta gran empresa.

El logro de esta etapa en mi vida quiero dedicárselo a mis padres, y a todos quienes en algún momento colaboraron directa o indirectamente en este maravilloso resultado, los quiero mucho. A mis hermanos y a toda mi familia por su apoyo absoluto en cada momento de mi vida.

Por su puesto no pueden faltar los amigos, compañeros y profesores, que siempre estuvieron ahí para animarme, alentarme y con todo esto lograr crecer como persona cada día más.

	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

ACRÓNIMOS

SSIS SQL Server Integration Services, es una plataforma que permite generar soluciones de integración de datos de alto rendimiento, entre las que se incluyen paquetes de extracción, transformación y carga de datos (ETL) para el almacenamiento de datos.

SGBD sistema gestor de base de datos, es un conjunto de programas que permiten el almacenamiento, modificación y extracción de la información en una base de datos, además de proporcionar herramientas para añadir, borrar, modificar y analizar los datos.

SQL Se trata del lenguaje de consulta estructurado, para realizar consultas a nivel de base de datos.

ORACLE Es un sistema de gestión de base de datos de tipo objeto-relacional.

ETL Estas siglas corresponden a Extraer, transformar y cargar.

BI Las siglas significan Inteligencia de negocios. Este es un conjunto de metodologías, aplicaciones y tecnologías que permiten reunir, depurar y transformar datos de los sistemas transaccionales.

FTP (File Transfer Protocol) es un protocolo de transferencia de archivos entre sistemas conectados a una red TCP basado en la arquitectura cliente-servidor.

XML eXtensible Markup Language, es un formato universal para datos y documentos estructurados.

 Institución Universitaria	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

TABLA DE CONTENIDO

1.	INTRODUCCIÓN	
1.1.	Objetivo General	
1.2.	Organización del trabajo final de practica	
2.	MARCO TEÓRICO	
2.1.	Aspectos Teóricos del Proceso de Integración de Datos	
2.2.	Extracción, Transformación y Carga	
3.	METODOLOGÍA	
4.	RESULTADOS Y DISCUSIÓN.....	
4.1.	Realización de una aplicación BI en SSIS	
4.2.	Otras funciones asignadas de TigoUne	
4.2.1.	Funciones (administrativas) en BI	
5.	CONCLUSIONES, RECOMENDACIONES Y TRABAJO FUTURO	
5.1.	Conclusiones	
5.2.	Recomendaciones	
5.3.	Trabajo Futuro	
	REFERENCIAS	
	APÉNDICE	

	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

TABLA DE ILUSTRACIONES

Ilustración 1: Arquitectura del servicio de integración de datos	21
Ilustración 2: Flujo de Datos de un Paquete.....	23
Ilustración 3: Macro a migrar de Access	28
Ilustración 4: Macro a migrar de Access ampliada	29
Ilustración 5: Consultas definidas en Access	29
Ilustración 6: Migración de Access a SSIS.....	30
Ilustración 7: Migración de Access a SSIS, esquema completo	31
Ilustración 8: Variables usadas en ETL migrada.....	32
Ilustración 9: Creación de reporte	32
Ilustración 10: Reporte, Parametrizar_Analytics_Medios_Magneticos_Open6	33
Ilustración 11: Reporte, Ejecutar_Analytics_Medios_Magneticos_Open6.....	34
Ilustración 12: Ruta de Servidor de Reportes.....	34
Ilustración 13: Reporte para ejecutar	35
Ilustración 14: Reporte en ejecución 1	35
Ilustración 15: Reporte en ejecución 2	35
Ilustración 16: Reporte ejecutado	36
Ilustración 17: Creación de Change Order	36
Ilustración 18: Campos importantes en la Creación de Change Order.....	37
Ilustración 19: Documentos requeridos en la Creación de Change Order.....	37
Ilustración 20: Documentos necesarios para Paso a Ambiente de Pruebas	38
Ilustración 21: Incidentes reportados para asignación	39
Ilustración 22: Detalle de un Incidente	40
Ilustración 23: Opción de Transferencia de un Incidente	40
Ilustración 24: Asignar persona para Transferencia de Incidente	41
Ilustración 25: Transferencia realizada	41
Ilustración 26: Opción de Catalogo de Servicios.....	42
Ilustración 27: Opción para asignación o retiro de permisos.....	43
Ilustración 28: Opción para gestión de cuentas de Red.....	44
Ilustración 29: Opción Ingreso/Retiro de usuarios	45
Ilustración 30: Formulario para Ingreso/Retiro de usuarios	46
Ilustración 31: Lista de Opciones Crear/Retirar.....	46
Ilustración 32: Diligenciamiento Formulario Ingreso/retiro de usuario	47

	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

1. INTRODUCCIÓN

La información es considerablemente importante en los procesos históricos y de trayectoria de las organizaciones; el avance de las tecnologías de Inteligencia de negocio obliga a migrar a herramientas más potentes para trabajar con grandes cantidades de datos que se acumulan de forma útil. Dado que la información está directamente relacionada con la calidad de la toma de decisiones y los riesgos asumidos en el cumplimiento de las metas de la organización, es importante que en la migración de los datos se mantenga su valor de una forma óptima. En este proceso de centralización de los datos debe hacerse todo lo posible por corregir los errores que puedan surgir, a pesar de la complejidad, la heterogeneidad, y el error humano en los mecanismos de captura de datos de un sistema.

Microsoft SQL Server Integration Services (SSIS), como gestor de datos en su evolución ha ido incorporando cada vez más herramientas que facilitan esta tarea y en particular a partir del SQL Server 2005 se ofrece el SSIS como plataforma en que se pueden desarrollar una serie de operaciones que contribuyen a la integración de datos a través de proyectos, permitiendo crear soluciones robustas y complejas de ETL de una forma fácil.

En el proceso de migración de datos, desde diversas fuentes, conocido como Extracción, Transformación y Carga (ETL); es importante que los datos posean calidad al ser migrados, por tanto, es necesario encontrar el programa lo suficientemente eficaz que lo permita.

Este trabajo contribuye a explicar distintas posibilidades que brinda esta herramienta para posteriormente se aprovecharía de comparativo con otras que han ido apareciendo en el mercado con este propósito.

	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

1.1. Objetivo General

El objetivo general de este informe consiste en realizar un estudio que permita describir el proceso de integración de datos y establecer en qué medida se corresponde el Servicio de integración del SQL Server Integration Services (SSIS) con esa caracterización.

Objetivos específicos

- Realizar el estudio de algunos de los aspectos teóricos relacionados con el proceso de integración de datos.
- Describir las posibilidades que brinda el servicio de integración de datos con respecto a la caracterización lograda en el objetivo anterior.
- Realizar y documentar un caso de estudio con el servicio de integración de datos del SQL Server Integration Services (SSIS).

1.2. Organización del trabajo final de practica

La estructura del documento está dividida en tres capítulos:

Capítulo 1. Aspectos teóricos del proceso de integración de datos: abordará algunos de los aspectos teóricos que tienen que ver con la investigación de Integración de datos y ETL, cuáles son los escenarios que demandan este proceso, y la caracterización del mismo. Capítulo 2. Presentación del SISS: describirá el servicio de la herramienta de SSIS, la arquitectura de los paquetes, las soluciones, proyectos, el paquete y sus componentes.

Capítulo 3. Presentación del caso de estudio: describirá a través de un caso de estudio una solución de integración con el uso del SSIS.

	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

2. MARCO TEÓRICO

2.1. Aspectos Teóricos del Proceso de Integración de Datos

Actualmente en todas las esferas de la vida tanto personales, empresariales como globales, el volumen de los datos crece exponencialmente, por lo que su organización se ha convertido en una tarea muy importante. Disciplinas como la integración de datos, la migración, la sincronización, la inteligencia de negocios(BI), entre otras, permiten hacerlo.

La integración de datos consiste en combinar datos de varias fuentes diferentes, que se almacenan utilizando diversas tecnologías, proporcionando a los usuarios una vista unificada de los datos.

El proceso de integración de datos es significativo en una variedad de situaciones, que incluyen tanto la comercial (cuando dos empresas similares deben fusionar sus bases de datos) así como en dominios científicos (por ejemplo, cuando se combinan los resultados de investigación de diferentes repositorios de bioinformática). Aparece con mayor frecuencia en la medida que el volumen y la necesidad de compartir los datos existentes aumenta (*Lane, 2006*). Este proceso se ha convertido en el centro de un extenso trabajo teórico y de numerosos problemas abiertos para la investigación científica.

Sobre la integración de datos se ha desarrollado una teoría que constituye un subconjunto de la teoría de base de datos y que formaliza los conceptos subyacentes del problema en la lógica de primer orden. La aplicación de la teoría da indicaciones sobre la viabilidad y las dificultades de la integración de datos. (*Lenzerini, 2002*)

Tipos de herramientas de integración de información: En (*Haas, 2008*), se relacionan una serie de tipos de herramientas que son usadas en la práctica en el proceso de integración de información, entre las cuales se encuentran las siguientes:

 Institución Universitaria	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

- Herramientas de ETL: estas herramientas están dirigidas a resolver los problemas de la carga de datos hacia un almacén (*Kimball, 2004*), simplificando la programación de ficheros scripts. Una típica herramienta de ETL debe incluir un repertorio de operaciones de limpieza y de reformulación al estilo SQL para recuperar, combinar y ordenar los datos, así como también operaciones de planificación para controlar cuando realizar la carga periódica o la actualización del almacén de datos.
- Herramientas de Enterprise Information Integration (EII): las cuales están dirigidas específicamente a resolver la integración de datos virtuales, que consiste en trasladar las solicitudes de los usuarios a las fuentes e integrar los resultados como si esto hubiera sido hecho sobre una única base de datos (*Halevy, 2005*). Estos productos EII varían en dependencias de los orígenes de datos que serán integrados. Por ejemplo, algunos se centran en la integración de bases de datos SQL, otros en la integración de servicios Web y otros en bases de datos de bioinformática.
- Herramientas de conversión objeto-relacional (object-relational mapper): Son herramientas que permiten el entendimiento entre los programas de aplicación, que generalmente están hechos en lenguajes orientados a objetos y los datos que ellos acceden, los que generalmente están en bases de datos relacionales. La correspondencia entre ambos se expresa a través de un programa en un lenguaje de alto nivel que al ser compilado traduce las solicitudes y actualizaciones hechas sobre la interfaz orientada a objeto a solicitudes y actualizaciones sobre la base de datos relacional (*Melnik, 2007*)
- Herramientas de gestión de documentos: que realizan la integración de los documentos de grandes empresas, facilitando la colaboración y evitando la duplicidad de trabajo. Las herramientas de este tipo hacen los documentos disponibles a través de páginas Web o de un sistema manejador de contenidos.

La integración de datos debe verse como la convergencia de varias tecnologías existentes y de algunas nuevas que están emergiendo. Este proceso reúne tecnologías que son necesarias desde el lado operacional del negocio con otras que son necesarias desde el lado de la toma de

	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

decisiones del negocio. La integración de datos trata con la incorporación de todos los tipos de datos organizacionales en un todo unificado. (*Landscape*)

En una estrategia de integración de datos deben estar presentes los aspectos de conectividad y metadatos. Cualquier solución debe brindar la posibilidad de conectividad a través de acceso nativo utilizando los utilitarios comunes y las formas de acceso abierto (tales como Open Database Connectivity: ODBC), a las principales fuentes de datos estructurados, como bases de datos relacionales, archivos planos, ERP (Enterprise Resource Planning, por sus siglas en inglés) y a los lenguajes de marcas como XML para leer y escribir. Las capacidades de conectividad de los datos deben también facilitar el acceso a la información desde diferentes sistemas operativos tales como UNIX, Windows, preferiblemente sin tener que hacer uso de ficheros intermedios. Además, la conectividad debe brindar soporte para leer y escribir desde colas de mensajes y la habilidad de recibir y enviar datos para y desde el servicio Web. Debe también facilitar la conectividad a fuentes de datos no estructurados. (*Landscape*)

Otras áreas que deben estar presentes en las estrategias de integración de datos son:

- Almacenamiento de datos (Data Warehousing) La utilización más difundida y conocida de integración de datos es la construcción de los almacenes de datos. El principal beneficio de un almacén de datos es que permite a una empresa llevar a cabo análisis basados en los datos del almacén, lo que no sería posible solo con los datos disponibles en los sistemas de origen, pues en estas fuentes, aun cuando los datos pueden tener los mismos nombres, podrán referirse a entidades diferentes.

En la construcción de un almacén de datos los datos no son “ingresados”, tal y como ocurre en los sistemas transaccionales u operacionales, en su lugar el almacén es poblado tomando los datos que ya existen en otras fuentes: sistemas operacionales y otros ficheros. El proceso de importación y manipulación de los datos transaccionales en el almacén de datos es referido como Extracción, Transformación y Carga (Extraction, Transformation and Loading, ETL). (*Alkis Simitsis National Technical University of Athens et al., 2005*)

	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

Cualquier estrategia de integración de datos debe posibilitar el construir y mantener almacenes y mercados de datos a través del proceso de ETL.

- Migración de datos (Data migration). La migración de datos hace referencia al traspaso de información entre bases de datos. Por ejemplo, si se tiene una aplicación sobre una base de datos como Access y posteriormente "crece" de manera que hace falta un sistema gestor de bases de datos potente, lo más seguro es que se use Oracle, SQL Server u otros sistemas similares.

En este caso, los datos, que están en formato "Access" deberán pasar al formato de las bases de datos del gestor escogido. La migración de los datos consiste en convertir los datos desde un sistema de base de datos a otro. Este proceso puede conducir a la creación de tablas nuevas o a la modificación de las existentes, y posiblemente a cambios en algunos tipos de datos que existen en un gestor base de datos, pero no en el otro, entre algunas de las transformaciones. *(Wikipedia, 2012b)*

Actualmente la mayoría de los SGBD incluyen herramientas de ayuda a la migración, el gestor Microsoft SQL Server la brinda como parte del SQL Server Integration Services (SSIS).

2.2. Extracción, Transformación y Carga

El proceso de Extracción, Transformación y Carga puede ser visto como la máxima expresión de la integración de datos en el escenario de la construcción de los almacenes de datos, aunque el mismo va más allá de este contexto, posibilitando la resolución de tareas rutinarias llevadas a cabo por los administradores en las bases de datos y en las tareas de integración de sistemas legados con las nuevas aplicaciones.

ETL, es el proceso que permite a las organizaciones mover datos desde múltiples fuentes, reformatearlos, limpiarlos, y cargarlos en otra base de datos, mercados de datos, o en almacenes de datos para analizarlos, o en otros sistemas operacionales para apoyar un proceso de negocio. *(Wikipedia, 2012)*

	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

En *(Prix, 2010)* y *(Alkis Simitsis, 2005)* se expresa que las herramientas ETL, son consideradas piezas de software fundamentales en la construcción, explotación y evolución de un almacén de datos, responsables de la extracción de los datos desde las diferentes fuentes, su limpieza, adecuación e inserción en los almacenes de datos proporcionando las siguientes funcionalidades:

- La extracción de la información relevante desde el lado de las fuentes.
- La transportación de esta información al área de datos intermedia (Data Staging Area, DSA).
- La transformación requerida de esta información (personalización e integración).
- La limpieza de datos sobre la base de las reglas de negocio de la base de datos, metadatos y restricciones.
- La propagación y carga de los datos al almacén de datos y la actualización de los mercados de datos.

