

ANÁLISIS DE LOS PROCESOS DE FACTURACIÓN E INVENTARIOS EN LA
PRESTACIÓN DE SERVICIOS DE *NETWORKING* MEDIANTE INTELIGENCIA DE
NEGOCIOS

RAMÓN OCTAVIO MARIÑO PARRA

UNIVERSIDAD PONTIFICIA BOLIVARIANA

ESCUELA INGENIERÍAS

FACULTAD DE INGENIERÍA EN TECNOLOGÍAS DE INFORMACIÓN Y
COMUNICACIÓN

MAESTRÍA EN TECNOLOGÍAS DE INFORMACIÓN Y COMUNICACIÓN

MEDELLÍN

2019

ANÁLISIS DE LOS PROCESOS DE FACTURACIÓN E INVENTARIOS EN LA
PRESTACIÓN DE SERVICIOS DE *NETWORKING* MEDIANTE INTELIGENCIA DE
NEGOCIOS

RAMÓN OCTAVIO MARIÑO PARRA

Trabajo de grado para optar al título de

Magíster en Tecnologías de la Información y la Comunicación

Asesor

IVÁN AMON URIBE

Magíster en ingeniería de Sistemas

UNIVERSIDAD PONTIFICIA BOLIVARIANA

ESCUELA INGENIERÍAS

FACULTAD DE INGENIERÍA EN TECNOLOGÍAS DE INFORMACIÓN Y
COMUNICACIÓN

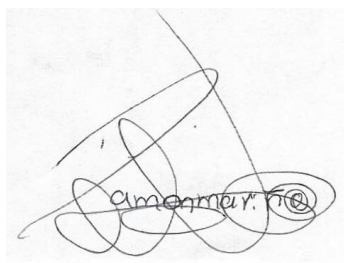
MAESTRÍA EN TECNOLOGÍAS DE INFORMACIÓN Y COMUNICACIÓN

MEDELLÍN

2019

DECLARACIÓN ORIGINALIDAD

“Declaro que esta tesis (o trabajo de grado) no ha sido presentada para optar a un título, ya sea en igual forma o con variaciones, en esta o cualquier otra universidad”. Art. 82 Régimen Discente de Formación Avanzada, Universidad Pontificia Bolivariana.

A handwritten signature in black ink, appearing to read 'C. M. M. M. M.', is written over a horizontal line. The signature is somewhat stylized and partially obscured by the line.

FIRMA AUTOR (ES)

Medellín Noviembre 12 de 2019

AGRADECIMIENTOS

Agradezco a mi familia por el incondicional y fraternal apoyo en cada momento de la elaboración de este proyecto.

Tabla de contenido

Tabla de contenido.....	V
1. INTRODUCCIÓN	13
1. PLANTEAMIENTO DEL PROBLEMA.....	15
1.1 Problema.....	15
2. JUSTIFICACIÓN.....	17
3. OBJETIVOS.....	19
3.1 Objetivo General.....	19
3.2 Objetivos Específicos	19
4. MARCO REFERENCIAL	20
4.1 Marco contextual.....	20
4.2 Marco conceptual.....	21
4.2.1 Inteligencia de negocios (BI).....	21
4.2.2 Soluciones de red Cisco.	25
4.2.3 <i>Secure Shell</i> (SSH)	26
4.2.4 Lenguaje Visual Basic Script (VBS)	27
4.2.5 Visualización de datos	27
4.2.6 Protocolo SNMP.....	27
4.3 Marco legal.....	28
4.4 Estado del arte.....	30
5. METODOLOGÍA.....	39

5.1 Desarrollo del proyecto basado las metodologías planteadas.....	42
6. PRESENTACIÓN Y ANÁLISIS DE RESULTADOS	45
6.1 Requerimientos de información del negocio	45
6.1.1 Entendimiento del negocio	45
6.1.2 Elaboración de <i>startnets</i>	47
6.1.3 Fuentes de los datos.	48
6.1.4 Especificaciones de aplicaciones	49
6.2 Construcción del modelo dimensional.....	51
6.2.1 Modelo lógico y dimensional.....	51
6.2.2 Diseño físico	54
6.2.3 Carga inicial de los datos.....	55
6.3 Desarrollo de aplicación de extracción de información	56
6.3.1 Diseño e implementación de ETL.....	56
6.3.1.1 Desarrollo de la interfaz gráfica	57
6.3.1.2 Conexión a la base de datos.....	59
6.3.1.3 Ingreso a los dispositivos de interés	59
6.3.1.4 Captura de información en variables locales	60
6.3.1.5 Escritura en la base de datos	61
6.3.1.6 Validación de datos.	61
6.3.1.7 Verificación.....	61
6.3.1.8 Mantenimiento.	63
6.4 Análisis de calidad de los datos.....	64

6.4.1	Análisis de calidad de datos Dimensión Puertos.....	64
6.4.2	Análisis de calidad de datos de la dimensión RT_PUERTOS.....	65
6.4.3	Análisis de calidad de datos de la dimensión SEDES.....	66
6.5	Implementar reportes y visualización.....	68
6.5.1	Herramienta de visualización.....	68
6.5.2	Creación de reportes y visualización.....	69
6.5.3	Actualización de los datos.....	74
6.5.4	Impacto corporativo de la solución de inteligencia de negocios.....	75
7.	CONCLUSIONES.....	77
7.1	Conclusiones de la ejecución del proyecto de BI.....	77
7.2	Conclusiones del conocimiento adquirido con la ejecución del proyecto.....	79
8.	TRABAJOS FUTUROS.....	81
9.	REFERENCIAS.....	82
ANEXO 1.....		87

LISTA DE FIGURAS

<i>FIGURA 1.</i> TAREAS DE LA METODOLOGÍA KIMBALL (KIMBALL & ROSS, 2011).....	22
<i>FIGURA 2.</i> MODELO KIMBALL (KIMBALL & ROSS, 2011).	23
<i>FIGURA 3.</i> TOPOLOGÍA DE RED DEL CLIENTE, AUTORÍA PROPIA	26
<i>FIGURA 4.</i> PRECIOS DE PTRG (“REQUERIMIENTOS PRTG,” N.D.).....	35
<i>FIGURA 5.</i> PRECIOS DE SOLARWINDS, NMP(“SOLARWINDS - PRICE LIST,” 2018).....	36
<i>FIGURA 6.</i> CVE APLICATIVO CACTI(MITRE CORPORATION, N.D.).....	37
<i>FIGURA 7.</i> TAREAS DE LA METODOLOGÍA KIMBALL (KIMBALL & ROSS, 2011).....	39
<i>FIGURA 8.</i> EJEMPLO GENERAL DE STARNET, AUTORÍA PROPIA	40
<i>FIGURA 9.</i> METODOLOGÍA EN CASCADA (S. PRESSMAN, 2010)	41
<i>FIGURA 10.</i> STARNET DEL PROYECTO, AUTORÍA PROPIA	48
<i>FIGURA 11.</i> ARQUITECTURA DE BI DEL PROYECTO, AUTORÍA PROPIA.	50
<i>FIGURA 12.</i> MODELO DIMENSIONAL, AUTORÍA PROPIA.	54
<i>FIGURA 13.</i> CREACIÓN DE TABLAS EN ACCESS., AUTORÍA PROPIA.	55
<i>FIGURA 14.</i> CARGA EN MICROSOFT ACCESS DESDE LIBROS DE EXCEL, AUTORÍA PROPIA.	55
<i>FIGURA 15.</i> ETAPAS DE CONSTRUCCIÓN DEL SCRIPT, AUTORÍA PROPIA.	57
<i>FIGURA 16.</i> FORMULARIO PARA LANZAR SCRIPTS, AUTORÍA PROPIA.	58
<i>FIGURA 17.</i> CÓDIGO DE CONEXIÓN A LA BASE DE DATOS, AUTORÍA PROPIA.....	59
<i>FIGURA 18.</i> TABLA IP, AUTORÍA PROPIA.....	60
<i>FIGURA 19.</i> CÓDIGO PARA CAPTURAR EL IOS, AUTORÍA PROPIA.....	61
<i>FIGURA 20.</i> CÓDIGO CON SENTENCIAS SQL PARA POBLAR LAS DIMENSIONES.....	61
<i>FIGURA 21.</i> ANÁLISIS CALIDAD DE DATOS DIMENSIÓN PUERTOS, AUTORÍA PROPIA.....	64
<i>FIGURA 22.</i> ANÁLISIS CALIDAD DE DATOS DE FRECUENCIA CAMPO INT_CATEGORIA, AUTORÍA PROPIA.	65
<i>FIGURA 23.</i> ANÁLISIS CALIDAD DE DATOS DIMENSIÓN RT_PUERTOS, AUTORÍA PROPIA.....	66

<i>FIGURA 24. ANÁLISIS CALIDAD DE DATOS DIMENSIÓN SEDES, AUTORÍA PROPIA.</i>	<i>67</i>
<i>FIGURA 25. ANÁLISIS CALIDAD DE DATOS NULOS, ÚNICOS Y DUPLICADOS DE LATITUD, AUTORÍA PROPIA.</i>	<i>67</i>
<i>FIGURA 26. MODELO DE DATOS EN QLIK, AUTORÍA PROPIA.</i>	<i>69</i>
<i>FIGURA 27. REPORTE DE INVENTARIO, AUTORÍA PROPIA.....</i>	<i>71</i>
<i>FIGURA 28. REPORTE OPERACIÓN, AUTORÍA PROPIA.</i>	<i>72</i>
<i>FIGURA 29. REPORTE FACTURACIÓN. AUTORÍA PROPIA.....</i>	<i>73</i>
<i>FIGURA 30. REPORTE CAJEROS, AUTORÍA PROPIA.</i>	<i>74</i>
<i>FIGURA 31. ACTUALIZAR DATA DE LA VISUALIZACIÓN. AUTORÍA PROPIA.</i>	<i>74</i>
<i>FIGURA 32. ACTIVIDAD DE REVISIÓN GENERADA POR RESULTADOS DE ESTE PROYECTO, AUTORÍA PROPIA.....</i>	<i>75</i>
<i>FIGURA 33. INDICADORES DE PUERTOS</i>	<i>76</i>

GLOSARIO

Almacén de datos: base de datos centralizada e integrada con contenido de diferentes fuentes para soportar inteligencia de negocios.