Proceso de Extracción: La primera parte del proceso ETL consiste en extraer los datos desde los sistemas de origen. La mayoría de los proyectos de almacenamiento de datos fusionan datos provenientes de diferentes sistemas de origen. Cada sistema separado puede usar una organización diferente de los datos o formatos distintos. Los formatos de las fuentes normalmente se encuentran en bases de datos relacionales o ficheros planos, pero pueden incluir bases de datos no relacionales u otras estructuras diferentes. La extracción convierte los datos a un formato preparado para iniciar el proceso de transformación.

Una parte interna del proceso de extracción es la de analizar los datos extraídos, de lo que resulta un chequeo que verifica si los datos cumplen la pauta o estructura que se esperaba. De no ser así los datos son rechazados. Un requerimiento importante que se debe exigir a la tarea de extracción es que esta cause un impacto mínimo en el sistema origen. Si los datos a extraer son muchos, el sistema de origen se podría ralentizar e incluso colapsar, provocando que éste no pueda utilizarse con normalidad para su uso cotidiano. Por esta razón, en sistemas grandes las operaciones de extracción suelen programarse en horarios o días donde este impacto sea nulo o mínimo. *(Wikipedia, 2012a)*

	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

Si los datos operacionales residen en un SGBD Relacional, el proceso de extracción se puede reducir a, por ejemplo, consultas en SQL o rutinas programadas. En cambio, si se encuentran en un sistema no convencional o fuentes externas, ya sean textuales, hipertextuales, hojas de cálculos, etc., la obtención de los mismos puede ser un tanto más dificultoso, debido a que, por ejemplo, se tendrán que realizar cambios de formato y/o volcado de información a partir de alguna herramienta específica y puede tener que realizarse a partir de informes o de datos proporcionados por los propietarios que deberán ser procesados posteriormente.

Una vez que los datos son extraídos con la calidad requerida se les lleva a un segundo proceso llamado Transformación, donde se le aplican una serie de reglas de negocio o funciones que los preparan adecuadamente para el posterior proceso de carga.

Después de la extracción un paso no menos importante debe ser realizado, la transportación (*Alkis Simitsis National Technical University of Athens et al., 2005*) Para la cual pueden emplearse las tareas de codificación/decodificación, compresión/descompresión, los FTP y otras.

Proceso de Transformación: Esta función es la encargada de convertir aquellos datos inconsistentes en un conjunto de datos compatibles y congruentes, para que puedan ser cargados en el almacén de datos. (*Ing. Bernabeu R. Dario*). Las principales transformaciones ocurren en el DSA. Las transformaciones pueden clasificarse según (*java.net*), en dos tipos fundamentales, las que ocurren a nivel de esquema y las que ocurren a nivel de los datos. Entre las que ocurren a nivel de esquemas son significativos los conflictos de nombres, los que se pueden presentar como homónimos (un mismo nombre es usado para representar objetos diferentes) o sinónimos (un mismo objeto es representado por diferentes nombres), y los conflictos estructurales que están dados porque en las diferentes fuentes un mismo objeto es representado de formas distintas.

En el nivel de los datos son muchas las transformaciones que pueden tener lugar debido a que la lista de problemas que se pueden presentar a este nivel es mucho más variada, entre ellas se pueden señalar la aparición de registros duplicados, diferente representación de un valor, interpretación diferente de un mismo valor, y otros problemas técnicos de bajo nivel como la conversión de tipos de datos, aplicación de máscara de formatos, etc.

	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

Limpeza de datos: La limpieza de datos (data cleansing, data scrubbing), es el proceso de descubrimiento, corrección o eliminación de datos erróneos de una base de datos (*López-Porrero, 2011*). El proceso de limpieza de datos permite identificar datos incompletos, incorrectos, inexactos, no pertinentes, etc. y luego substituir, modificar o eliminar estos datos sucios (data dirty). Después de la limpieza, la base de datos podrá ser compatible con otras bases de datos similares en el sistema.

Los almacenes de datos constantemente están cargando y refrescando cientos de datos de diferentes fuentes, por lo que la probabilidad de que algunas de las fuentes contengan datos sucios es alta. Además, como los almacenes de datos se usan en sistemas de ayuda a la toma de decisiones, la corrección de estos es vital para evitar llegar a conclusiones erróneas (*López-Porrero, 2011*). Por ejemplo, información duplicada o ausente podría producir estadísticas incorrectas; debido a esto, la limpieza de datos se considera uno de los mayores problemas en los almacenes de datos.

Proceso de Carga: La fase de carga es el momento en el cual los datos resultantes de la fase anterior (Transformación) son cargados en el sistema de destino. Dependiendo de los requerimientos de la organización, este proceso puede abarcar una amplia variedad de acciones diferentes. En algunas bases de datos se sobrescribe la información antigua con nuevos datos. En los almacenes de datos se mantiene un historial de los registros de manera que se pueda hacer una auditoría y disponer de un rastro de toda la historia de un valor a lo largo del tiempo.

Esta función se encarga, por un lado, de realizar las tareas relacionadas con la carga inicial, la cual se refiere precisamente a la primera carga de datos que se le realizará al almacén de datos, por lo general, esta tarea consume bastante tiempo, ya que se deben insertar registros que han sido generados aproximadamente, y en casos ideales, durante más de cinco años, por otro lado, se puede hacer la actualización o mantenimiento periódico (siempre teniendo en cuenta un intervalo de tiempo predefinido para tal operación), ya que los mismos mueven pequeños volúmenes de datos, y su frecuencia está dada en función de la granularidad del almacén de datos y los requerimientos de los usuarios.

	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

El objetivo de esta tarea es añadir al depósito aquellos datos nuevos que se fueron generando después del último refrescamiento. Existe siempre la posibilidad de cargar el almacén de datos desde cero, este proceso se denomina carga total. La fase de carga interactúa directamente con la base de datos de destino. Al realizar esta operación se aplicarán todas las restricciones y disparadores que se hayan definido en esta (por ejemplo, valores únicos, integridad referencial, campos obligatorios, rangos de valores). Estas restricciones y disparadores (si están bien definidos) contribuyen a que se garantice la calidad de los datos en el proceso ETL, y deben ser tomados en cuenta.

Procesamiento paralelo en sistemas ETL: El proceso de ETL, en un solo hilo, puede ser eficiente para algunos casos en que la cantidad de datos sea pequeña, pero cuando estos datos crecen exponencialmente, se hace necesario implementar una solución que permita realizar varias tareas a la vez, lo que se conoce como procesamiento paralelo.

Un desarrollo en el software ETL es la aplicación del procesamiento paralelo, lo que ha permitido utilizar una serie de métodos para mejorar el rendimiento general de los procesos ETL cuando se trata de grandes volúmenes de datos. Hay 3 tipos principales de paralelismos que se pueden implementar en las aplicaciones ETL:

- De datos: Consiste en dividir un único archivo secuencial en pequeños archivos de datos para proporcionar acceso paralelo.
- De segmentación (pipeline): Permitir el funcionamiento simultáneo de varios componentes en el mismo flujo de datos. Por ejemplo, buscar un valor en el registro número 1 a la vez que se suman dos campos en el registro número 2.
- De componente: Consiste en el funcionamiento simultáneo de múltiples procesos en diferentes flujos de datos, pertenecientes todos ellos a un único flujo de trabajo. Esto es posible cuando existen porciones dentro de un flujo de trabajo que son totalmente independientes entre ellas a nivel de flujo de datos.

Estos tres tipos de semejanza no son excluyentes, sino que pueden ser combinados para realizar una misma operación de ETL. No es recomendable usar muchos procesos en paralelo, pues el tiempo de procesamiento aumentaría en cuanto a procesador, pero caería en el llamado cuello de

	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

botella con la Entrada/Salida del disco, debido a que los procesos están leyendo simultáneamente del mismo disco. Una solución aconsejable en este caso es dispersar los archivos en discos y controladoras independientes o pasar a una tecnología de disco más rápida. *(Carroll)*

Los procesos ETL pueden ser muy complejos. Un sistema ETL mal diseñado puede provocar importantes problemas operativos, por lo que, por esta y otras muchas razones deben enfrentarse varios retos.

Conclusiones parciales:

Las herramientas de Integración de datos combinan datos de varias fuentes diferentes para proporcionar a los usuarios una vista unificada de los datos, de modo que son sumamente importantes en el proceso de toma de decisiones dentro de la inteligencia empresarial. Las herramientas de integración de datos de forma general deberán dar soporte para:

La conectividad con disímiles fuentes sobre distintas plataformas, el uso y mantenimiento de los metadatos, la migración de datos, la sincronización de datos y la construcción de almacenes de datos.

El proceso de ETL, donde la integración de datos es imprescindible, es la pieza de software responsable de la extracción de los datos desde algunas fuentes, su limpieza, personalización e inserción en un almacén de datos.

El desarrollo de un sistema de ETL es un proyecto complejo que demanda un plan previo para enfrentar su elaboración, a partir de los requerimientos impuestos por las necesidades del negocio.

	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

3. METODOLOGÍA

Servicio de integración de datos del SQL Server. El SSIS es una plataforma que reúne un motor y un ambiente de diseño para la creación de soluciones empresariales de transformación e integración de datos que da solución a disímiles y complejas tareas, desde una simple importación o exportación de datos hasta una solución completa de un sistema ETL, requerido en cualquier escenario de inteligencia de negocios.

Esta herramienta ofrece un ambiente de diseño gráfico muy intuitivo en el que se pueden crear complejos programas que resuelvan tareas de mantenimiento de las bases de datos, de comunicación con otros sistemas, como por ejemplo Servidores FTP o servicios de la Web que lean y escriban en ficheros XML e incluso resolver todo esto como parte de una única solución.

A diferencia de otros sistemas de ETL que se focalizan en el movimiento de los datos, hacia el área intermedia de datos, lugar donde se realizan las transformaciones, el servicio de integración del SQL Server realiza las transformaciones de los datos a su forma final dentro del propio proceso de extracción.

Usos típicos del servicio de integración de datos, el servicio de integración de datos sirve para realizar una gran variedad de procesos de migración, entre sus usos típicos se destacan:

- Combinar datos desde fuentes de datos heterogéneos.

Los datos suelen almacenarse en muchos formatos de almacenamiento distintos, por lo que extraer datos de todos los orígenes y combinarlos en un solo conjunto de datos coherente constituye un desafío. Esta situación puede producirse por diversas razones. Por ejemplo:

- Muchas organizaciones archivan información que está almacenada en sistemas de almacenamiento de datos antiguos. Estos datos pueden no ser importantes para las operaciones diarias, pero pueden resultar útiles para el análisis de tendencias, que requiere datos recopilados a lo largo de un período prolongado de tiempo.

	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

- Las sucursales de una organización pueden usar distintas tecnologías de almacenamiento de datos para almacenar los datos operativos. Es posible que el paquete tenga que extraer datos de hojas de cálculo y de bases de datos relacionales para poder combinar los datos.
 - Los datos pueden estar almacenados en bases de datos que usan distintos esquemas para los mismos datos. Es posible que el paquete tenga que cambiar el tipo de datos de una columna o combinar datos de varias columnas en una sola columna para poder combinar los datos.
 - El servicio de integración de datos puede conectarse a una gran variedad de orígenes de datos que incluyen desde ficheros planos hasta bases de datos relacionales. Una vez que se han combinado los datos de los diferentes orígenes, aplicando transformaciones que los expresan en formatos compatibles, estos son cargados a uno o varios destinos.
- Poblar almacenes y mercados de datos.

Los datos de los almacenes y mercados de datos suelen actualizarse frecuentemente y normalmente las cargas de datos son muy grandes. Mediante este servicio se pueden cargar las tablas de dimensiones y hechos en la base de datos. Si los datos de origen de una tabla de dimensión están almacenados en varias fuentes, estas se combinarán en un solo conjunto de datos y se podrá realizar la carga a la tabla de la dimensión en un solo proceso, en lugar de utilizar un proceso independiente para cada fuente.

La actualización de datos de almacenamientos de datos y mercados de datos puede ser compleja, ya que en ambos se suelen incluir dimensiones lentamente cambiantes que pueden ser difíciles de administrar mediante un proceso de transformación de datos. El servicio de integración de datos brinda un asistente que automatiza la compatibilidad para estas dimensiones.

En el proceso de carga a un almacén o mercado de datos es muy frecuente que se necesite consolidar la información de las fuentes, para ello el servicio de integración permite calcular funciones tales como SUM, AVERAGE y COUNT.

	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

- Limpiar y normalizar datos.

Independientemente de si los datos se van a cargar en una base de datos de procesamiento de transacciones en línea o de procesamiento analítico en línea, una hoja de cálculo o un archivo, hay que limpiarlos y normalizarlos antes de cargarlos. Puede ser necesario actualizar los datos porque los datos proceden de varias fuentes y en cada una se usan convenciones y estándares distintos. Con el SSIS los datos pueden limpiarse y normalizarse realizando operaciones para cambiar las mayúsculas y minúsculas de los datos, convertir datos a un tipo o formato distinto, crear nuevos valores de columna basados en expresiones, concatenar columnas, también puede limpiar datos reemplazando los valores de las columnas por valores de una tabla de referencia mediante una búsqueda exacta o aproximada y eliminando los registros duplicados.

- Automatizar las funciones administrativas y la carga de datos.

En un entorno de bases de datos se hace necesario automatizar las funciones administrativas como la copia de seguridad y la restauración de bases de datos, la copia de bases de datos de SQL Server y los objetos que contienen, la copia de objetos de SQL Server y la carga de datos. SSIS pueden realizar estas funciones. Además, facilita la copia de objetos de bases de datos de SQL Server como tablas, vistas y procedimientos almacenados, bases de datos, inicios de sesión, estadísticas y agregar, modificar y eliminar objetos y datos de SQL Server mediante instrucciones TransactSQL.

- Arquitectura del servicio de integración de datos.