Bodega de datos: Almacén de datos,

Business Intelligence: BI

Calidad de datos: Información que cumple los requerimientos y expectativas de quienes la usan.

Cisco: Empresa que provee equipos de Telecomunicaciones.

Datamart: Bodega de datos limitada a un área o proceso de la empresa.

Dato: Registro de un hecho.

ETL: Software o herramienta que extrae, convierte y carga datos al almacén.

Networking: Comunicación de dispositivos conectados en una red.

Nulls: Campos vacíos en una base de datos.

SSH: Protocolo de conexión seguro.

Stack: Varios equipos configurados como un *cluster* con gestión unificada.

Switches: Equipos de red a los que se conectan los dispositivos finales.

Tablero de control: Interfaz en la que se visualizan gráficas e indicadores resultantes de las herramientas de BI.

TI: Área de tecnología de información dentro de una compañía.

RESUMEN

La gran cantidad de dispositivos de *networking* en producción bajo el modelo comercial de IaaS hace un reto mantener la oportunidad y calidad de la información del detalle y descripción de cada puerto de red operativo. Se ejecutó un proyecto de inteligencia de negocios (BI) que permite obtener la información directamente de la fuente (*switches* Cisco) automatizando la captura de datos, integrado con hojas de cálculo de inventario, se revisó la calidad de estos y posteriormente se implementó una herramienta de visualización que apoya el proceso de facturación e inventario.

PALABRAS CLAVE: Inteligencia de negocios; modelamiento BI; bodega de datos; tableros de control.

ABSTRACT

The large number of networking devices in production under the IaaS business model make it a challenge to maintain the quality of the detail information and description of each operational network port. It seeks to execute a business intelligence (BI) project that allows obtaining the information directly from the source (Cisco Switches) by automating the capture of data, integrate it with other inventory spreadsheets, review the quality of them and subsequently implement an exploration data tool that supports the billing and inventory process.

KEY WORDS: Business intelligence, dimensional modeling, datawarehouse, dashboards,

1. INTRODUCCIÓN

Este proyecto buscó mediante técnicas de inteligencias de negocios dar solución a las necesidades de información de un servicio de *networking* prestado por una empresa grande a otra grande en Colombia, conocido como servicios *business-to-business* (B2B).

En esta compañía, la inteligencia de negocios no ha sido aplicada a los servicios de red prestados a los clientes, y tanto estos como la compañía misma ven la necesidad de avanzar en ese aspecto. Hoy día los clientes están más interesados en un servicio administrado que ofrezca analítica predictiva, descriptiva, correlación de eventos y mucho más que solamente un enlace de comunicaciones.

A lo largo de este documento se expone cómo construir una solución de inteligencia de negocios con énfasis en los procesos de operación, facturación e inventario. Siguiendo metodologías vigentes, se diseñó e implementó un modelo dimensional que permite mediante una herramienta de exploración de datos el análisis por parte de las personas interesadas en tomar decisiones, revisar indicadores o elaborar reportes internos o al cliente.

Los datos que componen una solución de BI deben tener la mejor calidad posible, por lo que para este proyecto se desarrolló código para obtener desde los mismos equipos toda la información operativa. Se encuentra el paso a paso para poblar el almacén de datos y los reportes generados que permiten el análisis de información. Finalmente se realizó un perfilamiento de datos con las

dimensiones ya pobladas que permitan desde el mismo trabajo generar la mejora continua de la solución.

1. PLANTEAMIENTO DEL PROBLEMA

1.1 Problema

La organización actualmente debe administrar más de 1500 dispositivos de red y más de 400 enlaces de comunicaciones para atender 379 sedes del cliente distribuidas por todo el país. El esquema contratado por el cliente se conoce como modelo administrado por puertos, por tanto, la facturación final de la operación del proyecto se reduce a un cobro por cada puerto de red usado más el cobro por el ancho de banda del enlace. A su vez se tienen tipificados diversos tipos de puertos que generan o no cobro y dependiendo del tipo es diferente su valor.

El problema para la organización consiste básicamente en la dificultad para integrar y analizar la información de los puertos en uso por los clientes y poder hacer los cobros correspondientes. La ausencia de esto, está generando inconformidad con el cliente por las diferencias entre la realidad (los puertos realmente en uso) y lo facturado.

El indicador de factor de calidad y cumplimiento general (FCCG) es conciliado con el cliente mes a mes con revisión quincenal y permite realizar renovación de prestación de servicios de forma automática sin necesidad de salir a licitación, siempre y cuando esté en un valor mínimo establecido por el cliente durante todo el contrato. En las actas generadas durante estas revisiones el cliente ha registrado quejas por la demora en reportes e información que ellos consideran que se deberían tener de inmediato como en el acta 11 de seguimiento de Jun 2018, y repetidas inconformidades debido a inconsistencia de puertos facturados vs usados realmente como en las actas 8 y 18 de 2018 donde inclusive se generan demoras del proceso de pago hasta tener información confiable. Esto ha derivado en trabajos urgentes para levantar información, pero con

el tiempo se desactualiza y se vuelve un problema recurrente. Las actas de seguimiento no se pueden mostrar por confidencialidad.

Toda la gestión se realiza de manera manual en documentos de office y al no tener una base de datos de configuración (*CMBD*) bien implementada como lo propone (Yamada, Yada, & Nomura, 2011), llevar a cabo tareas de inventario y reportes particulares para el cliente interno o cliente final toma un tiempo importante. Los informes pueden requerir tres (3) días de un ingeniero dedicado y algunas solicitudes particulares generan “planes de choque” en las que se distribuyen tareas y sitios para tener resultados en dos a cuatro semanas. Se deben abrir múltiples hojas de cálculo e incluso ingresar a muchos dispositivos para buscar la información en la que se requiere información actualizada.

Los acuerdos de niveles de servicio bajo el modelo contratado con el cliente contemplan una disponibilidad cercana al 100%, lo que se dificulta debido a que la información actual de inventario está en hojas de Excel para tomar decisiones sobre cantidades y distribuciones de equipos en las bodegas, mantener equipos por mucho tiempo en inventario sin uso es un indicador negativo para la organización y la optimización de inventario y posible redistribución de los equipos adquiridos incluso en otros proyectos es lo deseado para hacer uno de los bienes de capital en lo que ya se ha invertido.

2. JUSTIFICACIÓN

Mantener un alto nivel de exactitud, oportunidad y calidad de los datos de puertos operativos e inventario permite tener la tranquilidad del cumplimiento de acuerdos de servicio con el cliente y mejorar la gestión de la operación. Al mantener lo más actualizados posibles los datos de puertos en uso y distribución de inventario, se puede lograr disminuir el tiempo invertido de los ingenieros del proyecto a estas tareas, además de tomar mejores decisiones de distribución de equipos de respaldo en las bodegas del país mejorando los indicadores económicos del área.

Una visualización apropiada de estos aspectos apoya la elaboración de informes, reportes ágiles y toma decisiones de una manera oportuna sin dedicar demasiadas horas hombre para la elaboración de estos ítems. Adicionalmente aporta al buen funcionamiento del flujo de trabajo y esto es percibido directamente por el cliente, lo cual contribuye con el aumento de la confianza de éste, generando así ingresos para la compañía por instalación de sitios nuevos o renovación de contratos.

Como podrá observarse más adelante en el estado del arte, varias herramientas para gestión de red analizadas, además de costosas, no cubren las necesidades expuestas debido a que se enfocan en resolver otros temas como gestión de fallas, configuración, entre otras, consideradas mejores prácticas mencionadas por Cisco (Cisco Systems, 2007). Algunas herramientas gratuitas tienen el inconveniente de que no cuentan con el soporte y tampoco desarrollo continuo para resolver vulnerabilidades.

Dado que no existe comercialmente una herramienta de gestión de red que cubra todas las necesidades expuestas o ésta implicaría un costo demasiado alto por el licenciamiento tan grande requerido, se justifica construir una herramienta que obtenga e integre los datos de diferentes sedes y permitir así un buen análisis mediante un prototipo de visualización usada en inteligencia de negocios que contenga los datos importante para la facturación y para la toma de decisiones de inventario y de respaldos acordes a los SLA del servicio. No se justifica realizar instalaciones de diversas aplicaciones con los recursos y procesos administrativos que esto implica para obtener información parcial y tener en todo caso que integrar las bases de datos debido a las limitaciones de aplicativos que usan SNMP como se menciona más adelante.

Luego de la integración de la información, todo el proceso se vio enriquecido por el dinamismo, interacción y poderosa visualización de las herramientas de Inteligencia de Negocios potenciando la toma de decisiones al respecto.

3. OBJETIVOS

3.1 Objetivo General.

Desarrollar un prototipo de Inteligencia de Negocios para el análisis de los procesos de facturación e inventarios en la prestación de servicios de *Networking* mediante técnicas de Modelamiento Multidimensional y Visualización de Información.

3.2 Objetivos Específicos

- Definir los requerimientos de información del negocio.
- Construir el modelo multidimensional.
- Desarrollar una aplicación que permita la extracción de la información de los dispositivos Cisco en producción de forma automática.
- Implementar reportes y visualización con la base de datos levantada.

4. MARCO REFERENCIAL

4.1 Marco contextual

El crecimiento tecnológico ha llevado a que hoy en día se tengan muchísimos datos y se esté creciendo a cada momento en la cantidad de estos por el auge generado con el internet de las cosas (Chen, Mao, & Liu, 2014). El área de *Networking* no ha sido la excepción y bajo el modelo *IaaS (Infrastructure as a service)*, es crucial la correcta calidad de los datos críticos (Merino, Caballero, Rivas, Serrano, & Piattini, 2016).

Este proyecto se orientó a atender necesidades de gestión de red de servicios de *networking* prestados hoy día por terceros como servicio a grandes empresas. Pudiese ser aplicado a cualquier empresa con solución Cisco u otro fabricante con pequeños ajustes, muy común en las grandes empresas en Medellín, para este caso consiste en un servicio prestado de empresa a empresa conocido como *business to business (B2B)* entre un importante proveedor de servicios de telecomunicaciones y una entidad financiera con más de 600 sedes en el país. Para la prestación de estos servicios el alcance usualmente se limita a prestar toda la infraestructura necesaria para la conectividad y cobrar un canon por tipo de puerto utilizado, o por tipos de dispositivos utilizados según los acuerdos comerciales, se negocian diversos niveles de acuerdos de servicio (SLA) de la disponibilidad, atención de incidentes y capacidad y rendimiento de la solución prestada.

Una compañía prestadora de servicios de comunicaciones busca tener un *framework* de gobierno y gestión para operar con calidad, actualmente se está implementando lentamente procesos basados en *IT Infrastructure Library (ITIL)*, y no ha tenido el impacto y la velocidad que se quisiera principalmente por la resistencia al cambio (Esteves & Alves, 2013), actualmente se cuenta con información en hojas de cálculo con campos manuales, almacenadas en carpetas compartidas con privilegios asignados según el rol con las dificultades administrativas que esto conlleva para visualizar y obtener información.

Toda la solución de red para el cliente es de Cisco Systems por solicitud de éste (exceptuando equipos de enlaces de comunicaciones), lo que facilita la configuración y administración. Son factores claves de éxito para mejorar el servicio y procesos de TI la buena administración de TI basados en estándares (Diirr & Santos, 2014), se busca impactar positivamente el servicio creando y aplicando estándares y se apoya la idea .

4.2 Marco conceptual

4.2.1 Inteligencia de negocios (BI)

Business Intelligence (BI) se puede definir como el conjunto de sistemas, aplicaciones y procesos de gobierno de toda la empresa que permite análisis sofisticados, al permitir que fluyan datos, contenido y análisis a quienes, lo necesitan cuando lo necesitan (Davenport & Harris, 2007)

BI se ha mantenido como prioridad de inversión para los *CEO* (Kalakota, n.d.), durante varios años, lo cual significa dos cosas, la primera que la buena información es la clave para tomar acertadas decisiones, y que se han desarrollado diversas estrategia y metodologías para llevar a cabo proyectos de BI (Holsapple, Lee-Post, & Pakath, 2014). Dentro de estas, los modelos más famosos son los de Inmon y Kimball. El mismo Inmon ha tratado de defender su modelo comparando con Kimball (Inmon, 2011), cada cual tiene sus ventajas, Rivadera recomienda utilizar el modelo propuesto por Kimball por la versatilidad del enfoque en la que se empieza por los *Datamarts* (Rivadera, 2010) y permite agilizar la puesta en marcha de las soluciones. La figura 1 muestra la metodología de Ralph Kimball para la ejecución de un proyecto de BI y la figura 2 muestra el modelo (Kimball & Ross, 2011).

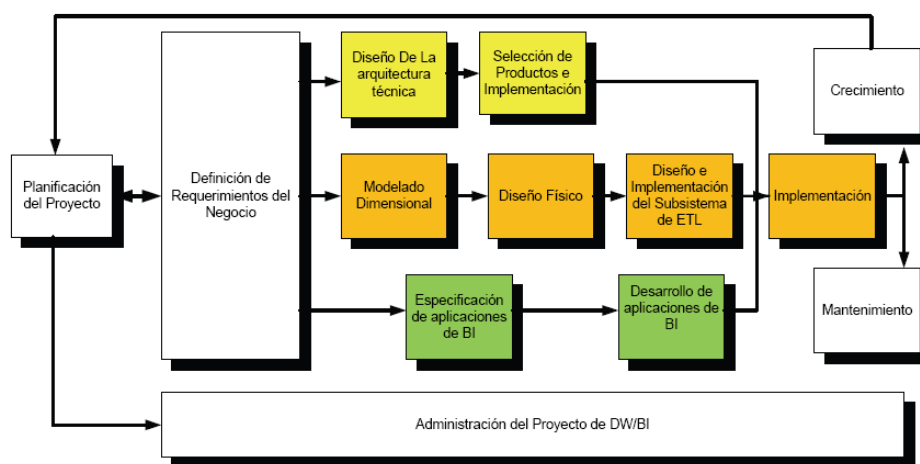


Figura 1. Tareas de la metodología Kimball (Kimball & Ross, 2011).

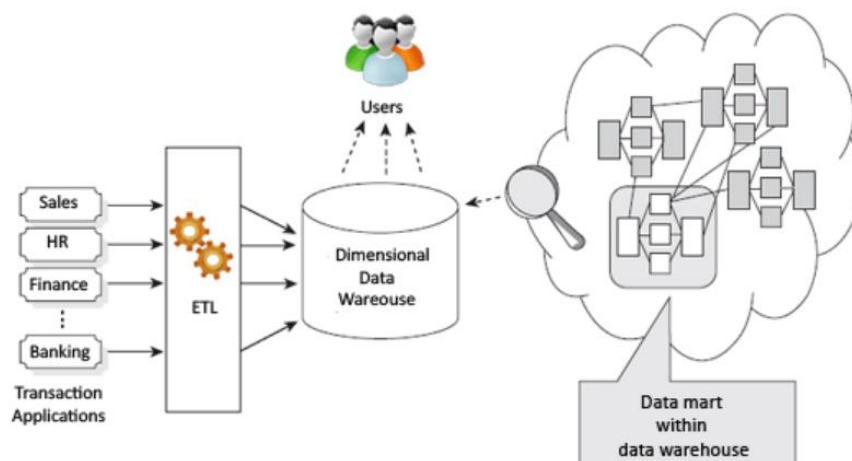


Figura 2. Modelo Kimball (Kimball & Ross, 2011).

Un *Datamart* es una bodega de datos limitada a una área de interés utilizando un modelo de datos dimensional (Dama, 2008).

Un *Datawarehouse* es una base de datos centralizada e integrada con contenido de diferentes fuentes para soportar BI. R. Kimball la define como una versión de los datos de transacción específicamente estructurados para consultas y análisis, compuesto por data marts (Dama, 2008).

Los datamarts se poblan mediante procesos de ETL (*Extract-Transform-Load*). Se define como el conjunto de software o herramientas que extrae datos de la fuente, lo convierte, transforma y carga a otra base de datos diferente (Dama, 2008).

Los proyectos de BI buscan convertir los datos en información valiosa. El concepto de dato se puede definir como el registro de un hecho, no es información, pero es su materia prima, “El

valor de almacenar grandes cantidades de datos depende de nuestra habilidad para extraer reportes útiles” (Mezzanica, Boselli, Cesarini, & Mercurio, 2015).

Para generar valor se requiere entonces una buena calidad en los datos, información que cumpla los requerimientos y expectativas de quienes la usan (Dama, 2008), Una alta calidad de datos es considerado como un factor crítico de éxito desde el punto de vista de la inteligencia de negocios (Rud et al., 2009), por esto los datos deben ser tratados como activos (Finch et al., 2014) y no solo se deben guardar datos sino adicionalmente datos de los datos, conocido como metadatos (Stanton, 2012).

Los datos cuentan con características (Juran & Godfrey, 1998), dimensiones y categorías (Karkouch, Mousannif, Al Moatassime, & Noel, 2016). A continuación, se presentan seis dimensiones (6) de calidad críticas (DAMA, 2013):

Compleitud: La porción de registros contra el total posible, se miden registros vacíos o nulos.

Los datos maestros o críticos deben ser chequeados por completitud con prioridad.

Unicidad: Registros que no deben repetirse, esto es, que no deben estar más de una vez.

Oportunidad: Registro en el momento que se requiere, y su validez en el tiempo.

Conformidad: Los registros deben estar de acuerdo con el formato para su correcta gestión, se aconseja particular cuidado con las fechas y formatos definidos desde el diseño.

Exactitud: Se refiere a que el dato esté libre de errores y sea una correcta descripción de la realidad.

Consistencia: Un registro es consistente cuando no existe otro que le contradiga o haya lugar a dudar de cuál es el correcto. Usualmente se valida la consistencia comparando datos incluso de diferente fuente de información.

Si bien para mantener la calidad de datos hoy día se propone crear, procesos, buenas prácticas y procedimientos (Mosley, Brackett, Earley, & Henderson, 2009), la actividad de *data cleaning* o *data cleansing*, se refiere a detectar y eliminar errores, inconsistencias (Lee, Hsu, & Kothari, 2004). Para este proyecto se limitó a la detección de campos que no se encuentren en los estándares definidos por la operación más no a establecer procesos de gobierno sobre los datos.

4.2.2 Soluciones de red Cisco.

Cisco Systems es un fabricante de tecnología líder mundial en fabricación de equipos de *Networking* y es reconocido además por ser un fabricante que apunta a la innovación y generación de conocimiento a tal punto que es reconocido por generar las mejores prácticas de implementación de soluciones de tecnología de conectividad incluso con desarrollo de protocolos propietarios (Sullivan, 2015).

Comercialmente es conocido el concepto de *Networking as a service (NaaS)* e *Infrastructure as a Service (IaaS)*. Es la prestación de servicios en la que los clientes no quieren tener infraestructura propia para el caso de estudio de este proyecto consiste en un modelo *IaaS* en el

que se presta un servicio integral de conectividad: *Lan*, *WLAN*, *Wan*, *Aceleración Wan*, *VoIP* y *Colaboración*, facturando por puerto de red que esté utilizando el cliente y por enlaces *WAN* en producción. La figura 3 muestra la solución típica de una sede y representa el diseño en sitios remotos del cliente, se remarcan los dispositivos en modalidad *IaaS* responsabilidad del prestador de servicio en verde. La cantidad de *Switches* puede aumentar dependiendo del tamaño de la sede.