El SSIS incluye diversos componentes, los que se muestran en la siguiente figura:

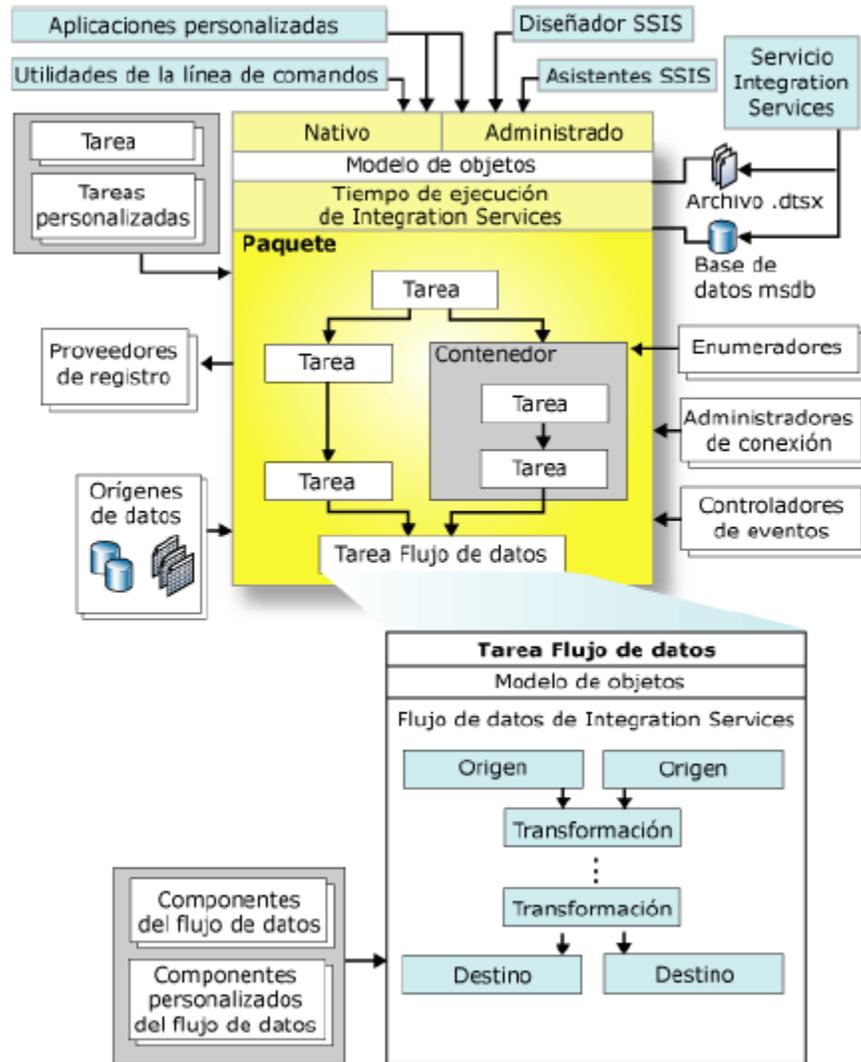


Ilustración 1: Arquitectura del servicio de integración de datos (2014). Recuperado de <http://slideplayer.es/slide/123948/>

	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

De estos, los más importantes para utilizar SSIS correctamente son:

- El diseñador.

El diseñador es una herramienta gráfica que se puede usar para crear y mantener soluciones. El Diseñador SSIS está disponible en Business Intelligence Development Studio como parte de un proyecto.

- Motor de Motor de Motor de ejecución.

El motor de ejecución del SSIS guarda el diseño de las soluciones, las ejecuta y brinda soporte para la autenticación, puntos de interrupción, configuración, conexiones y transacciones.

- Tareas y otros ejecutables.

El SSIS incluye como ficheros ejecutables varios objetos: los paquetes, los paquetes, los contenedores, contenedores, contenedores, contenedores, contenedores, contenedores, las tareas, los controladores de eventos y otras tareas las tareas desarrollados por los usuarios.

- El Motor de flujo de datos y los componentes de flujo de datos.

La tarea de flujo de datos encapsula el motor de flujo de datos. El motor de flujo de datos proporciona los búferes en memoria que mueven datos desde el origen hasta el destino y llama a las fuentes que extraen los datos desde archivos y bases de datos relacionales. El motor de flujo de datos también administra las transformaciones que modifican los datos y los destinos que cargan datos o los ponen a disposición de otros procesos. Los componentes de un flujo de datos son los orígenes, transformaciones y destinos que se incluyen en el SSIS, pudiéndose añadir otros componentes personalizados en un flujo de datos.

- Servicio de integración.

El servicio de integración permite usar SQL Server Management Studio para supervisar paquetes en ejecución y para administrar el almacenamiento de los paquetes.

- Asistente para importación y exportación de SQL.

	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

El Asistente para importación y exportación de SQL Server puede copiar datos entre orígenes de datos para los que esté disponible un proveedor de datos de .NET Framework o un proveedor OLE DB nativo. El Asistente ofrece también el método más simple para crear paquetes que copien datos de un origen en un destino.

- Arquitectura de un paquete de SSIS.

El paquete, en sí mismo, constituye el núcleo central del SSIS y no es más que una colección organizada de conexiones, elementos de flujo de control, elementos de flujo de datos, controladores de eventos, variables y configuraciones que se pueden ensamblar con la ayuda de las herramientas gráficas de diseño proporcionadas por SSIS o mediante programación. Una vez completado el paquete, se guarda en SQL Server, en el almacén de paquetes SSIS o en el sistema de archivos. El paquete es la unidad de trabajo que se recupera, ejecuta y guarda.

El siguiente diagrama muestra un paquete genérico que contiene un flujo de control con una tarea Flujo de datos que, a su vez, contiene un flujo de datos.

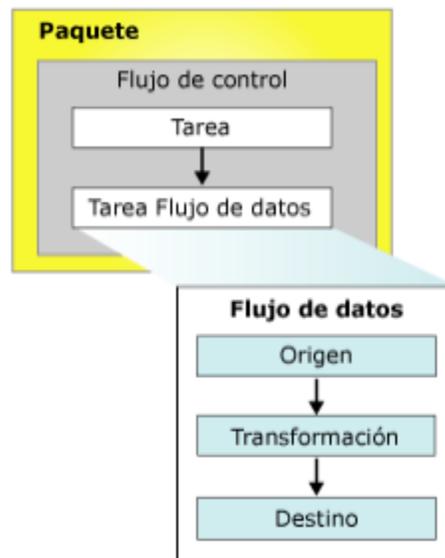


Ilustración 2: Flujo de Datos de un Paquete (2013). Recuperado de [https://technet.microsoft.com/es-es/library/ms141134\(v=sql.110\).aspx](https://technet.microsoft.com/es-es/library/ms141134(v=sql.110).aspx)

Una vez creado el paquete básico, puede agregársele otras características avanzadas para extender su funcionalidad y puede configurarse estableciendo propiedades a nivel de paquete que

	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

permiten implementar la seguridad, reiniciar paquetes desde puntos de rupturas e incorporar transacciones en el flujo de trabajo del paquete.

- Administradores de conexiones.

Los paquetes utilizan conexiones para realizar diferentes tareas y para implementar características, SSIS utiliza los administradores de conexiones para ello. Un ejemplo con dos conexiones:

- Un administrador de conexiones de archivo que se conecta con un archivo que contiene el código SQL que se usará en el paquete.
- Un administrador de conexiones OLE DB, que se conecta con la base de datos que está creada localmente en el equipo.

Sobre esta base datos se harán chequeos de existencia, de los registros de acuerdo con la referencia y se crearán tablas de trabajo temporales que requieren algunas transformaciones para hacer el trabajo.

Al mismo tiempo, SSIS ofrece otras posibilidades de conexiones:

- Conectar con orígenes y destinos de datos tales como archivos de texto, archivos XML, libros de Excel y bases de datos relacionales para extraer y cargar datos.
- Conectar con bases de datos relacionales que contienen datos de referencia para realizar búsquedas exactas o aproximadas.
- Conectar con bases de datos relacionales para ejecutar instrucciones SQL, tales como los comandos SELECT, DELETE e INSERT, así como procedimientos almacenados.
- Conectar con SQL Server para realizar tareas de transferencia y mantenimiento, tales como la realización de copias de seguridad de bases de datos y la transferencia de inicios de sesión.
- Escribir entradas del registro en archivos XML, de texto y tablas de SQL Server y configuraciones de paquete en tablas de SQL Server.

SSIS usa el administrador de conexiones como una representación lógica de una conexión.

En tiempo de diseño, se establecen las propiedades del administrador de conexiones para que describa la conexión física que crea cuando se ejecuta el paquete. Un administrador de conexiones

	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

incluye la propiedad `ConnectionString` que se establece en la etapa diseño. En el tiempo de ejecución, se crea una conexión física mediante el valor de la propiedad de la cadena de conexión.

Un paquete puede usar varias instancias de un tipo de administrador de conexiones y se pueden establecer las propiedades de cada instancia. En el tiempo de ejecución, cada instancia de un tipo de administrador de conexiones crea una conexión que tiene diferentes atributos.

SSIS proporciona diferentes tipos de administradores de conexión que permiten que los paquetes se conecten a una serie de orígenes de datos y servidores, los cuales se instalan cuando el servicio es instalado siendo posible que los usuarios puedan crear otros administradores de conexión.

- Variables de SSIS.

Para hacer realizar algunas tareas, es necesario que el SSIS use variables para almacenar valores que se puedan usar en tiempo de ejecución, en un paquete, sus contenedores, tareas y controladores de eventos. También pueden ser usadas en los componentes de la tarea Script y en las restricciones de precedencia que ordenan tareas y contenedores en un flujo de trabajo cuando estas incluyen expresiones. Existen dos tipos de variables: variables definidas por el usuario y variables del sistema.

Las variables pueden ser usadas con los siguientes fines:

- Actualizar propiedades de elementos del paquete en tiempo de ejecución. Por ejemplo, puede establecer dinámicamente el número de ejecutables simultáneos que admite un contenedor de bucles `Foreach`.
- Incluir una tabla de búsqueda almacenada en la memoria. Por ejemplo, un paquete puede ejecutar una tarea Ejecutar SQL que carga una variable con valores de datos.
- Cargar variables con valores de datos y usarlas posteriormente para especificar una condición de búsqueda en una cláusula `WHERE`.
- Cargar una variable con un número entero y luego usar el valor para controlar bucles dentro de un flujo de control de paquetes.
- Llenar valores de parámetros para instrucciones `Transact-SQL` en tiempo de ejecución. Por ejemplo, un paquete puede ejecutar una tarea Ejecutar SQL y luego usar variables para establecer dinámicamente los parámetros en una instrucción `Transact-SQL`.

	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

- Generar expresiones que incluyen valores de variable.
- Cuando el paquete es creado en el diseñador del SSIS, pueden verse los miembros de las colecciones de variables en la carpeta Variables en la ficha Explorador. Las carpetas listan las variables definidas por el usuario y variables del sistema.

- Destinos.

Un destino es el componente de flujo de datos que escribe los datos desde un flujo de datos en un almacenamiento de datos especificado, o crea un conjunto de datos en la memoria.

Un destino de SSIS debe tener por lo menos una entrada. La entrada contiene columnas de entrada, que proceden de otro componente de flujo de datos. Las columnas de entrada se asignan a columnas en el destino.

Muchos destinos también tienen una salida de error. La salida de error de un destino contiene columnas de salida, que normalmente contienen información sobre errores que se producen mientras se escriben datos en el almacén de datos de destino. Los errores se producen por muchos motivos diferentes, por ejemplo, una columna puede contener un valor NULL, mientras que la columna de destino no se puede establecer como NULL. El modelo de objetos de SSIS no limita la cantidad de entradas regulares y salidas de error que los destinos pueden tener, y se pueden crear destinos personalizados que implementan varias entradas y salidas de error.

- Metadatos externos.

Al crear un flujo de datos en un paquete mediante el Diseñador SSIS, los metadatos de los orígenes y destinos se copian en las columnas externas en orígenes y destinos, actuando como una instantánea del esquema. Cuando SSIS valida el paquete, el diseñador compara esta instantánea con el esquema del origen o destino y expone errores y advertencias, según las modificaciones.

SSIS proporciona un modo de trabajo sin conexión. Cuando se trabaja sin conexión, no se realizan conexiones a los orígenes o destinos que usa el paquete y los metadatos de las columnas externas no se actualizan. Tanto las entradas como las salidas de orígenes, transformaciones y destinos pueden configurarse para lograr diferentes objetivos, por ejemplo, una entrada se puede configurar para monitorear y controlar el flujo de datos, una salida se puede configurar para

	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

ofrecer información útil sobre el procesamiento de los datos. Por otra parte, las salidas de error, como opción predeterminada, contienen las columnas de salida y dos columnas de error: ErrorCode y ErrorColumn. Las columnas de salida contienen los datos de la fila que generó el error, ErrorCode proporciona el código de error y ErrorColumn identifica la columna que genera el error.

Es SSIS es una herramienta potente que permite una gran variedad de facilidades en los intentos de Integración de datos, específicamente en la Extracción, Transformación y Carga.

Aunque es una herramienta que se distribuye con el SQL Server, no limita su alcance en cuanto a la integración a fuentes de datos de este gestor, proporcionando una gran variedad de conexiones con fuentes de datos relaciones, de archivos de texto sin formato o con algunos formatos especializados.

Cuenta además con una amplia cantidad de transformaciones para lograr una limpieza óptima de los datos que serán almacenados en los destinos.

	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

4. RESULTADOS Y DISCUSIÓN

4.1. Realización de una aplicación BI en SSIS

Se construye la solución que permite realizar el proceso de extracción y publicación de las solicitudes pendientes por aprovisionar en Fenix Nal de los productos asociados a IDC.

En el área de BI de TigoUne, se requiere que se actualicen los datos asociados al pedido, al cliente y datos técnicos para su aprovisionamiento en plataformas tales como el Dominio, Login, Password, Megas, Plan Facturación, etc.

Este proceso se manejaba con Access y se requiere la migración a SSIS. (Por motivos de políticas de seguridad de la empresa, se restringen las imágenes de algunas capturas de pantalla de lo realizado durante los procesos, tanto de migración y gestión de incidencias, entre otras.)

Creación de ETL para migración de proyecto de Access a Integration Services Colas IDC.