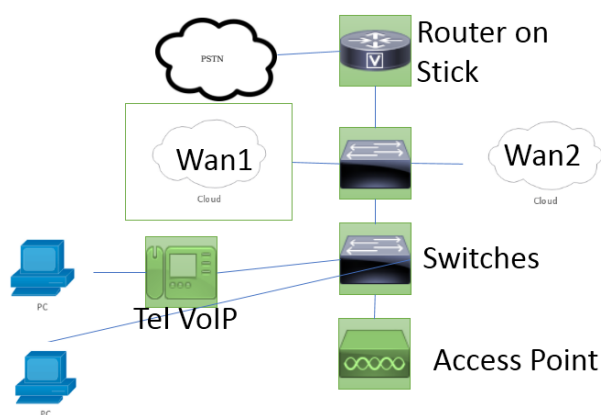


Figura 3. Topología de red del cliente, autoría propia

4.2.3 Secure Shell (SSH)

SSH es el protocolo de comunicación por excelencia para los dispositivos Cisco por tratarse de un protocolo seguro (Liu, 2009), funciona bajo el modelo cliente servidor y los dispositivos Cisco pueden ser configurados de manera segura para usarlo.

Existen diversos programas para conectarse usando SSH, uno de ellos SecureCRT de VanDyke, una de sus mayores cualidades es la funcionalidad de ejecutar Scripts en lenguaje Python o en Visual Basic Script (VBS) lo cual es de vital importancia para este proyecto.

4.2.4 Lenguaje Visual Basic Script (VBS)

Es un lenguaje de programación utilizado por soluciones con Microsoft Windows y es considerado una variación de Visual Basic, muy utilizado para scripts con Internet Explorer y como lenguaje para automatización (Lomax, Childs, 2003). Este se considera un lenguaje interpretado y no compilado por lo que se puede utilizar un procesador de texto para las aplicaciones. Este proyecto contempla el uso del lenguaje VBS para crear una aplicación que permita extraer la información de los dispositivos Cisco en producción de forma automática y rápida para su posterior análisis y visualización.

4.2.5 Visualización de datos

Escoger adecuadamente la presentación de los datos es fundamental para transmitir un mensaje claro, los *dashboards* se encuentran en auge y son los favoritos de los clientes, siendo un deber ofrecer una recomendación acertada. Stephen Few (Few, 2014) explica las diferentes opciones de visualización y sugiere escoger según el propósito de información. Dentro de diversas opciones para una alta interacción flexible y fluida del usuario con filtros particulares, una urgencia de respuesta baja y uso regular, se recomienda una herramienta que permita la exploración de los datos.

4.2.6 Protocolo SNMP

SNMP es un protocolo que se encuentra en la categoría de estándar. Se originó para apoyar la administración de redes bajo el modelo cliente-servidor, en el que el servidor se conoce como administrador SNMP y a los clientes como Agentes SNMP. Para lograr la gestión del

dispositivo se utilizan las MIBs que son identificadores que tienen cada Agente (dispositivo) para poder realizar consultas o modificaciones. Las MIB a su vez está compuesta por objetos que se identifican con un identificador estándar (OID), el cual es una secuencia de números.

El protocolo SNMP cuenta solo con 6 instrucciones: GetRequest, GetNextRequest, GetResponse, SetRequest, Trap y GetBulkRequest. El grupo GET sirve para que el administrador SNMP obtenga información particular solicitándolo al Agente; los mensajes TRAP envían información sin ser solicitada cuando un evento ocurre y los mensajes SET se usan para realizar modificaciones en la configuración de los dispositivos. Finalmente, la instrucción BulkRequest es simplemente una petición de información grande y es más eficiente realizar una grande que muchos GetRequest pequeños (Terplan, 1992).

4.3 Marco legal

Este proyecto se encuentra en términos generales bajo la normativa de Superintendencia Financiera de Colombia expedida por medio de la circular 0052 de 2007 (Financiera, 2007) que aplica para sedes de entidades financieras. En ningún momento se realizará algún procedimiento que comprometa algún ítem de la circular. Resaltan el manejo confidencial de la información, el uso de la corporación MITRE para revisar CVE tal como se realiza en este documento con una de las posibles herramientas, también la necesidad de identifiquen apropiadamente los usuarios que hacen uso de los sistemas de información y poder auditar los cambios o consultas que realizan.

Adicionalmente este proyecto debe cumplir con las siguientes normativas:

No divulgación de datos a terceros por el capítulo de privacidad y confidencialidad celebrado en el contrato entre el proveedor y el cliente. Este ítem aplica para todos los datos del proyecto que hagan parte del inventario o configuración de los dispositivos.

Políticas de uso aceptable de los usuarios utilizados para acceder a los dispositivos y adicionalmente políticas de contraseña para el usuario que ejecuta el script para obtener la data. Mínimo 12 caracteres con mínimo una letra mayúscula, una letra minúscula, un carácter especial y un número, no puede ser ninguna de las cuatro (4) anteriores y no se debe compartir con otras personas. Tampoco debe incluir frases que incluyan el nombre de usuario, de las personas o de las organizaciones.

No divulgación de costos y precios ofrecidos por el proveedor al cliente según se define en la Política de confidencialidad de ofertas comerciales con propuesta económica definida. El proyecto contempla realizar una visualización para apoyar la facturación en la que se incluirá los costos para que se logre visualizar cuanto aporta la selección filtrada por el usuario.

4.4 Estado del arte

El Tema del BI ha tenido una gran aceptación y como se mencionó anteriormente se volvió una prioridad de muchos ejecutivos desarrollar proyectos de BI y analítica en sus organizaciones, debido a que presentan un importante impacto en la buena toma de decisiones.

Si bien se habla de BI desde décadas atrás, toma más fuerza hoy que el precio del hardware y del almacenamiento es más bajo al contar con mejores procesadores y cantidades grandes de almacenamiento (Chaudhuri, Dayal, & Narasayya, 2011), lo que ha permitido que el desarrollo de aplicaciones y soluciones se encuentre en auge.

Es muy usado el modelo de Kimball en el que se realizan modelos de base de datos con una tabla de hechos en el centro y dimensiones alrededor. Hoy día se utiliza un modelo conocido como *star flake*, que es la combinación de un modelo de base de datos en estrella y copo de nieve (Sulaiman & Yahaya, 2013). Además, una vez definida su granularidad, permite analizar más de una tabla de hechos para la base de datos.

Jones y Song presentan un interesante modelo sistemático para el diseño de dimensiones por patrones, en búsqueda de encontrar un método sistemático que apoye el diseño de almacenes de datos. Presentan los DDP (Patrones de diseño de dimensiones) y sus integraciones con la metodología Kimball, aumentando la eficiencia a la hora de modelar al ser usado como

herramienta para tal propósito. Se busca incluir el cuándo, donde, quién, qué, porqué y la acción, acompañado de preguntas y respuestas para evaluar la pertinencia de cada clase. El modelamiento es conceptual y como un lineamiento para la creación de otros modelos dimensionales siguiendo como guía o *framework* para hacerla mejor (Jones & Song, 2008).

En el entorno de *newtorking* en general, la analítica y el enfoque de herramientas de BI se presentan orientadas a analizar el comportamiento de los usuarios, sus preferencias, Cisco propone BI y analítica para analizar soluciones de Voz, Contact Center y Wlans tanto para tomar decisiones corporativas como ofrecer servicios extras incluyendo anuncios personalizados.

Los modelos predictivos analíticos y el *big data* empiezan a tomar fuerza y desarrollarse en conjunto por proyectos de BI, las visualizaciones descriptivas de datos continúan teniendo un fuerte auge por el terreno ya abonado durante décadas pero el BI se está transformando con la entrada de la Big Data y nuevas oportunidades de analítica, se ha incluido el BI creando sinergia con las metodologías ágiles para acelerar los proyectos de BI y abordar las dificultades comunes especialmente la colaboración en interacción con los altos rangos de la compañía e incluso es útil las metodologías ágiles a nivel operacional. (Larson & Chang, 2016)

Li, Shue y Lee aplican BI a *un Internet Service Provider* (ISP), buscando obtener la información más relevante que permita ver la ocupación de la red realizada por los usuarios finales, analizando el comportamiento de los usuarios en términos de tráfico utilizado, buscando

encontrar patrones de usuarios para optimizar recursos del ISP. Se utiliza la técnica *Clustering* de minería de datos, que permite agrupar comportamientos mostrando un ejemplo práctico de la incursión de la analítica a este tipo de proyectos (Li, Shue, & Lee, 2008). La captura de muchos de los datos es por herramientas de *networking* de monitoreo de dispositivo y tráfico apoyado en SNMP, aun así, se evidencia en el desarrollo que el modelamiento dimensional y las etapas de un proyecto de BI continúan siendo en esencia las mismas.

En los proyectos de BI, uno de los temas importantes es la visualización de la información, con el objetivo de encontrar la mejor visualización que satisfaga las necesidades de información e interacción de los clientes finales. (Tokola, Gröger, Järvenpää, & Niemi, 2016) presentan algunos principios importantes para el diseño de visualizaciones y particularmente *dashboard*, como que cada usuario pudiera requerir ver información diferente y ver los KPI que para ellos son relevantes, siendo fundamental para el diseño de cualquier visualización la realimentación permanente con el usuario desde la fase de pruebas.

En (Sulaiman & Yahaya, 2013) se presentan diversos trabajos de diseño de *datawarehouse* y realizan el modelo dimensional en estrella. Algunos van más allá y se conocen como *Decision Support Systems* (DSS) los que incluyen *On-Line Analytical Processing* (OLAP) y visualizaciones, se resalta la importancia de realizar verificaciones en la integridad, consistencia y estandarización de los datos o puede haber muchas dificultades y la necesidad de una fuerte etapa de *staging*. (Rai, Dubey, Chaturvedi, & Malhotra, 2008).

Los autores (Khalilian, Ibrahim, & Nilashi, 2017), en su artículo *La construcción de un Datawarehouse (DW)* para apoyar la gestión de TI comparten parte de los problemas los reportes que no son precisos y confiables. Generar buenos reportes requiere un esfuerzo manual importante. Ellos proponen un DW híbrido que permita generar información con calidad al cual se llegue con realimentaciones constantes de los usuarios incluyendo los procesos operaciones e ITIL, la mayor dificultada para este proyecto fue el cambio de procesos, aún más de lo esperado por el tiempo invertido en documentación, en cambiar la cultura hacia los procesos entre otros.

En parte, este trabajo se enfocó en traer información directamente de los dispositivos lo cual la hace confiable y busca también generar las alertas cuando encuentre información fuera del estándar establecido y así mantener la calidad de la data. Tener aplicados campos y medidas estándar permite realizar un benchmarking (Completo, Cruz, Coheur, & Delgado, 2012) y los proyectos de BI incluyen este tipo de procedimientos para poder ser funcionales y es una buena práctica que se busca tener para mitigar alguna de las dificultades de otros proyectos de BI.

En el área de *networking* es una buena práctica conocida el uso de herramientas que saquen provecho del protocolo SNMP, actualmente existen un número importante de herramientas que puede traer información importante, las mejores de estas con pago por su licenciamiento. Lamentablemente estas herramientas presentan una serie de limitaciones que se resumen en presentación visual, costo, y seguridad. La seguridad es una de las más importantes debido a que al tratarse de una entidad financiera todos los equipos deben estar protegidos y más un equipo que use SNMP o Net Flow dado que con acceso indebido se puede acceder a mucha información

y cambiar configuración que puede derivar en sitios fuera de servicio, este alto nivel de seguridad deseado limita a las aplicaciones *open source* más conocidos.

En el mercado existen diversas aplicaciones que utilizan SNMP para apoyar la gestión de red, algunas de estas de pago que almacenan la información en bases de datos para permitir realizar análisis históricos de la data, usualmente también realizan monitoreo de los dispositivos y generan acciones ante comportamientos no deseados en la red.

La mayoría de estas aplicaciones están orientadas a la operación de la red, seguimiento de las variables de CPU, Memoria, Temperatura y tráfico cursado por las interfaces.

Dentro de las herramientas más populares se encuentran PRTG, CACTI, y Solarwinds, utilizan SNMP en conjunto con otros protocolos para el monitoreo y gestión teniendo en cuenta las buenas prácticas de redes.

Para este proyecto que se buscó tener información de inventario por puertos donde se superan la cantidad de 20.000, estos aplicativos no satisfacen las necesidades dado que no están diseñados para esto, ellos buscan apoyar la gestión desde otros puntos de la operación, como configuraciones automáticas, respaldo de configuración, monitoreo de red y servidores, entre otros que no hacen parte del alcance de este proyecto. Estas herramientas no son buenas levantando información de inventario, de módulos y HWIC instalados en los dispositivos y sería

una mala práctica instalar o utilizar un servidor robusto y pagar una licencia importante para obtener información incompleta.

Con PRTG por ejemplo sería necesario contar con una máquina robusta y la licencia de más alto nivel de 14.500 USD (“PRTG - Price List,” n.d.) para satisfacer solamente el inventario por puerto, dado que para PRTG se necesitaría 1 sensor por cada puerto que se quisiera monitorear. Aun si se tuviese ya operando, haría falta tener que migrar la información en dicha base de datos a una nueva para realizar BI sobre esta información. La figura 4 presenta los precios de la herramienta PRTG.

Sensors per Core Server	License	Recommended Core Server Hardware	Disk Space (1 year data retention)	User Accounts	Remote Probes	Virtualization	PRTG Cluster
up to 1,000 sensors (~ 100 devices)	PRTG 1000	2 CPU Cores, 3 GB RAM	250 GB	< 30	< 30	✓	✓
1,000 – 2,500 sensors (~ 250 devices)	PRTG 2500	3 CPU Cores, 5 GB RAM	500 GB	< 30	< 30	✓	✓
2,500 – 5,000 sensors (~ 500 devices)	PRTG 5000	5 CPU Cores, 8 GB RAM	1 TB	< 20	< 30	✓	!
5,000 – 10,000 sensors (~ 1,000 devices)	PRTG XL1	8 CPU Cores, 16 GB RAM	2 TB	< 10	< 30	!	!
More than 10,000 sensors	We recommend that you set up additional PRTG core server installations and contact your presales team .						

Figura 4. Precios de PTRG (“Requerimientos PRTG,” n.d.)

Solarwinds, actualmente está en proceso de instalación para la red del ISP involucrado en este proyecto. Es una herramienta más completa que PRTG dado que permite incluir no solamente equipos de red sino servidores, clústeres, WLAN, entre otros con la adición de diferentes módulos. Aun así, esta herramienta no ofrece visualizaciones personalizables y no cuenta con un buen desarrollo para obtener inventario de dispositivos, se limita al chasis y no trae información

completa para módulos, submódulos y equipos en *stack*. La figura 5 presenta una ilustración con el precio de solo el módulo de *network performance monitor (NPM)*.

Network Performance Monitor	
	Preço de tabela
SolarWinds Network Performance Monitor SL100 (up to 100 elements) - License with 1st-year Maintenance	\$2,955
SolarWinds Network Performance Monitor SL250 (up to 250 elements) - License with 1st-year Maintenance	\$6,720
SolarWinds Network Performance Monitor SL500 (up to 500 elements) - License with 1st-year Maintenance	\$10,445
SolarWinds Network Performance Monitor SL2000 (up to 2000 elements) - License with 1st-year Maintenance	\$19,345
SolarWinds Network Performance Monitor SLX (unlimited elements-Standard Polling Throughput) - License with 1st-year Maintenance	\$32,525

Figura 5. Precios de Solarwinds, NMP(“SolarWinds - Price List,” 2018)

Existen herramientas sin costo como la mencionada CACTI, pero aún haría falta el hardware para correrlo, recursos para el soporte y mantenimiento del aplicativo y el proceso administrativo para ejecutarlo y pasar las pruebas de seguridad que usualmente no aprueban este tipo de herramientas, por este ítem de seguridad y soporte en el tiempo las herramientas free no son bien vistas en organizaciones financieras. En la figura 6 se observan 62 Vulnerabilidades para el aplicativo CACTI, al no tener un responsable directo mitigarlas es un riesgo que se evita asumir y se prefieren soluciones licencias y soportadas con respuesta rápida.

Cacti : Security Vulnerabilities

CVSS Scores Greater Than: 0 1 2 3 4 5 6 7 8 9

Sort Results By : [CVE Number Descending](#) [CVE Number Ascending](#) [CVSS Score Descending](#) [Number Of Exploits Descending](#)

Total number of vulnerabilities: **62** Page : 1 (This Page) 2

[Copy Results](#) [Download Results](#)

#	CVE ID	CWE ID	# of Exploits	Vulnerability Type(s)	Publish Date	Update Date	Score	Gained Access Level	Access	Complexity	Authentication	Conf.	Integ.	Avail.
1	CVE-2018-10061	79		XSS	2018-04-12	2018-04-26	3.5	None	Remote	Medium	Single system	None	Partial	None
Cacti before 1.1.37 has XSS because it makes certain htmlspecialchars calls without the ENT_QUOTES flag (these calls occur when the htmlspecialchars function in lib/html.php is not used).														
2	CVE-2018-10060	79		XSS	2018-04-12	2018-04-26	3.5	None	Remote	Medium	Single system	None	Partial	None
Cacti before 1.1.37 has XSS because it does not properly reject unintended characters, related to use of the sanitize_uri function in lib/functions.php.														
3	CVE-2018-10059	79		XSS	2018-04-12	2018-04-26	3.5	None	Remote	Medium	Single system	None	Partial	None
Cacti before 1.1.37 has XSS because the get_current_page function in lib/functions.php relies on \$_SERVER['PHP_SELF'] instead of \$_SERVER['SCRIPT_NAME'] to determine a page name.														
4	CVE-2017-1000032	79		XSS	2017-07-17	2017-07-19	4.3	None	Remote	Medium	Not required	None	Partial	None
Cross-Site scripting (XSS) vulnerabilities in Cacti 0.8.8b allow remote attackers to inject arbitrary web script or HTML via the parent_id parameter to tree.php and drp_action parameter to data_sources.php.														
5	CVE-2017-1000031	89		Exec Code Sql	2017-07-17	2017-07-19	6.5	None	Remote	Low	Single system	Partial	Partial	Partial
SQL injection vulnerability in graph_templates_inputs.php in Cacti 0.8.8b allows remote attackers to execute arbitrary SQL commands via the graph_template_input_id and graph_template_id parameters.														
6	CVE-2017-16785	79		XSS	2017-11-10	2017-11-27	4.3	None	Remote	Medium	Not required	None	Partial	None
Cacti 1.1.27 has reflected XSS via the PATH_INFO to host.php.														
7	CVE-2017-16661	200		+Info	2017-11-08	2017-11-28	4.0	None	Remote	Low	Single system	Partial	None	None
Cacti 1.1.27 allows remote authenticated administrators to read arbitrary files by placing the Log Path into a private directory, and then making a clog.php?filename= request, as demonstrated by filename=password (with a Log Path under /etc) to read /etc/passwd.														
8	CVE-2017-16660	264		Exec Code	2017-11-08	2017-11-28	9.0	None	Remote	Low	Single system	Complete	Complete	Complete
Cacti 1.1.27 allows remote authenticated administrators to conduct Remote Code Execution attacks by placing the Log Path under the web root, and then making a remote_agent.php request containing PHP code in a Client-ip header.														

Figura 6. CVE Aplicativo CACTI(Mitre Corporation, n.d.)

Actualmente, la compañía se encuentra en fase de negociación para adquirir Solarwinds por medio del proveedor EDEA como herramienta para el monitoreo de las soluciones de TI. Dicha herramienta se encuentra como *Challenger* en el cuadrante de Gartner (Koutsoukos, 2018), esta y otras herramientas ayudan para gestionar el servicio y obtener los indicadores mensuales pero queda faltando para el tema de inventario de puertos y dispositivos en su totalidad.

Debido a otras limitantes adicionales al costo ya mencionado como visualización, limitación de inventario con *switches* en *stack*, inventario de módulos SFP, Cue y otras HWIC, se encuentra más sencillo la elaboración de un script a necesidad que traiga la información de la fuente. Pues bien, aún con varias herramientas que combinen SNMP y SSH se tendría que trabajar la base de datos que generen dichos aplicativos en pro de organizarlos, filtrarlos y convertirlos en información valiosa recurrentemente.