Se visualiza la estructura conformada en Access,

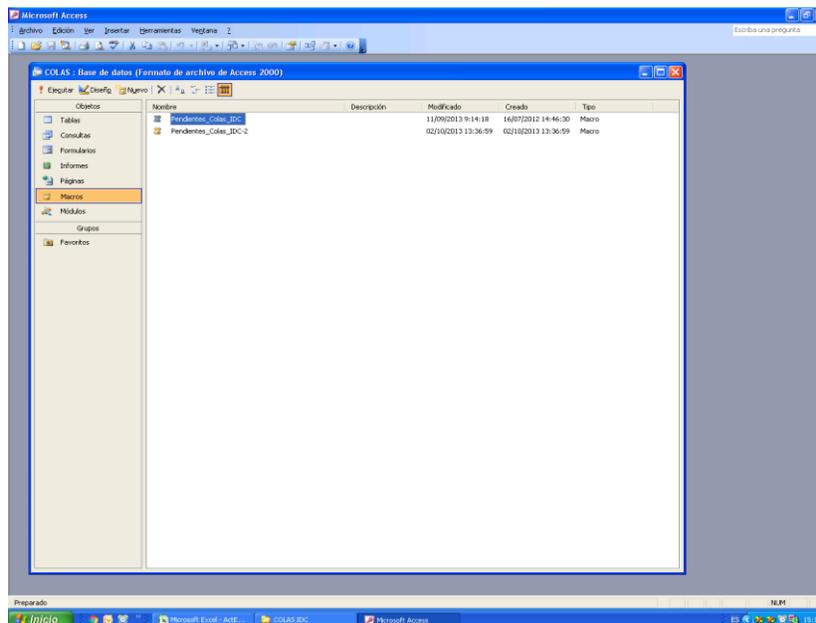


Ilustración 3: Macro a migrar de Access

 Institución Universitaria	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

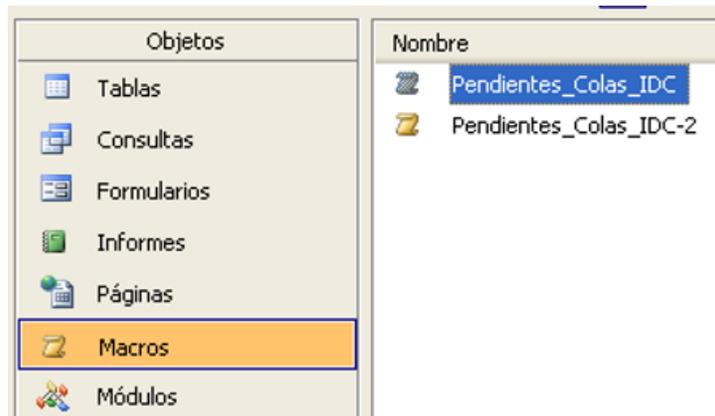


Ilustración 4: Macro a migrar de Access ampliada

Se valida cuantos procesos de consulta tiene,

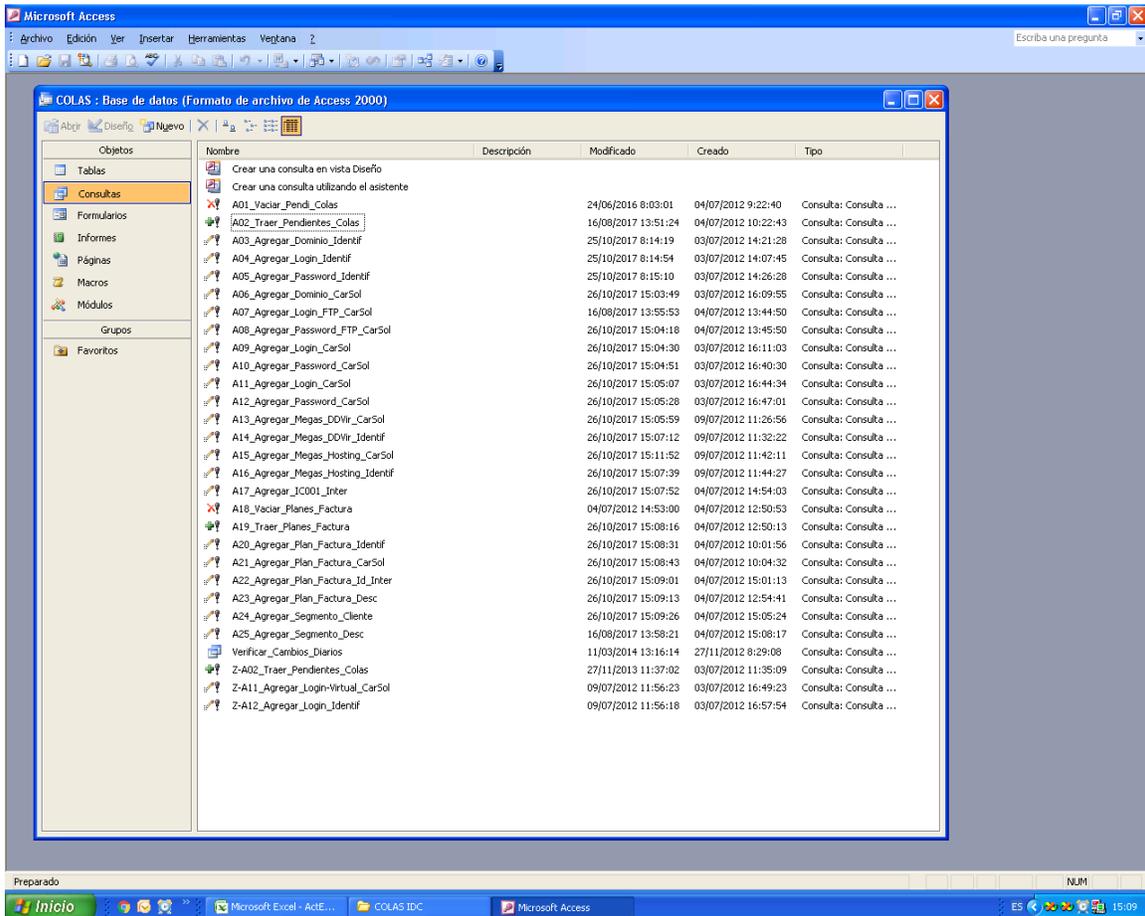


Ilustración 5: Consultas definidas en Access

 Institución Universitaria	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

En las siguientes imágenes se evidencia la migración realizada con ETL en integration services,

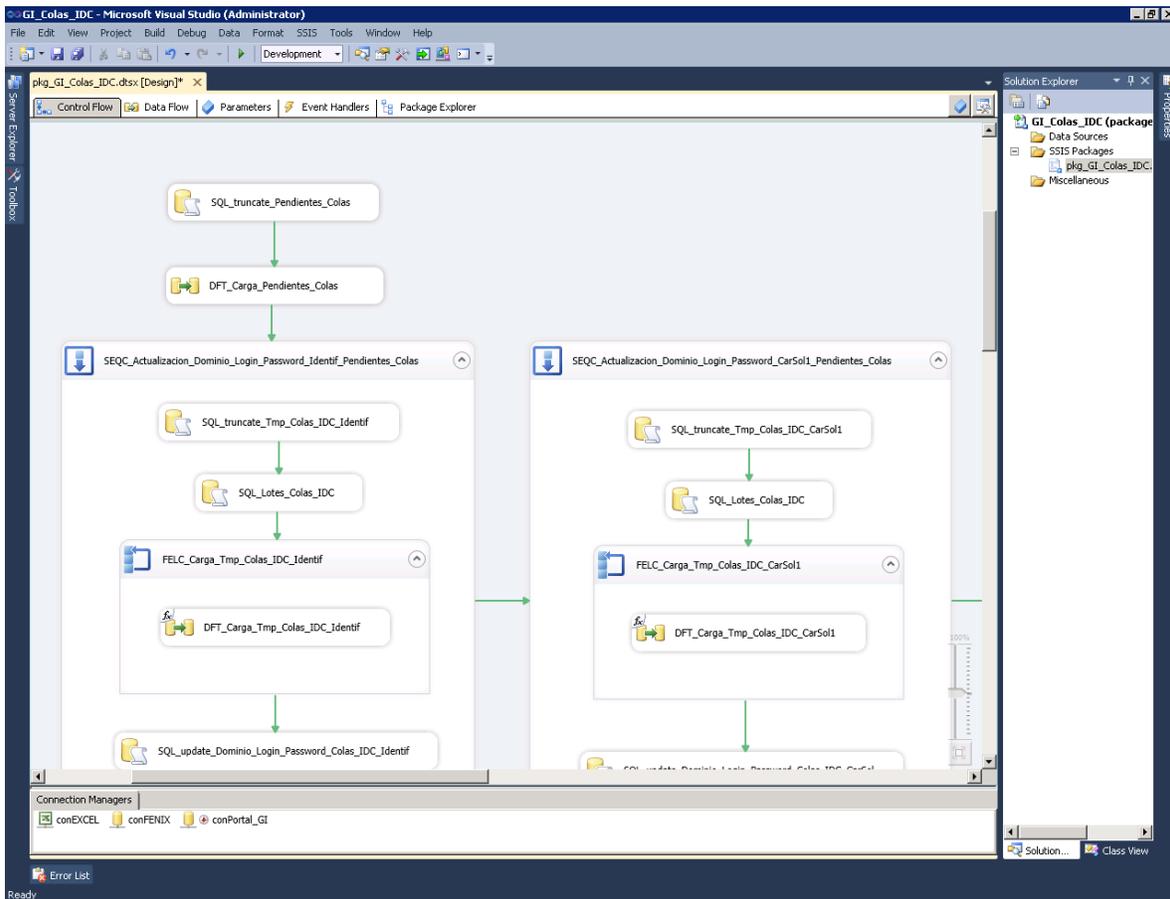


Ilustración 6: Migración de Access a SSIS

	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

Presentación de ETL completa,

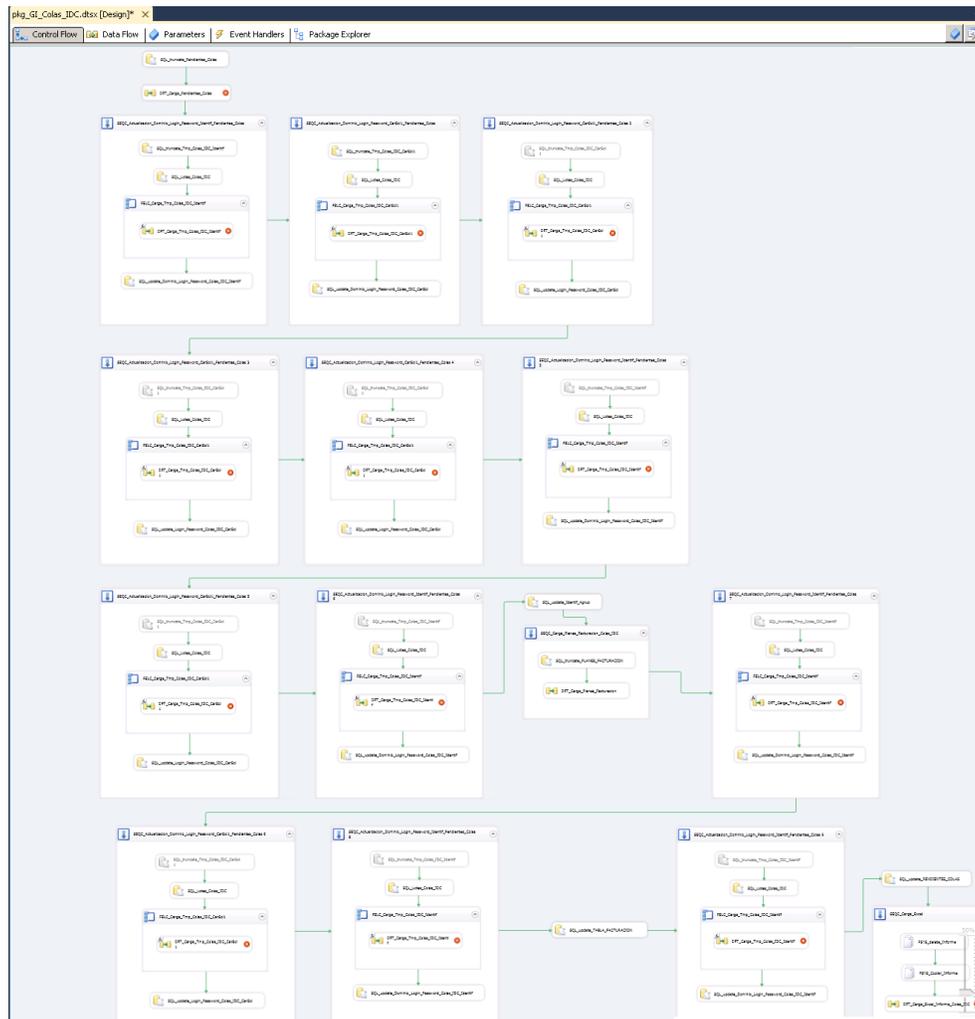


Ilustración 7: Migración de Access a SSIS, esquema completo

Variables generadas que se usaron en la ETL,

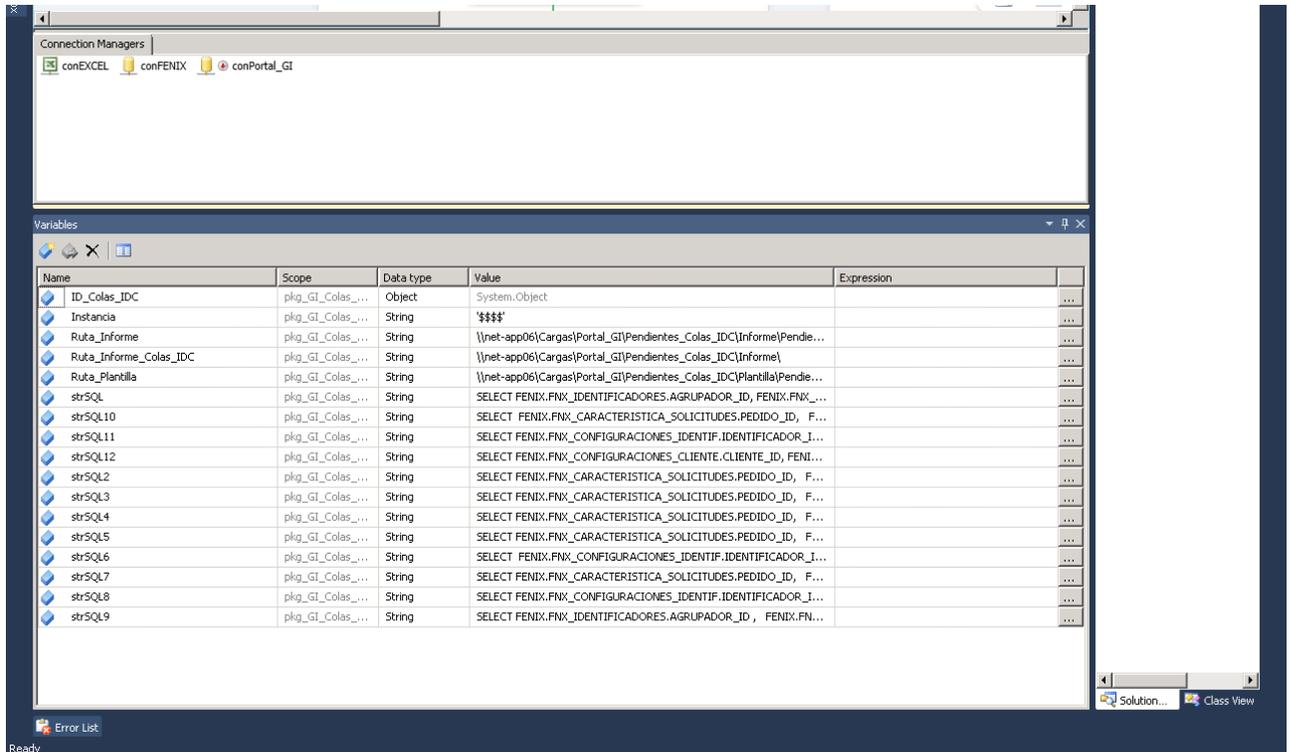


Ilustración 8: Variables usadas en ETL migrada

4.2. Otras funciones asignadas de TigoUne

Creación de reportes con Reporting Services,

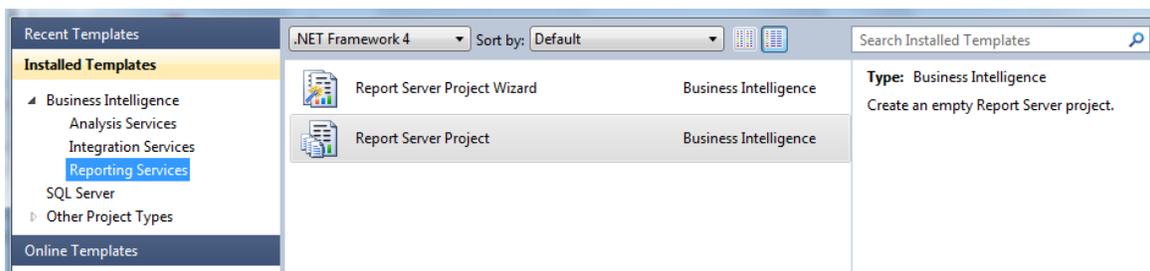


Ilustración 9: Creación de reporte

 Institución Universitaria	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

Se crean reportes para ejecución de Jobs (para medios magnéticos). Se crean dos rdl, donde “Parametrizar_Analytics_Medios_Magneticos_Open6.rdl” invoca al “Ejecutar_Analytics_Medios_Magneticos_Open6.rdl” y es ejecutado el job,

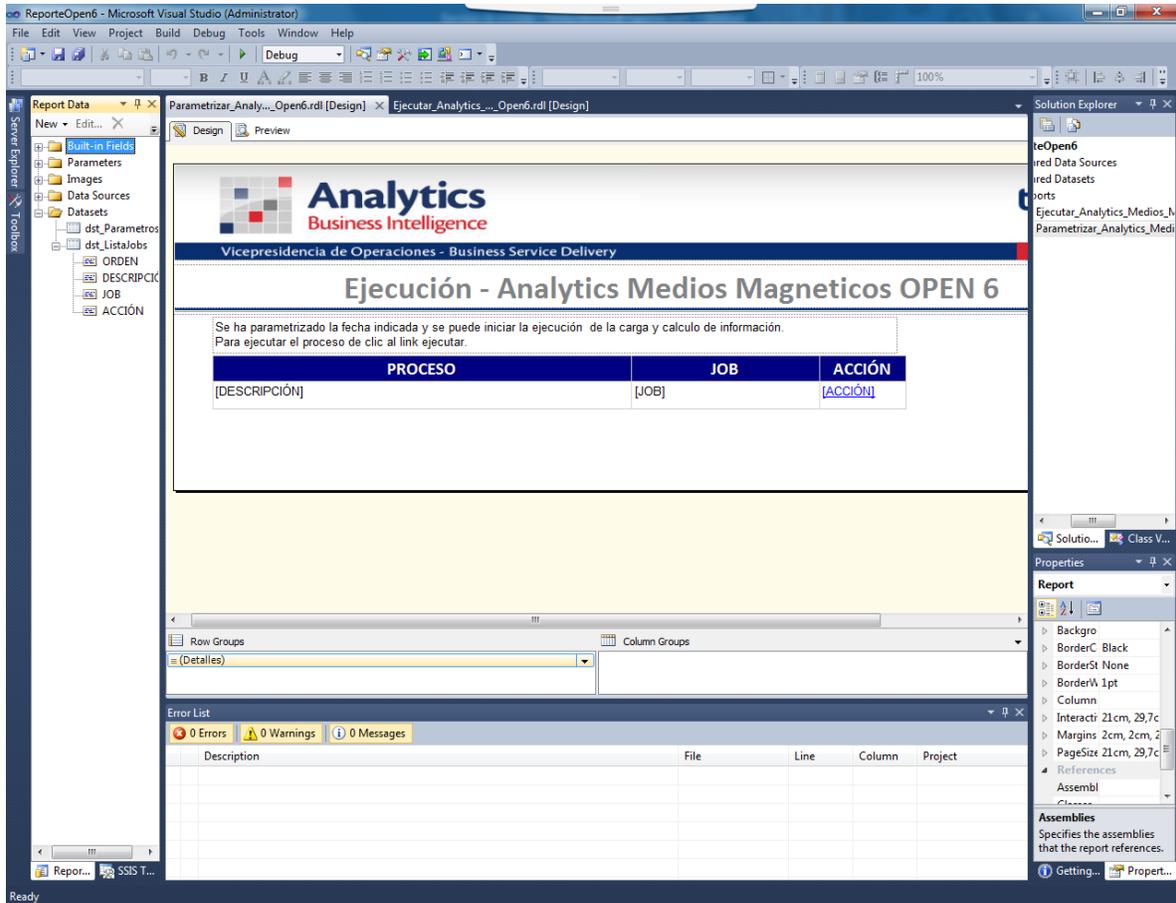


Ilustración 10: Reporte, Parametrizar_Analytics_Medios_Magneticos_Open6

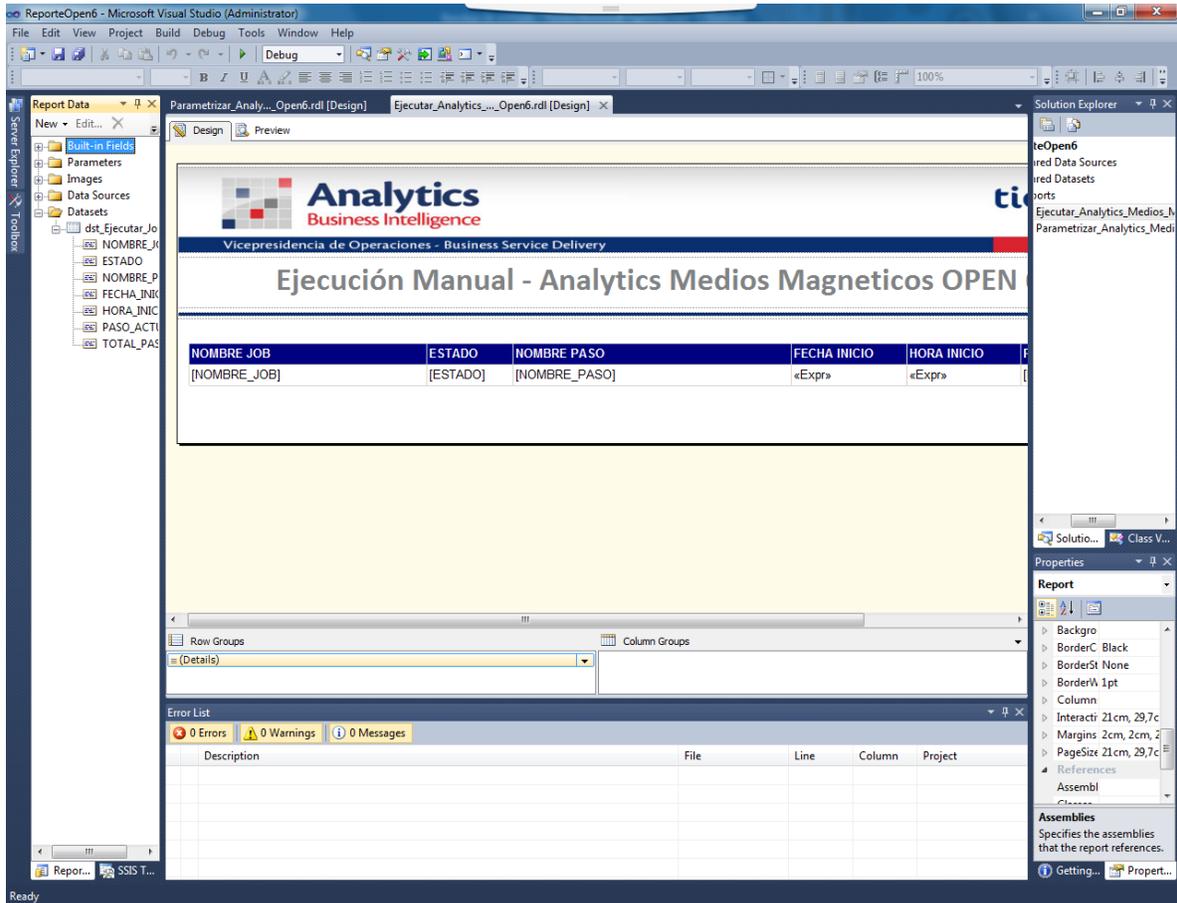


Ilustración 11: Reporte, Ejecutar_Analytics_Medios_Magneticos_Open6

Una vez se tiene configurado los reportes, se llevan al servidor de desarrollo para validar su correcto funcionamiento,

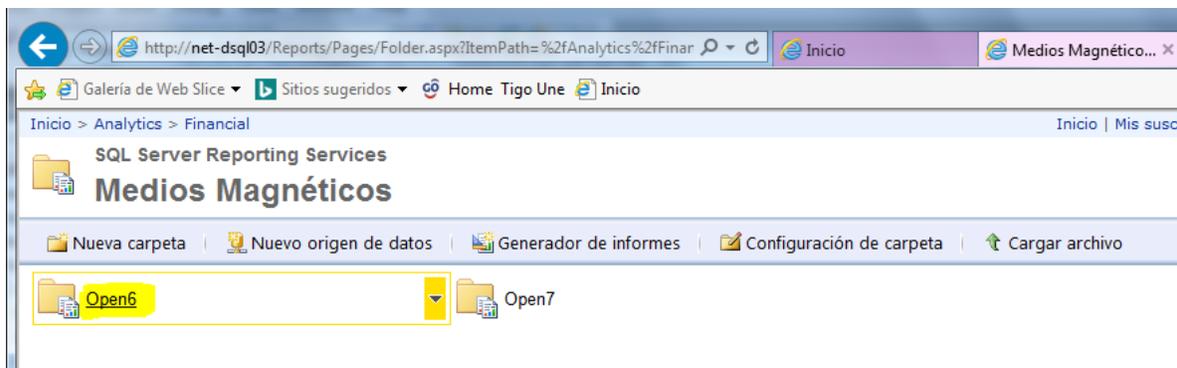


Ilustración 12: Ruta de Servidor de Reportes

	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

Se ejecuta “Parametrizar_Analytics_Medios_Magneticos_Open6.rdl” el cual invocará a “Ejecutar_Analytics_Medios_Magneticos_Open6.rdl”,

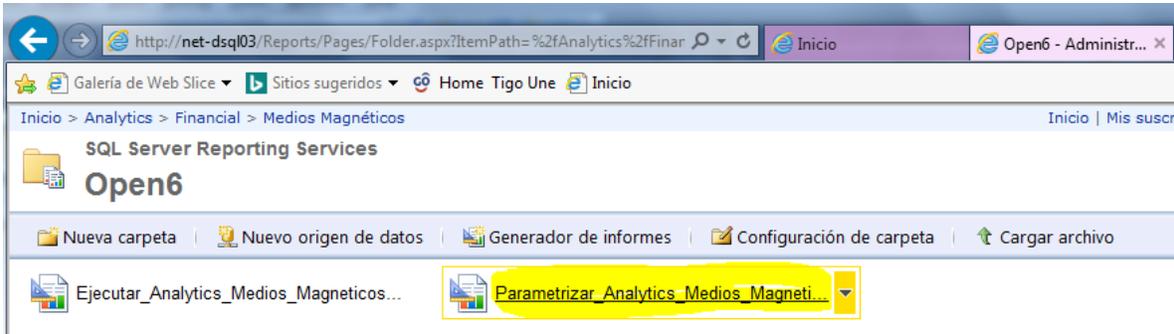


Ilustración 13: Reporte para ejecutar

Se presenta la siguiente pantalla,



Ilustración 14: Reporte en ejecución 1

Se ingresa la fecha solicitada y se selecciona el botón ver informe,



Ilustración 15: Reporte en ejecución 2

Se da clic en la acción Ejecutar y es iniciado el proceso; con esto se determina que los reportes fueron creados y configurados correctamente,

 Institución Universitaria	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22



Ilustración 16: Reporte ejecutado

4.2.1. Funciones (administrativas) en BI

- **Pasos a ambiente** de pruebas, este debe ser solicitado a través de Service Desk y se debe crear un Change Order,

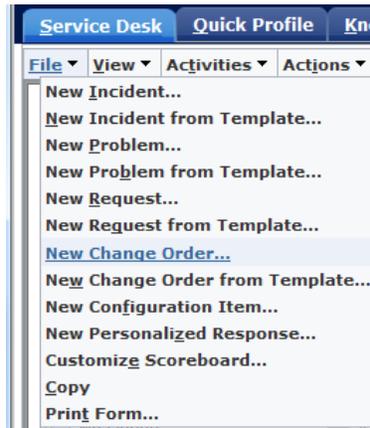


Ilustración 17: Creación de Change Order

Los campos que se encuentra marcados con rojo son los campos que se deben tener en cuenta para la generación de la OC,

1741032 Change Order Detail Edit(\$) Create Incident Quick Profile View Scheduler

Requester	Affected End User	Category	Status	Type
Bravo Londono, Leidy Johanna	Bravo Londono, Leidy Johanna	Cambios TIC Interno.Actualizacion de Objetos en Base de Datos UAT	RFC	
Priority	Impact	Risk	Causa Mayor	Active?
None			Mayor	YES

Detail

Created By	Assignee	Group	CAB
Bravo Londono, Leidy Johanna		Grupo Services Delivery	
Need By Date	Company	Organization	Call Back Date/Time
Closure Code	ID Catalogo	ID ITPAM	External_system_ticket
0			
Indisponibilidad	Root Causes		

Summary Information

Order Summary
 Por favor realizar despliegue en el servidor de pruebas NET-TSQL03 para los reportes de Medios Magnéticos. Solo deben ser ejecutados los objetos nombrados: "Parametrizar_Analytics_Medios_Magneticos_...", ya que este es el encargado de envi

Order Description
 Por favor realizar despliegue en el servidor de pruebas NET-TSQL03 para los reportes de Medios Magnéticos. Solo deben ser ejecutados los objetos nombrados: "Parametrizar_Analytics_Medios_Magneticos_...", ya que este es el encargado de enviar la variable y de invocar el reporte llamado "Ejecutar_Analytics_Medios_Magneticos_..."

Schedule Start Date	Schedule Duration	Schedule End Date
	00:00:00	
Open Date	Resolve Date	Close Date
01/25/2018 02:04 pm		

1. Additional Information | 2. Workflow Task | 3. Configuration Management | 4. Related Tickets | 5. Logs

1. Properties | 2. Service Type | 3. Attachments | 4. Conflicts | 5. Costs / Plans

Attach Document Attach URL(\$)

Attachment List

Ilustración 18: Campos importantes en la Creación de Change Order

1. Additional Information | 2. Workflow Task | 3. Configuration Management | 4. Related Tickets | 5. Logs

1. Properties | 2. Service Type | 3. Attachments | 4. Conflicts | 5. Costs / Plans

Attach Document Attach URL(\$)

Attachment List

1-2 of 2

Repository	Document	Description	Attached On	Status
files	20180125_01 Analytics_Medios_Magneticos- SQL.xls		01/25/2018 02:06 pm	Installed
files	OC_SSRS_20180125.zip		01/25/2018 02:06 pm	Installed

1-2 of 2

Ilustración 19: Documentos requeridos en la Creación de Change Order

 Institución Universitaria	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

En el campo **Requester**: Se ingresa el usuario solicitante de la OC.

En el campo **Category**: Se debe buscar siempre la categoría “Cambios TIC Interno.Actualizacion de Objetos en Base de Datos UAT”, que es la destinada para el paso de ambiente de desarrollo a pruebas.

En el campo **Organization**: Se borra lo que se genere, ya que no es necesario un valor en dicho campo.

En los campos **Order Summary** y **Order Description**: Se describe que se debe realizar, ejemplo: “Por favor realizar despliegue en el servidor de pruebas NET-TSQL03 para los reportes de Medios Magnéticos. Solo deben ser ejecutados los objetos nombrados: "Parametrizar_Analytics_Medios_Magneticos_....", ya que este es el encargado de enviar la variable y de invocar el reporte llamado "Ejecutar_Analytics_Medios_Magneticos_....".”.