También puede observarse en los diversos trabajos que la metodología propuesta por Kimball continúa siendo vigente, desarrollando el modelo dimensional que reúne dimensiones como usuarios, tiempo, ubicación y estos a su vez se relacionan con una tabla de hechos.

Tal como concluyen Rikhardssona y Yigitbasioglu, si bien existe gran cantidad de *papers* entorno al BI, no se ofrecen soluciones específicas, siendo la mayoría conceptuales (Rikhardsson & Yigitbasioglu, 2018). Este mismo pensamiento aplica para el desarrollo de este proyecto apoyando la gestión de TI, con la mayoría de los estudios enfocados al área comercial o comportamiento de los usuarios para buscar nuevas oportunidades de negocio no se encuentran muchos trabajos similares del que podamos extraer información valiosa más allá de las metodologías utilizadas. Aun así, es suficiente para sentar las bases conceptuales y metodológicas para llevar a cabo este u otro proyecto de BI.

5. METODOLOGÍA

Para este proyecto se utilizaron partes de diversas metodologías de TI, descritas en el marco referencial de este proyecto. Tiene un enfoque cuantitativo, cada paso de este proyecto busca obtener datos de forma confiable que respondan las necesidades de información y llevarlas a una visualización para interpretarlas.

El proyecto a nivel macro está basado en la metodología propuesta por Kimball para el desarrollo de proyectos de BI, compuesta por las siguientes etapas: Planificación, Análisis de requerimientos. Modelado Dimensional, Diseño Físico, Diseño ETL, Especificación y desarrollo de aplicaciones de BI, adicionalmente diseño arquitectónico y selección e instalación de productos de BI.

Se adjunta la figura 7 indicando con puntos rojos cuales etapas no se ejecutan y con puntos verde las que sí.

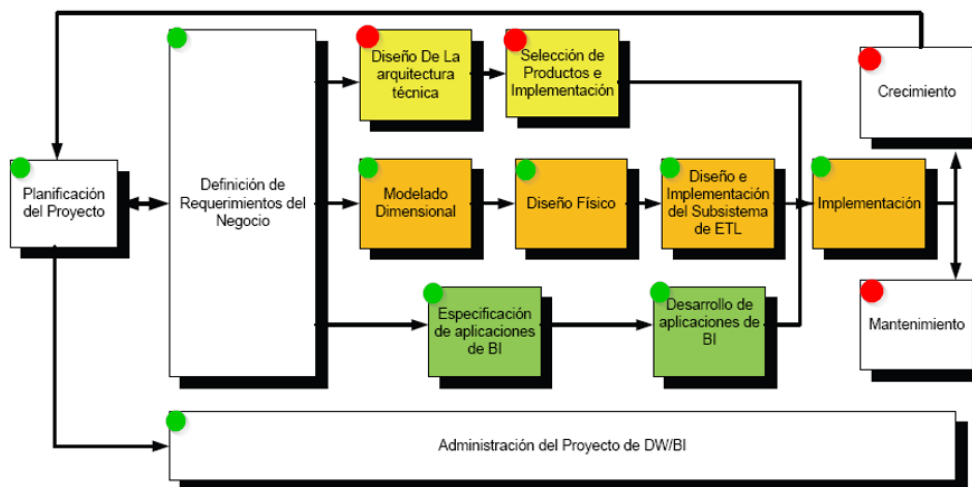


Figura 7. Tareas de la metodología Kimball (Kimball & Ross, 2011).

Luego se utilizó la técnica de creación de *starnet*. Una *starnet* busca plasmar un hecho (actividad objeto de análisis), dimensiones (puntos de vista desde los que se analiza el hecho) y medidas (métricas o indicadores) para cuantificar el hecho que surgen en el levantamiento de información y requerimientos; es el paso previo al modelo dimensional y su objetivo es tener un primer diagrama que apoye la construcción final del modelo dimensional. La figura 8 da una idea de lo que consiste un diagrama *starnet* para un hecho de ventas.

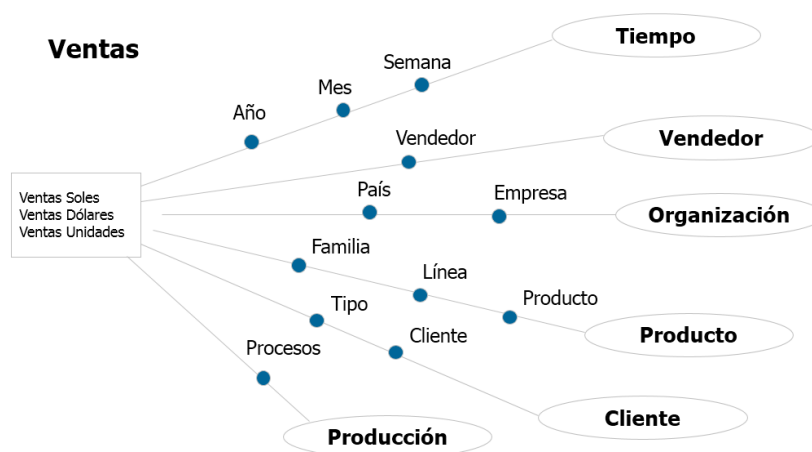


Figura 8. Ejemplo general de *Starnet*, autoría propia

Se desarrolló una aplicación con código en VBS para obtener los datos desde los equipos Cisco mediante la metodología en cascada para desarrollo de software. Esta metodología, si bien es considerada de las más antiguas en desarrollo de software, encajó bien en el desarrollo del proyecto al tener los requisitos muy claros. La Figura 9 presenta el ciclo de un desarrollo en cascada. (S. Pressman, 2010)

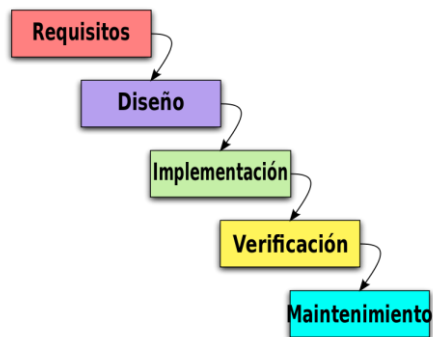


Figura 9. Metodología en cascada (S. Pressman, 2010)

Durante la ejecución de este código, se validaron las siguientes dimensiones de los datos: completitud, exactitud y conformidad. Por ejemplo, comparando cuales descripciones están estándar y cuáles no, y guardando este valor en la base de datos. Se utilizaron técnicas propias de verificación y limpieza de datos, se realizó un perfilamiento utilizando DQAnalyzer. El alcance del proyecto implicó realizar una verificación de calidad de datos, más no la solución a los errores encontrados.

El programa en VBS, dada la dirección IP de un dispositivo ingresó a este por SSH, contó los tipos de puertos por dispositivo y guardó dicha información y adicionalmente se verificó la calidad de los datos, luego de finalizar la sesión continua automáticamente con otro dispositivo.

5.1 Desarrollo del proyecto basado las metodologías planteadas.

Hay una sinergia entre las metodologías de Kimball y el modelo de desarrollo de tal manera que se completan, se resume las actividades más importantes de cada etapa.

Planificación del proyecto: Se busco el soporte y el apoyo del gerente de proyecto y se empezó a plantear el proyecto con alcances específicos y las necesidades más críticas que dieron idea a este proyecto. La etapa de planificación se cubre desde la elaboración misma de este documento, se socializa la idea a nivel corporativo, cliente y se buscó cumplir los key success factors antes de empezar.

Definición de requerimiento del negocio: Se realizaron entrevistas, reuniones y apoyados en el conocimiento del cliente se plasmaron las necesidades puntuales de información. Esta etapa tiene sinergia con la etapa de requerimientos del desarrollo de código dado que el resultado es el input para la información que se quiere ver. El análisis de requerimientos corresponde a lo planteado en el primer objetivo específico en. Las respuestas de estas preguntas cumplieron con la actividad de requerimientos y entendimiento del negocio.

Modelado dimensional: Se ejecutaron las actividades de definición de granularidad, métricas, diseño de dimensiones y relaciones.

Diseño Físico: Se diseñaron las tablas de la base de datos. Se revisaron tipos de archivos.

Subsistema de ETLs: Se obtuvieron datos de Excel y se modificaron manualmente y se guardó un documento con fórmulas utilizadas por si es requerido a futuro. Se desarrolló código para traer los datos desde la fuente, esta parte hace sinergia con la fase de diseño e implementación del desarrollo de código.

Implementación: Crear las bases de datos, y poblarla cargando archivos de Excel, ejecutar el código desarrollado, haciendo sinergia con la fase de implementación, verificación y mantenimiento.

Dentro de la implementación Se realizó la visualización de la información mediante una herramienta de exploración de los datos pues según se pudo observar en el marco conceptual otorga gran capacidad de filtrado, interacción fluida para buscar información relevante y facilidad para el usuario final.

La metodología Kimball para un proyecto de BI incluye como documentos entregables: Modelo de datos inicial de alto nivel, lista de atributos, diagrama de tablas de hechos, definición de campos de medida, diagrama de tablas de dimensiones, descripción de los atributos de las dimensiones. A esta información se suma el reporte de calidad de datos y el reporte de obtención de datos de los equipos Cisco, todos estos entregables se plasman en este documento en el siguiente capítulo de presentación y análisis de resultados.

Aplicaciones de BI: La actividad más importante se enfocó en elegir Hardware y Software basados en la practicidad y lo que más rápido disponible para desarrollar el prototipo sin tener restricciones corporativas o legales.

Las etapas de Diseño de la Arquitectura Técnica y selección de productos e implementación no hicieron parte del proyecto dado que el objetivo de este es desarrollar un prototipo y no requiere la ejecución de estos. Se deberá retomar y ejecutar estas dos etapas para definir las políticas, directrices estratégicas y definiciones de Hardware y procesos con responsables en caso de querer potencializar este prototipo y llevarlo a más procesos y más personas simultáneas. Para cumplir los objetivos del proyecto se pueden suplir con otras etapas y usando Hardware y Software disponibles.