En el campo **Attachment**: Se agregan los archivos necesarios para el paso a ambiente de pruebas, primero se carga un formato de Excel diligenciado y luego el .zip(Objetos y manuales).

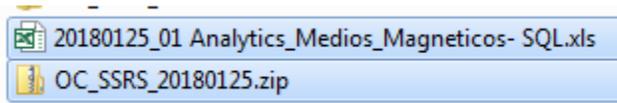
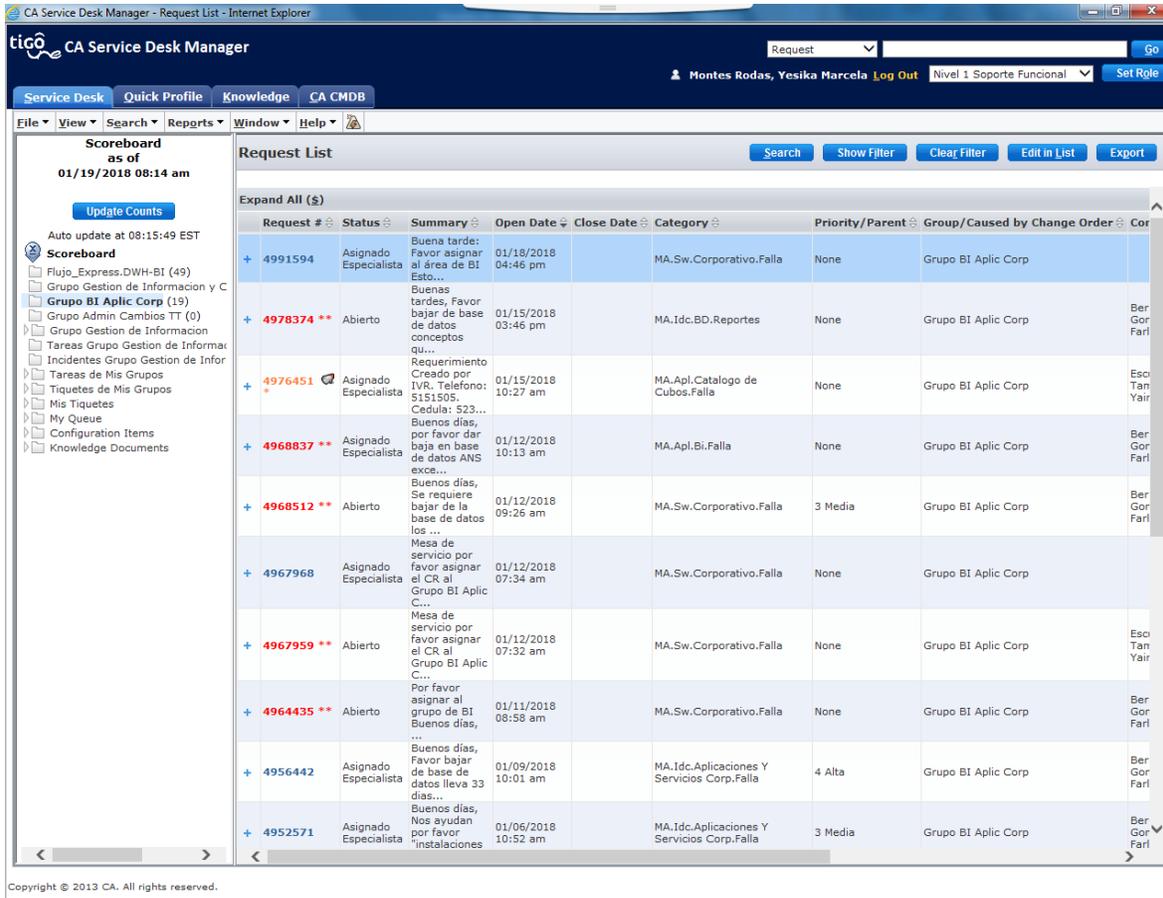


Ilustración 20: Documentos necesarios para Paso a Ambiente de Pruebas

Luego de esto se genera la OC y se debe esperar hasta que las personas encargadas del proceso realicen lo solicitado.

- **Asignación de incidentes, del grupo “Grupo BI Aplic Corp”**



The screenshot displays the CA Service Desk Manager interface. On the left, a 'Scoreboard' shows the date '01/19/2018 08:14 am' and a list of groups, with 'Grupo BI Aplic Corp' selected. The main area is titled 'Request List' and contains a table of open requests. The table columns are: Request #, Status, Summary, Open Date, Close Date, Category, Priority/Parent, Group/Caused by, Change Order, and Cor. The requests listed are for the 'Grupo BI Aplic Corp' group, with various summaries such as 'Buena tarde: Favor asignar al área de BI Esto...', 'Buenas tardes, Favor bajar de base de datos conceptos qu...', 'Requerimiento Creado por IVR. Telefono: 5151505. Cedula: 523...', 'Buenos dias, por favor dar baja en base de datos ANS exce...', 'Buenos dias, Se requiere bajar de la base de datos los ...', 'Mesa de servicio por favor asignar el CR al Grupo BI Aplic C...', 'Mesa de servicio por favor asignar el CR al Grupo BI Aplic C...', 'Por favor asignar al grupo de BI Buenos dias, ...', 'Buenos dias, Favor bajar de base de datos lleva 33 dias...', and 'Buenos dias, Nos ayudan por favor instalaciones'.

Request #	Status	Summary	Open Date	Close Date	Category	Priority/Parent	Group/Caused by	Change Order	Cor
4991594	Asignado Especialista	Buena tarde: Favor asignar al área de BI Esto...	01/18/2018 04:46 pm		MA.Sw.Corporativo.Falla	None	Grupo BI Aplic Corp		
4978374 **	Abierto	Buenas tardes, Favor bajar de base de datos conceptos qu...	01/15/2018 03:46 pm		MA.Idc.BD.Reportes	None	Grupo BI Aplic Corp		Ber Gor Farl
4976451 *	Asignado Especialista	Requerimiento Creado por IVR. Telefono: 5151505. Cedula: 523...	01/15/2018 10:27 am		MA.Apl.Catalogo de Cubos.Falla	None	Grupo BI Aplic Corp		Eso Tan Yair
4968837 **	Asignado Especialista	Buenos dias, por favor dar baja en base de datos ANS exce...	01/12/2018 10:13 am		MA.Apl.Bi.Falla	None	Grupo BI Aplic Corp		Ber Gor Farl
4968512 **	Abierto	Buenos dias, Se requiere bajar de la base de datos los ...	01/12/2018 09:26 am		MA.Sw.Corporativo.Falla	3 Media	Grupo BI Aplic Corp		Ber Gor Farl
4967968	Asignado Especialista	Mesa de servicio por favor asignar el CR al Grupo BI Aplic C...	01/12/2018 07:34 am		MA.Sw.Corporativo.Falla	None	Grupo BI Aplic Corp		
4967959 **	Abierto	Mesa de servicio por favor asignar el CR al Grupo BI Aplic C...	01/12/2018 07:32 am		MA.Sw.Corporativo.Falla	None	Grupo BI Aplic Corp		Eso Tan Yair
4964435 **	Abierto	Por favor asignar al grupo de BI Buenos dias, ...	01/11/2018 08:58 am		MA.Sw.Corporativo.Falla	None	Grupo BI Aplic Corp		Ber Gor Farl
4956442	Asignado Especialista	Buenos dias, Favor bajar de base de datos lleva 33 dias...	01/09/2018 10:01 am		MA.Idc.Aplicaciones Y Servicios Corp.Falla	4 Alta	Grupo BI Aplic Corp		Ber Gor Farl
4952571	Asignado Especialista	Buenos dias, Nos ayudan por favor instalaciones	01/06/2018 10:52 am		MA.Idc.Aplicaciones Y Servicios Corp.Falla	3 Media	Grupo BI Aplic Corp		Ber Gor Farl

Ilustración 21: Incidentes reportados para asignación

Se valida que solicitan en el incidente para realizar la correspondiente asignación, ya que cada compañero tiene a cargo diferentes temas y a mí me corresponde asignar permisos en grupos específicos en el directorio activo y en algunas rutas de carpetas.

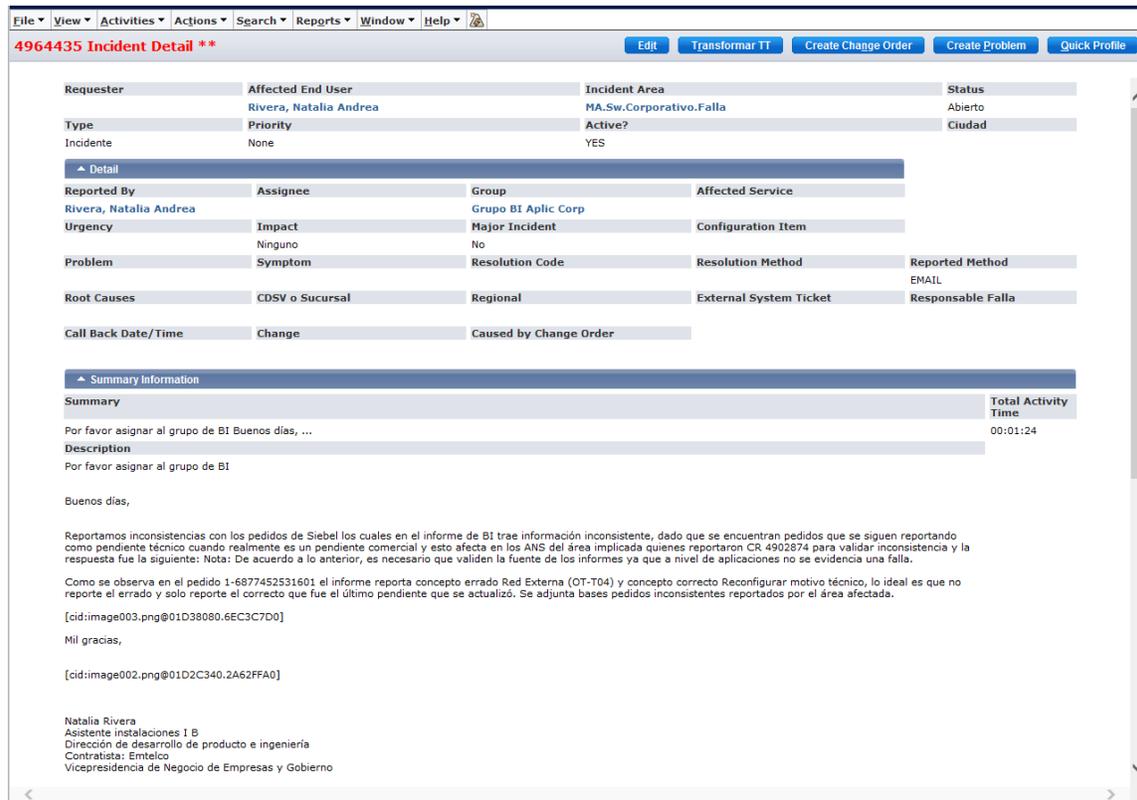


Ilustración 22: Detalle de un Incidente

Ya sabiendo a quien se debe asignar el incidente o CR, se transfiere a la persona encargada del tema

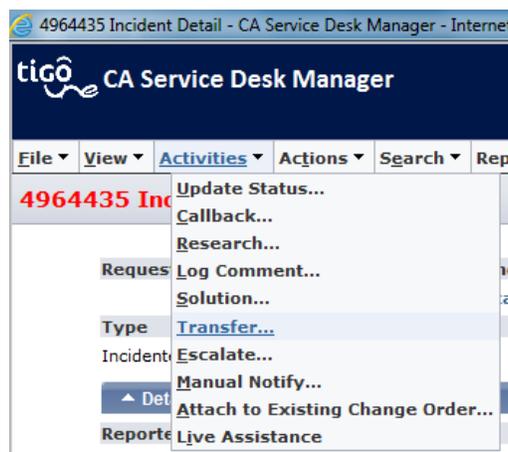
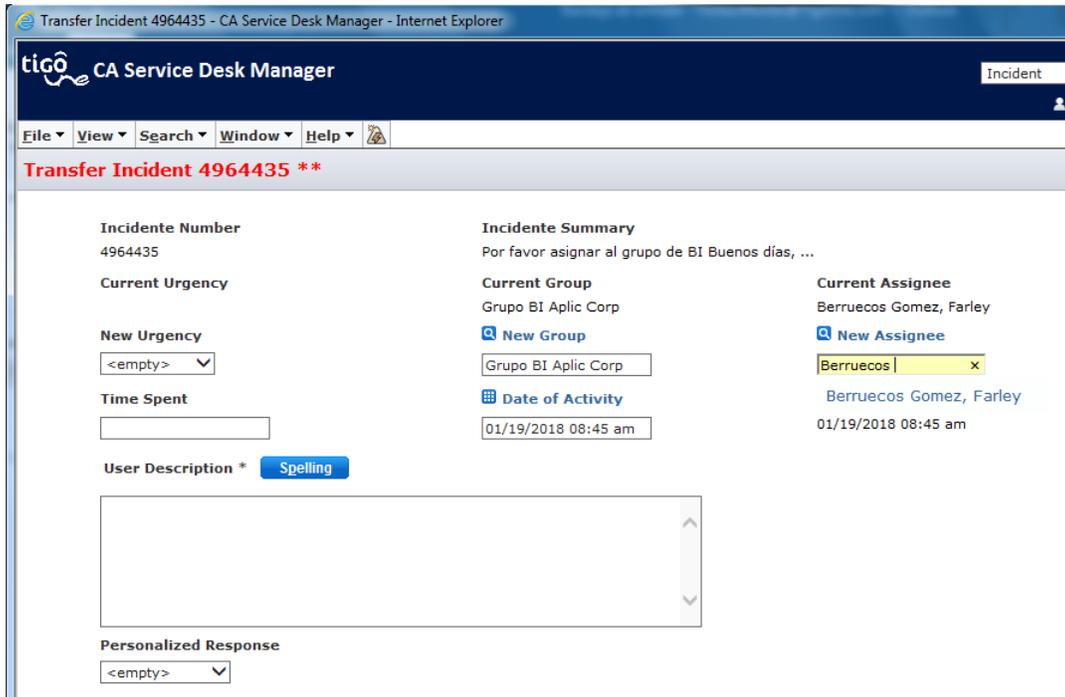


Ilustración 23: Opción de Transferencia de un Incidente

 Institución Universitaria	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

Se selecciona la persona encargada, se da una descripción para la asignación y se guarda