Las etapas de Mantenimiento y Crecimiento no hacen parte del alcance de este proyecto debido a están enfocadas en la mejora continua de la solución del BI en la que se mejora de forma iterativa con la retroalimentación y nuevas necesidades de información, lo cual queda en responsabilidad del equipo de operación del servicio de red.

6. PRESENTACIÓN Y ANÁLISIS DE RESULTADOS

6.1 Requerimientos de información del negocio

Esta sección da cumplimiento al objetivo específico 1: Definir los requerimientos de información del negocio.

6.1.1 Entendimiento del negocio

Para comprender a fondo las necesidades de información, es importante conocer a profundidad los pormenores de cada negocio y las necesidades de los diferentes procesos que intervienen. Por tratarse de un servicio prestado a un cliente cuyos procesos están basados en ITIL se realizaron reuniones con los roles más importantes de la operación como son: líder de mesa de servicio actuando como Gestor de servicio, gerente de operaciones actuando como Propietario del servicio, gestor de problemas, gestores del conocimiento, y el autor de este documento, gestor de cambios y líder técnico.

En el anexo 1 se encuentran las preguntas realizadas con sus respectivas respuestas que surgieron durante estas sesiones. A continuación, se resumen los requerimientos indicando a cuál proceso beneficiaría directamente:

Gestión de Inventario

Se desea ver por cada sitio cual es el inventario, cantidad de equipos y modelos.

Se desea conocer el inventario agrupado en diferentes ciudades y departamentos o zonas.

Facturación

Se desea conocer la cantidad de *switches* de 24 o 48 puertos por sede, por ciudad y departamento.

Según el cobro de cada puerto acordado con el cliente se desea tener esta información en detalle, y que se pueda ver por sede, por ciudad, por departamento, por zona o en su totalidad.

Se desea visualizar cambios en los puertos en el último mes que puedan generar un cambio en la facturación.

Se desea ver la utilización de puertos FXO (*Foreign Exchange Office*).

Operación

Se desea ver la ubicación de todas las sedes geográficamente.

Se desea visualizar que enlaces tiene la sede, cual es el proveedor de última milla y poder filtrar por código de sede, nombre de sede o por nombre del proveedor o del enlace.

Se desea poder buscar una dirección mac y que aparezca cual sede, *switch* y cual puerto contiene este dispositivo.

Se desea poder visualizar que puerto de *switch* tiene una extensión en particular.

Se desea ver la cantidad de puertos disponibles y utilizados para así mirar el porcentaje de eficiencia.

Se desea ver la cantidad de puertos que tienen una inactividad mayor a 4 semanas.

Cajeros

Debido a la cantidad de cajeros se desea una visualización particular donde se visualice información relativa a la ubicación, transportadora, marca e indicadores de cantidades de cajeros según el tipo de sede.

6.1.2 Elaboración de *starnets*

En la figura 10 se plasmó el *starnet* desarrollado para este proyecto. A la izquierda dentro de un rectángulo se encuentran los indicadores que el cliente desea; encerrados en óvalos se observan las dimensiones desde las cuales se analizará la actividad de los puertos en operación, encima de las líneas que unen las dimensiones con los indicadores se encuentran los atributos más relevantes para resolver las necesidades de información.

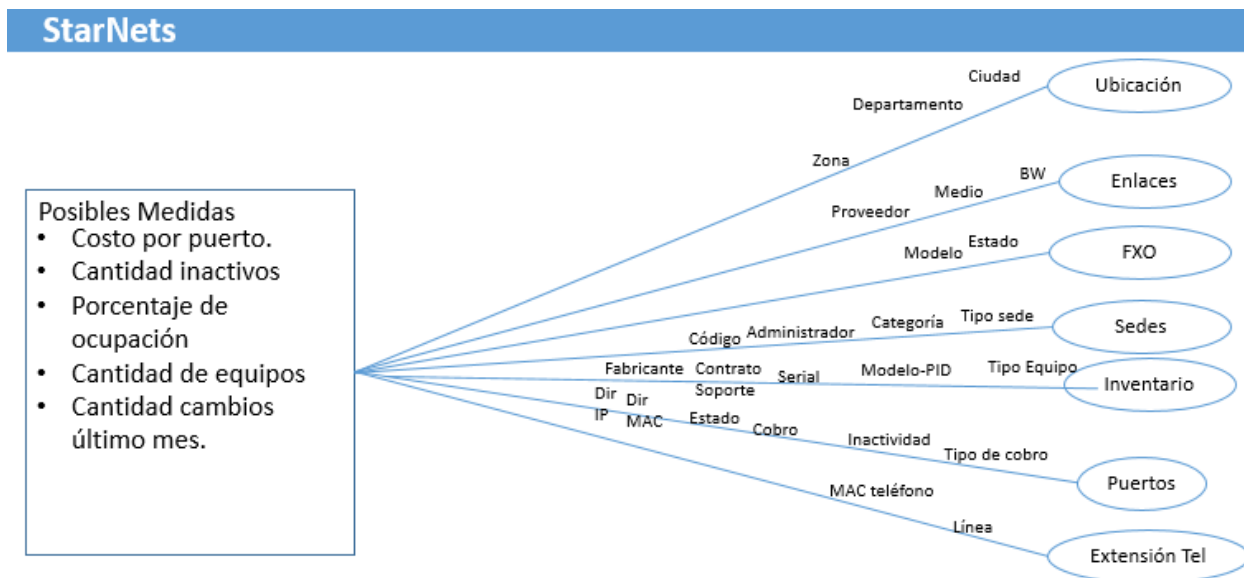


Figura 10. Starnet del proyecto, autoría propia

6.1.3 Fuentes de los datos.

Los datos necesarios se encontraron en diversos archivos de Excel, dentro de formatos definidos por la organización los cuales se modificaron para adaptarlos al esquema de una bodega de datos. Estos se encuentran reunidos todos en una carpeta compartida y con acceso exclusivo del personal que presta el servicio y no está disponible para otras áreas de la compañía.

Se definió que los datos de ubicación se tomarán desde el DANE, dado que allí se encuentran todos los municipios del país y está disponible para la descarga en la página web¹.

¹ <https://www.datos.gov.co/api/views/xdk5-pm3f/rows.csv?accessType=DOWNLOAD&bom=true&format=true>

Cada uno de los equipos Cisco que hacen parte de la prestación del servicio es considerado una fuente de datos. Se puede entender como un conjunto de fuentes y será accedido por SecureCRT como se explica en la sección 7.3.

El área de facturación compartió su archivo de conciliación para ser incluido en este proyecto. Es un archivo de Excel que contiene las sedes y cantidad de puertos facturados mensualmente con aprobación del cliente. Por solicitud de la compañía se trabajó para unificar los datos en esta base de datos y que esta se convierta en la nueva fuente de información y el punto donde se actualicen los futuros cambios.

6.1.4 Especificaciones de aplicaciones

Para la ejecución de este proyecto se utilizó Qlik Sense como herramienta de visualización, dado que la compañía cuenta con experiencia en dicha herramienta en el área de ciberseguridad. Cuenta con visualizaciones aptas para cumplir los requisitos determinados en el informe de requerimientos, algunas ventajas como son los mapas con diversas capas con facilidad y filtros que permiten buscar textos, adicionalmente Qlik se conecta fácilmente a la bodega de datos construida en Microsoft Access.

Al no contar con una CMBD ni con datos en otras bases de datos se realizó la construcción en Microsoft Access, se cuenta con la facilidad de que es una herramienta incluida en la suite de office 365 y la base de datos permite su modificación por medio de sentencias SQL. Microsoft

Access, cuenta con la facilidad de exportar o importar rápidamente tablas desde el aplicativo Excel, teniendo en cuenta que hay mucha información distribuida en este formato y tanto internamente como con el cliente se acostumbra el uso de Excel. Adicionalmente, permite fácilmente la migración hacia MS SQLSERVER en caso de ser requerido en el futuro. La figura 11 muestra la arquitectura de la solución de BI que se desarrolló en este proyecto.

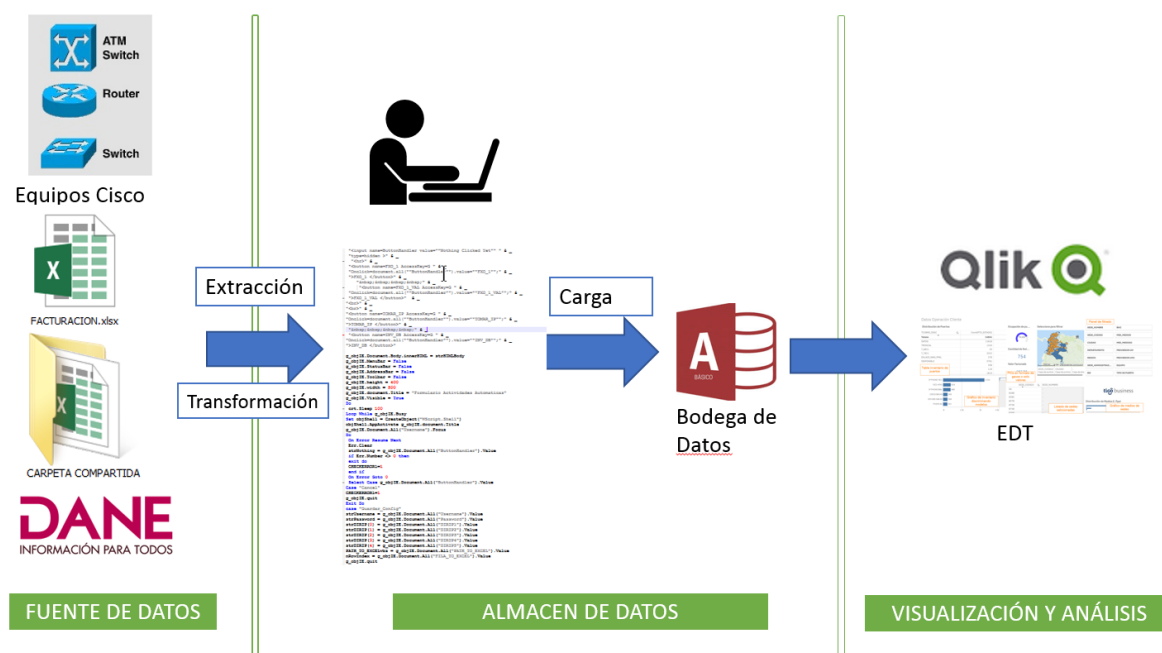


Figura 11. Arquitectura de BI del proyecto, autoría propia.