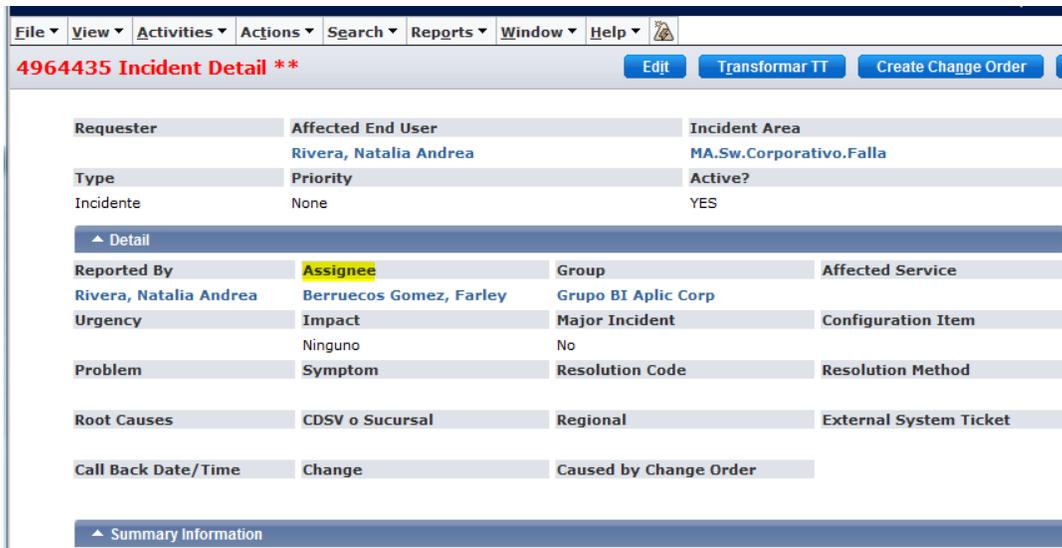


The screenshot shows the 'Transfer Incident 4964435' form in the CA Service Desk Manager. The form includes the following fields and values:

- Incidente Number:** 4964435
- Incidente Summary:** Por favor asignar al grupo de BI Buenos días, ...
- Current Urgency:** (empty)
- Current Group:** Grupo BI Aplic Corp
- Current Assignee:** Berruecos Gomez, Farley
- New Urgency:** <empty>
- New Group:** Grupo BI Aplic Corp
- New Assignee:** Berruecos Gomez, Farley
- Time Spent:** (empty)
- Date of Activity:** 01/19/2018 08:45 am
- User Description *:** (empty text area with a 'Spelling' button)
- Personalized Response:** <empty>

Ilustración 24: Asignar persona para Transferencia de Incidente

Una vez asignada la persona al CR, se verá reflejada esta asignación en el campo "Assignee"



The screenshot shows the '4964435 Incident Detail' view. The 'Assignee' field is highlighted in yellow, indicating the assigned person. The details are as follows:

- Requester:** Affected End User: Rivera, Natalia Andrea; Incident Area: MA.Sw.Corporativo.Falla
- Type:** Priority: None; Active?: YES
- Reported By:** Rivera, Natalia Andrea; **Assignee:** Berruecos Gomez, Farley; **Group:** Grupo BI Aplic Corp; **Affected Service:** Configuration Item
- Urgency:** Impact: Ninguno; **Major Incident:** No
- Problem:** Symptom; **Resolution Code:** (empty); **Resolution Method:** (empty)
- Root Causes:** CDSV o Sucursal; **Regional:** (empty); **External System Ticket:** (empty)
- Call Back Date/Time:** Change; **Caused by Change Order:** (empty)

Ilustración 25: Transferencia realizada

 Institución Universitaria	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

- Cuando me corresponde realizar una asignación o retiro de permisos, se realiza lo siguiente, Estos se ingresan a través del catálogo de servicios,

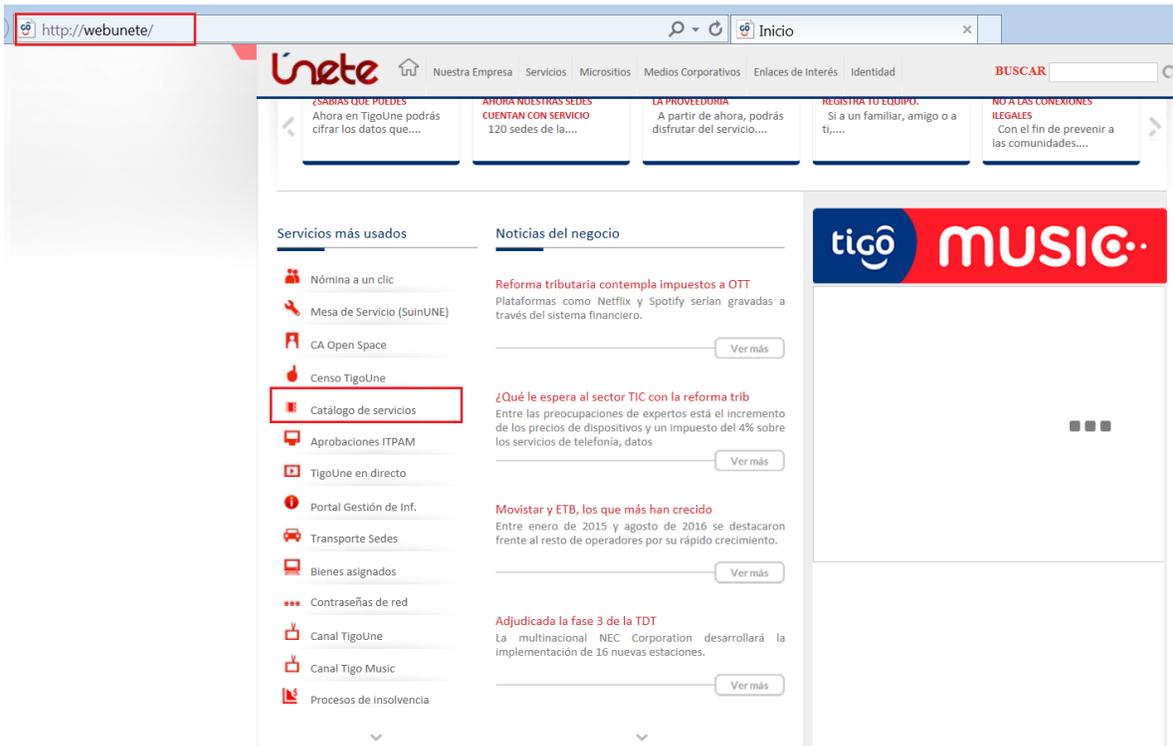
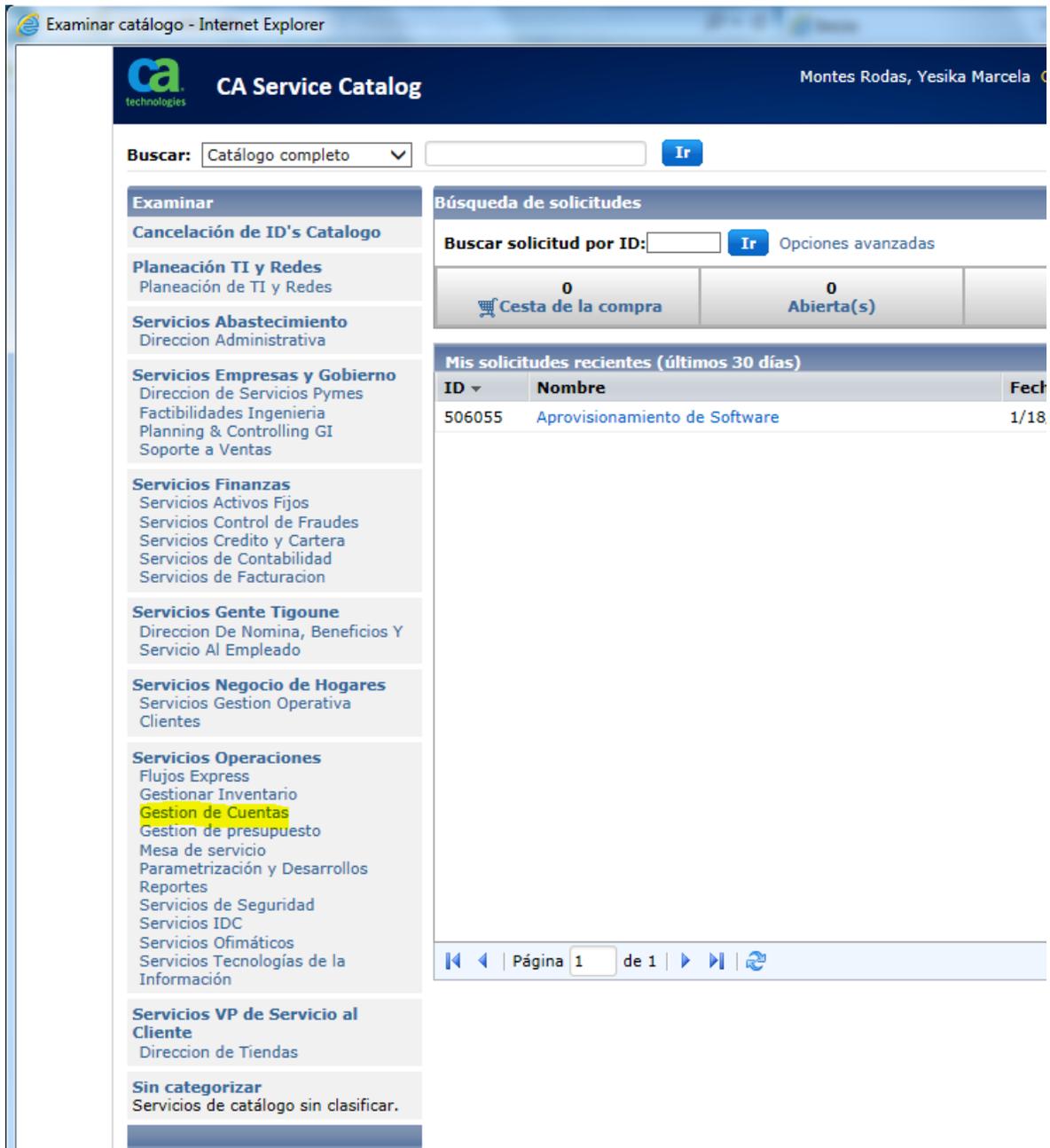


Ilustración 26: Opción de Catalogo de Servicios

 Institución Universitaria	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

Cuando se realice una solicitud de permisos con grupo de directorio activo, se selecciona la opción “Gestión de Cuentas”



The screenshot shows the CA Service Catalog interface. The left sidebar contains a list of service categories. The 'Gestión de Cuentas' option is highlighted in yellow. The main content area shows search results for 'Aprovisionamiento de Software' with ID 506055.

CA Service Catalog (Montes Rodas, Yesika Marcela)

Buscar:

Examinar

- Cancelación de ID's Catalogo**
- Planeación TI y Redes**
Planeación de TI y Redes
- Servicios Abastecimiento**
Direccion Administrativa
- Servicios Empresas y Gobierno**
Direccion de Servicios Pymes
Factibilidades Ingenieria
Planning & Controlling GI
Soporte a Ventas
- Servicios Finanzas**
Servicios Activos Fijos
Servicios Control de Fraudes
Servicios Credito y Cartera
Servicios de Contabilidad
Servicios de Facturacion
- Servicios Gente Tigoune**
Direccion De Nomina, Beneficios Y
Servicio Al Empleado
- Servicios Negocio de Hogares**
Servicios Gestion Operativa
Clientes
- Servicios Operaciones**
Flujos Express
Gestionar Inventario
Gestion de Cuentas
Gestion de presupuesto
Mesa de servicio
Parametrización y Desarrollos
Reportes
Servicios de Seguridad
Servicios IDC
Servicios Ofimáticos
Servicios Tecnologías de la Información
- Servicios VP de Servicio al Cliente**
Direccion de Tiendas
- Sin categorizar**
Servicios de catálogo sin clasificar.

Búsqueda de solicitudes

Buscar solicitud por ID: Opciones avanzadas

0 Cesta de la compra 0 Abierta(s)

Mis solicitudes recientes (últimos 30 días)

ID	Nombre	Fecha
506055	Aprovisionamiento de Software	1/18

« | « | Página 1 de 1 | » | » | ↻

Ilustración 27: Opción para asignación o retiro de permisos

	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

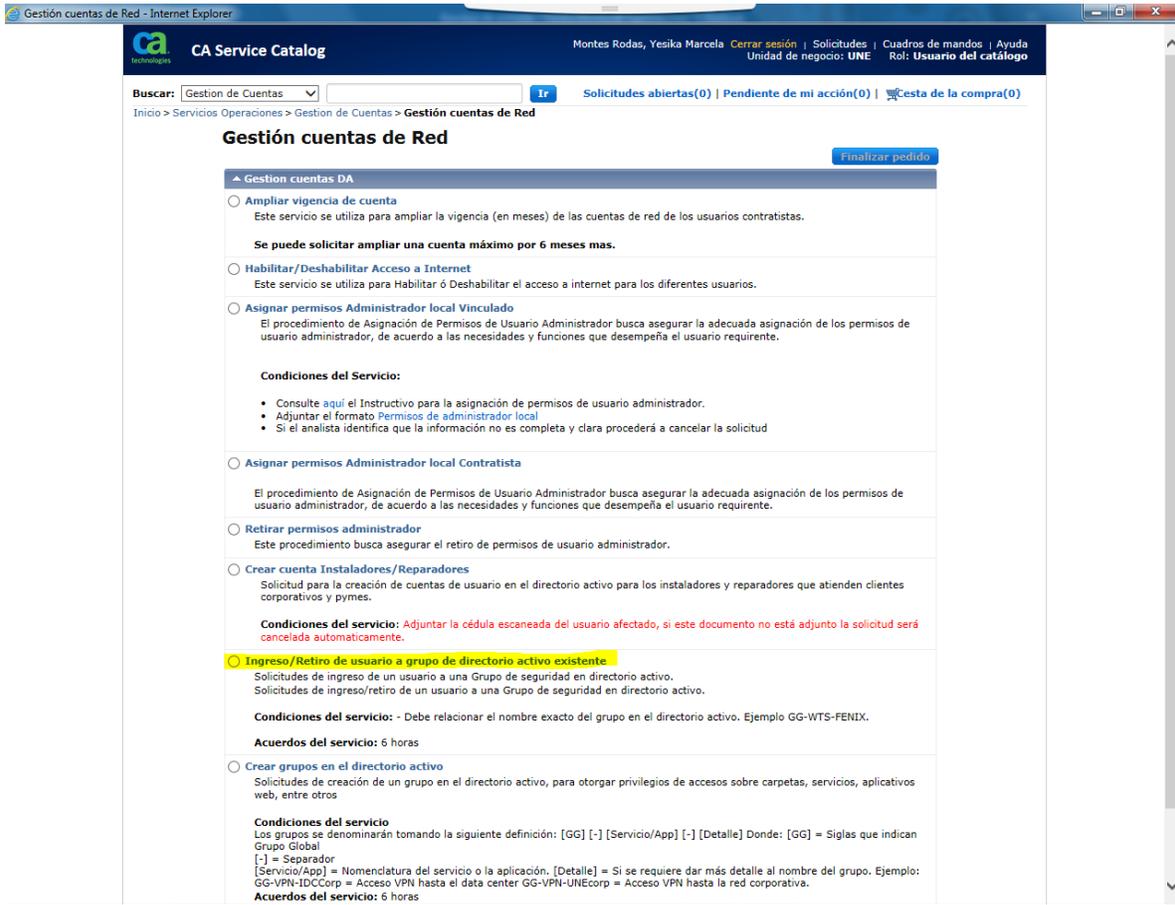
Luego se selecciona “Gestión cuentas de Red”



Ilustración 28: Opción para gestión de cuentas de Red

En el formulario que se abre, seleccionar la opción “Ingreso/Retiro de usuario a grupo de directorio activo existente”

 Institución Universitaria	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22



CA Service Catalog

Montes Rodas, Yesika Marcela Cerrar sesión | Solicitudes | Cuadros de mandos | Ayuda
 Unidad de negocio: UNE Rol: Usuario del catálogo

Buscar: Gestión de Cuentas Ir Solicitudes abiertas(0) | Pendiente de mi acción(0) | Cesta de la compra(0)

Inicio > Servicios Operaciones > Gestión de Cuentas > **Gestión cuentas de Red**

Gestión cuentas de Red

[finalizar pedido](#)

Gestión cuentas DA

- Ampliar vigencia de cuenta**
 Este servicio se utiliza para ampliar la vigencia (en meses) de las cuentas de red de los usuarios contratistas.
Se puede solicitar ampliar una cuenta máximo por 6 meses mas.
- Habilitar/Deshabilitar Acceso a Internet**
 Este servicio se utiliza para Habilitar ó Deshabilitar el acceso a internet para los diferentes usuarios.
- Asignar permisos Administrador local Vinculado**
 El procedimiento de Asignación de Permisos de Usuario Administrador busca asegurar la adecuada asignación de los permisos de usuario administrador, de acuerdo a las necesidades y funciones que desempeña el usuario requirente.
Condiciones del Servicio:
 - Consulte [aquí](#) el Instructivo para la asignación de permisos de usuario administrador.
 - Adjuntar el formato [Permisos de administrador local](#)
 - Si el analista identifica que la información no es completa y clara procederá a cancelar la solicitud
- Asignar permisos Administrador local Contratista**
 El procedimiento de Asignación de Permisos de Usuario Administrador busca asegurar la adecuada asignación de los permisos de usuario administrador, de acuerdo a las necesidades y funciones que desempeña el usuario requirente.
- Retirar permisos administrador**
 Este procedimiento busca asegurar el retiro de permisos de usuario administrador.
- Crear cuenta Instaladores/Reparadores**
 Solicitud para la creación de cuentas de usuario en el directorio activo para los instaladores y reparadores que atienden clientes corporativos y pymes.
Condiciones del servicio: Adjuntar la cédula escaneada del usuario afectado, si este documento no está adjunto la solicitud será cancelada automáticamente.
- Ingreso/Retiro de usuario a grupo de directorio activo existente**
 Solicitudes de ingreso de un usuario a una Grupo de seguridad en directorio activo.
 Solicitudes de ingreso/retiro de un usuario a una Grupo de seguridad en directorio activo.
Condiciones del servicio: - Debe relacionar el nombre exacto del grupo en el directorio activo. Ejemplo GG-WTS-FENIX.
Acuerdos del servicio: 6 horas
- Crear grupos en el directorio activo**
 Solicitudes de creación de un grupo en el directorio activo, para otorgar privilegios de accesos sobre carpetas, servicios, aplicativos web, entre otros
Condiciones del servicio
 Los grupos se denominarán tomando la siguiente definición: [GG] [-] [Servicio/App] [-] [Detalle] Donde: [GG] = Siglas que indican Grupo Global
 [-] = Separador
 [Servicio/App] = Nomenclatura del servicio o la aplicación. [Detalle] = Si se requiere dar más detalle al nombre del grupo. Ejemplo: GG-VPN-IDCCorp = Acceso VPN hasta el data center GG-VPN-UNECorp = Acceso VPN hasta la red corporativa.
Acuerdos del servicio: 6 horas

Ilustración 29: Opción Ingreso/Retiro de usuarios

Y se habilita un nuevo formulario,

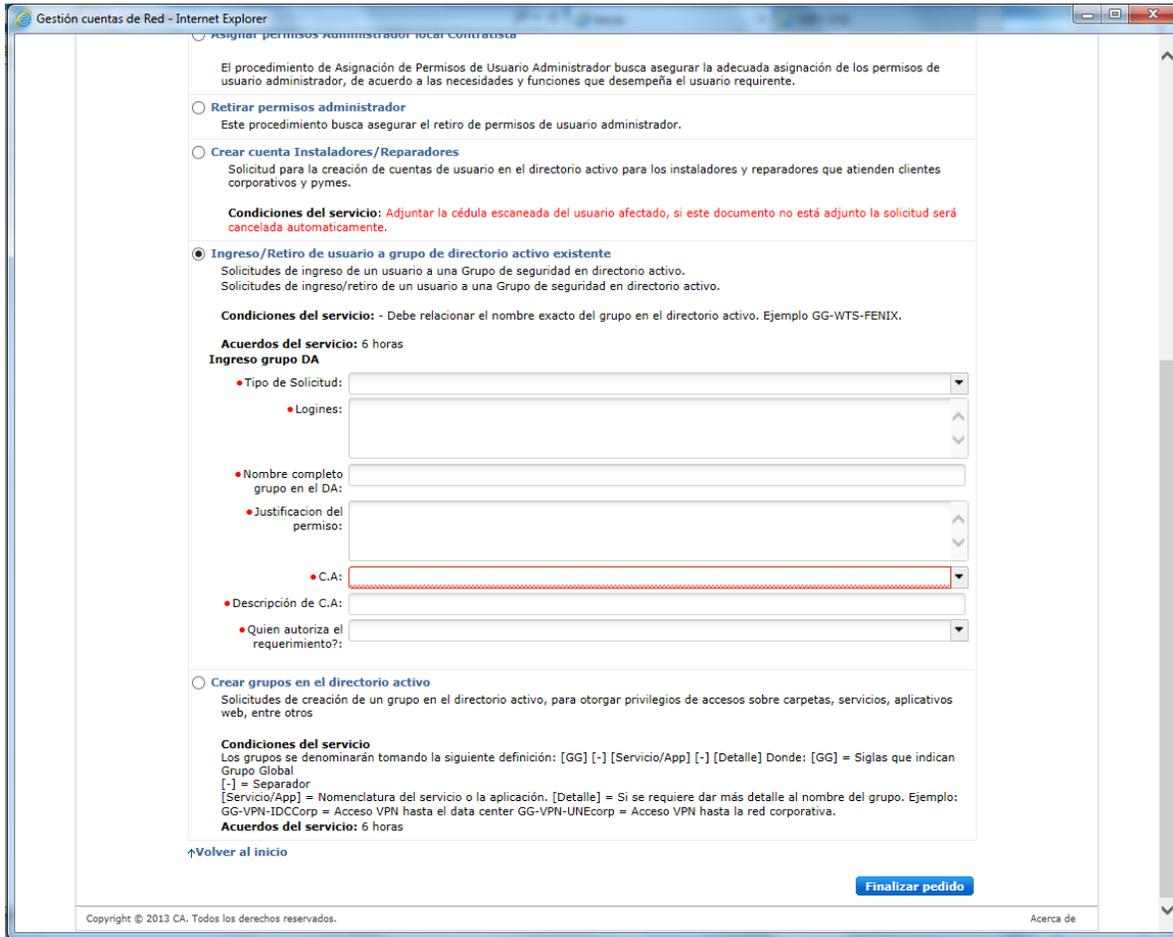


Ilustración 30: Formulario para Ingreso/Retiro de usuarios

Se debe tener en cuenta si es una solicitud de acceso o de retiro, para generar permiso

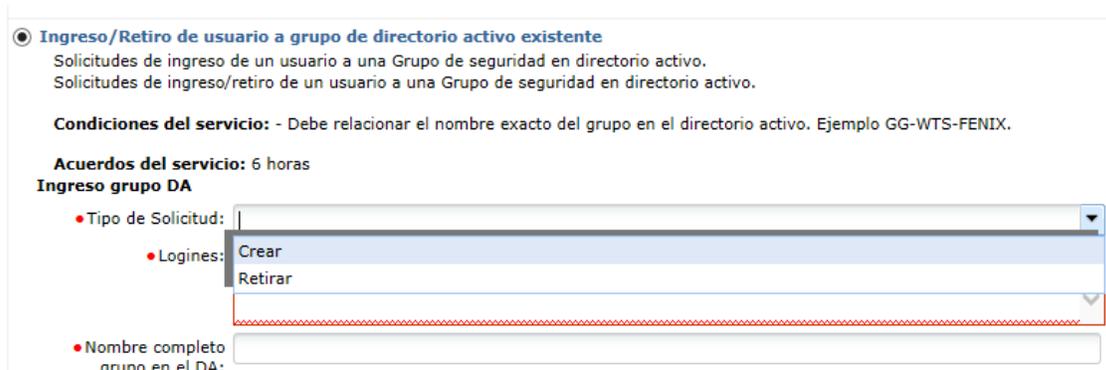


Ilustración 31: Lista de Opciones Crear/Retirar

 Institución Universitaria	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

Se termina de ingresar y seleccionar la información de los demás campos y se da clic en el botón “Finalizar pedido”,

Ingreso/Retiro de usuario a grupo de directorio activo existente
 Solicitudes de ingreso de un usuario a una Grupo de seguridad en directorio activo.
 Solicitudes de ingreso/retiro de un usuario a una Grupo de seguridad en directorio activo.

Condiciones del servicio: - Debe relacionar el nombre exacto del grupo en el directorio activo. Ejemplo GG-WTS-FENIX.

Acuerdos del servicio: 6 horas
Ingreso grupo DA

- Tipo de Solicitud:
- Logines:
- Nombre completo grupo en el DA:
- Justificación del permiso:
- C.A:
- Descripción de C.A:
- Quien autoriza el requerimiento?:

Crear grupos en el directorio activo
 Solicitudes de creación de un grupo en el directorio activo, para otorgar privilegios de accesos sobre carpetas, servicios, aplicativos web, entre otros

Condiciones del servicio
 Los grupos se denominarán tomando la siguiente definición: [GG] [-] [Servicio/App] [-] [Detalle] Donde: [GG] = Siglas que indican Grupo Global
 [-] = Separador
 [Servicio/App] = Nomenclatura del servicio o la aplicación. [Detalle] = Si se requiere dar más detalle al nombre del grupo. Ejemplo: GG-VPN-IDCCorp = Acceso VPN hasta el data center GG-VPN-UNEcorp = Acceso VPN hasta la red corporativa.

Acuerdos del servicio: 6 horas

[↑Volver al inicio](#)

[Finalizar pedido](#)

Ilustración 32: Diligenciamiento Formulario Ingreso/retiro de usuario

Luego se da en una opción de “Guardar y enviar a la cesta de compra” y queda pendiente para aprobación de quien autoriza.

Una vez lo den por aprobado, se genera un numero para buscarlo en service desk

 Institución Universitaria	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

5. CONCLUSIONES, RECOMENDACIONES Y TRABAJO FUTURO

5.1. Conclusiones

Desde el punto de vista teórico la integración de datos es el proceso que combina la información desde múltiples fuentes para lograr una vista unificada. Además, debe responder a diferentes procesos, tales como la migración, la sincronización, el almacenamiento de datos (Data Warehousing), entre otros; lo que debe ser posible de lograr a través de las herramientas que implementan la integración.

El SQL Server Integration Services es una herramienta que cumple con los aspectos señalados anteriormente, pues brinda una gran variedad de elementos para lograr soluciones complejas en estos temas, auxiliando a los usuarios con un entorno de desarrollo integrado que intenta que el proceso sea óptimo y fácil de implementar, dando amplísimas posibilidades de conexión a fuentes y destinos de datos muy diversos en tipo e incluso en plataformas que los soportan, expandiendo así la escala de soluciones, para lograr que el proceso de extracción, transformación y carga de los datos se realice con éxito.

En el desarrollo del trabajo se da cumplimiento al objetivo general y los objetivos específicos, ya que se logra después del estudio de la herramienta realizar y documentar la solución de ETL que requería migrar y optimizar el proceso de extraer y publicar las solicitudes pendientes por aprovisionar en Fenix Nal de los productos asociados a IDC.

	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

5.2. Recomendaciones

Para darle continuidad al trabajo comenzado en este proyecto de grado planteamos las siguientes recomendaciones:

- Continuar al proceso de estudio del SSIS para profundizar en otros aspectos.
- Comparar el servicio de integración de datos del SQL Server con otras herramientas que sirven a este propósito para poder sacar conclusiones de cuál es más factible usar en determinadas condiciones de integración de datos.

5.3. Trabajo Futuro

Se recomienda realizar otros temas de proyectos de grado derivados de este tema agregando más fuentes de datos, utilización de transformadores de herramientas ETL, ya que el campo de estas es muy amplio y necesario para la integración de datos.

	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

REFERENCIAS

ALKIS SIMITSIS NATIONAL TECHNICAL UNIVERSITY OF ATHENS, G. (G. 2005), PANOS VASSILIADIS UNIVERSITY OF IOANNINA, G. & TIMOS SELLIS NATIONAL TECHNICAL UNIVERSITY OF ATHENS, G. Extraction, Transformation, Loading Processes.

ALKIS SIMITSIS, P. V. A. T. S. (2005). Extraction- Transformation-Loading Process. In: INC, I. G. (ed.).

CARROLL, M. (2005). Designing for parallelism whit Integration Service.

HAAS, P. A. B. A. L. M. (2008). A guide to the tools and core technologies for merging information from disparate sources, communications of the acm.

HALEVY, A. Y., ASHISH, N., BITTON, D., CAREY, M.J., DRAPER, D., POLLOCK, J., ROSENTHAL, A., AND SIKKA, V.(2005). Enterprise information integration: Successes, challenges, and controversies. In: SIGMOD, I. P. O. T. A., ed. (2005). 778-787.

ING. BERNABEU R. DARIO, C. D. L. R. O. B. (2009). Load manager [Online]. Available: <http://www.dataprix.com/33-load-manager>.

JAVA.NET. (2013). Data integration [Online]. java.net. Available: <http://wiki.opensb.java.net/Wiki.jsp?page=DataIntegration>.

KIMBALL, R. A. C., J (2004). The Data Warehouse ETL Toolkit.

LANDSCAPE, T. N. D. I. (2013) Moving beyond ad_hoc ETL to an enterprise data integration strategy.

	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

LANE, F. (2006). IDC: World Created 161 Billion Gigs of Data in 2006.

LENZERINI, M. (2002). Data Integration: A Theoretical Perspective, 2002.

LÓPEZ-PORRERO, B. (2011). Limpieza de datos: reemplazo de valores ausentes y estandarización. Investigación, Universidad Central Martha Abreus de la Villas.

MELNIK, S., ADYA, A., AND BERNSTEIN, P.A. (2007). Compiling mappings to bridge applications and databases. In: SIGMOD, I. P. O. T. A., ed., 2007. 461–472.

PRIX, D. (2010). Herramientas ETL. ¿Que son, para que valen?. Productos mas conocidos. ETL's Open Source. [Online]. Available:
http://www.dataprix.com/category/categorias/Herramientas_ETL_Que_son_para_que_valen_Productos_mas_conocidos_ETL's_Open_Source.

WIKIPEDIA. (2012). Extract, transform and load [Online]. Available:
https://es.wikipedia.org/wiki/Extract,_transform_and_load.

WIKIPEDIA. (2012a). ETL [Online]. Available: <http://es.wikipedia.org/wiki/ETL>.

WIKIPEDIA. (2012b). Migración de Datos [Online].

 Institución Universitaria	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

APÉNDICE