6.2 Construcción del modelo dimensional

Esta sección da cumplimiento al objetivo específico 2: Construir el modelo dimensional

6.2.1 Modelo lógico y dimensional

6.2.1.1 Definición de la granularidad

Se requiere conocer con el mayor nivel de detalle posible la información contenida en los dispositivos por lo que se define que la granularidad para esta solución de BI debe ser a nivel de puertos de red, en consonancia con la facturación que se hace por puerto, igualmente llegar a conocer cada dispositivo conectado requiere este nivel de detalle. Por ende, cada registro equivale a un puerto de red apto para prestar servicio al usuario final.

6.2.1.2 Definición de los hechos

Un servicio de *networking* se puede definir como la interconexión por enlaces WAN de sitios dispersos geográficamente por medio de dispositivos para tal propósito. Por tanto, el diseño consiste en interrelacionar estos tres elementos fundamentales (enlaces, sedes y dispositivos) con puertos y otras dimensiones que vengan a lugar y permitan obtener la granularidad deseada.

El estado de los puertos en operación es el hecho que se evalúa durante este proyecto para cumplir las diversas expectativas desde Operación, Facturación e Inventario. Se plantea la creación de la bodega de datos con diferentes tablas relaciones formando esquemas en estrella y en copo de nieve. Los cambios que se generan desde el cliente o desde la operación misma impactan a los puertos. Ya sea cambios de sede, fusiones, cambio de rol de los puertos, bajas,

altas, etc. Existen puertos *Foreign Exchange Office (FXO)*, puertos de *Routers*, Puertos de *Switches*.

6.2.1.3 Definición de métricas

Los siguientes indicadores serán incluidos en la visualización:

- Cantidad de *Switches* de 24
- Cantidad de *Switches* de 48
- Cuáles son los cambios que impactan facturación en el último mes.
- Cantidad de puertos inactivos más de 4 semanas.
- Porcentaje de eficiencia= Cantidad de puertos disponibles / Total de puertos
- Optimización= Cantidad de puertos disponibles / 24
- Porcentaje de enlaces en red propia.
- Cantidad de equipos a vencer soporte en los próximos 6 meses.

6.2.1.4 Implementación del modelo multidimensional.

Las dimensiones se crearon partiendo de lo revisado en la starnet y los requisitos funcionales, estas son las siguientes:

- DIM_ENLACE: Contiene todos los enlaces WAN de las sedes

- DIM_SEDES: Contiene todas las sedes que hacen parte de la base de datos, del cliente o propias.
- DIM_U_CIUDAD: Contiene las ciudades donde se presta servicio.
- DIM_U_DPTO: Contiene los departamentos donde se presta servicio.
- DIM_U_REGION: Contiene las regiones de las sedes según el cliente.
- DIM_PUERTOS: Contiene información de los puertos de *switch*.
- DIM_RT_PUERTOS: Contiene información de los puertos de *router*.
- DIM_INV: Contiene el inventario de dispositivos.
- DIM_FXO: Contiene el detalle de los puertos FXO.
- DIM_FACT_COBRO: Contiene el detalle del tipo de puerto y si se incluye en facturación o no.
- DIM_CONCILIACION: Contiene los detalles del proceso de conciliación y facturación con el cliente.
- DIM_ATMS: Contiene información de ATMs.

Las relaciones entre las tablas se crean con integridad referencial para garantizar que no queden datos en una tabla que no existan en otra y por ende no se pueda visualizar correctamente, dando lugar así a un esquema en copo de nieve.

LA DIM_CONCILIACION está conformada por el archivo de Excel entregado por facturación, no se realizan modificación con el script, pero se incluye el campo SEDE_COD_INT para relacionarlo con el resto de los elementos., se incluye en el modelo para permitir su visualización y comparación con los datos obtenidos en este proyecto.

La figura 12 muestra el modelo dimensional creado para ese proyecto.

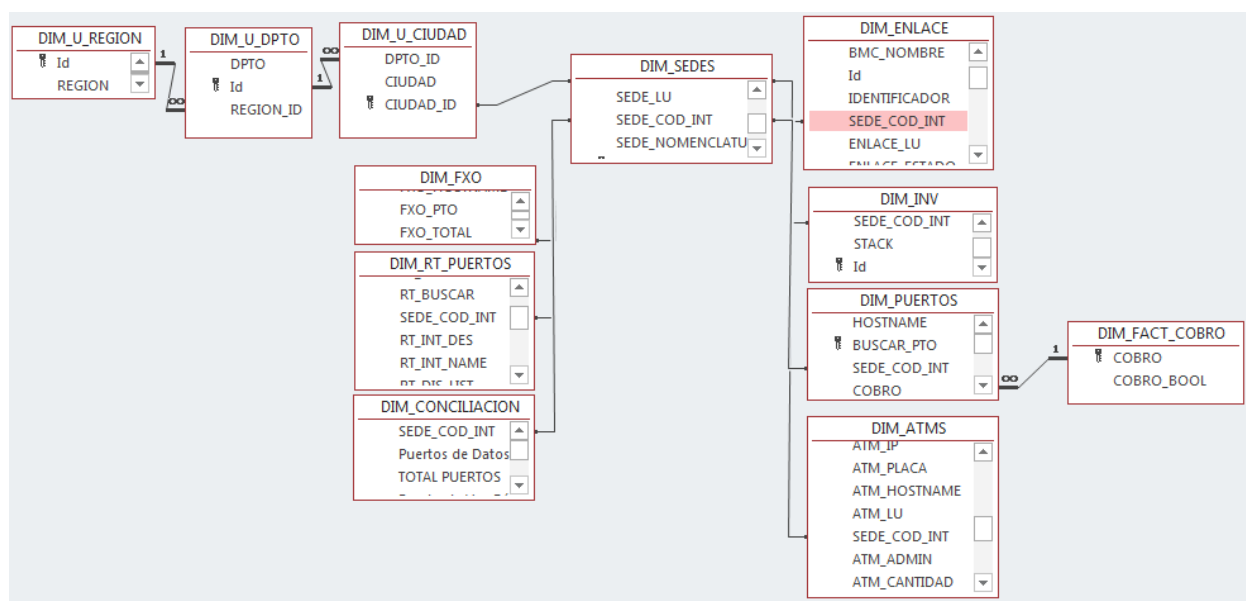


Figura 12. Modelo Dimensional, autoría propia.

6.2.2 Diseño físico

Se procede a crear las tablas en el aplicativo y se cambia a la vista de diseño en donde se nombran todos los atributos como muestra la figura 13.

Nombre del campo	Tipo de datos
CIUDAD_ID	Número
SEDE_COD_BANCO	Texto corto
CATEGORIA	Texto corto
SEDE_NOMBRE	Texto corto
ADMINISTRADOR	Texto corto
LATITUD	Número
SEDE_ESTADO	Texto corto
LONGITUD	Número
SEDE_LU	Fecha/Hora
SEDE_NOMENCLATURA	Texto corto
IVR	Sí/No

Figura 13. Creación de tablas en Access., autoría propia.

Los indicadores o métricas no fueron incluidos en la bodega de datos y se calculan directamente en la herramienta de visualización, lo cual es posible sin problema dado que el volumen de datos no es muy alto.

6.2.3 Carga inicial de los datos

La información unificada en archivos de Excel se cargó a la bodega de datos mediante la pestaña de “Datos Externos” como se muestra en la figura 14.

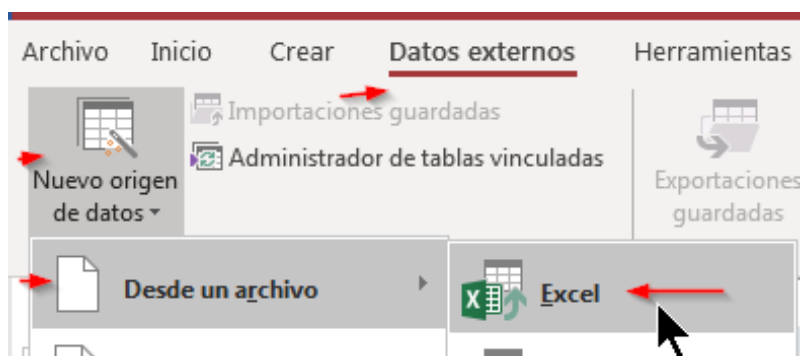


Figura 14. Carga en Microsoft Access desde libros de Excel, autoría propia.

Las dimensiones cargadas fueron: DIM_ENLACE, DIM_FACT_COBRO, parte de DIM_INV, DIM_SEDES, DIM_U_CIUADAD, DIM_U_DPTO, DIM_U_REGION, DIM_CONCILIACION parte de DIM_ATM.

Las otras dimensiones, DIM_FXO, DIM_PUERTOS, DIM_RT_PUERTOS y algunos campos de DIM_INV y DIM_ATM son pobladas por medio del script desarrollado durante este proyecto para capturar la información desde los equipos, el cual se detalla en la sección 7.3.

6.3 Desarrollo de aplicación de extracción de información

Esta sección da cumplimiento al objetivo específico 3: Desarrollar una aplicación que permita la extracción de la información de los dispositivos Cisco en producción de forma automática.

6.3.1 Diseño e implementación de ETL

Como objetivo de este proyecto se planteó desarrollar un script para obtener información directamente desde los dispositivos Cisco y de esta manera extraer los datos de la fuente, realizar las transformaciones necesarias para así cargarla en la bodega de datos, por tanto, este realiza todo el proceso ETL para esta fuente de datos.

Las otras fuentes de datos consistían en archivos de Excel, los cuales fueron preparados y ajustados de forma manual a los campos definidos para las diferentes dimensiones.

El desarrollo de código se desarrolla bajo la metodología en cascada explicada anteriormente.

Los requisitos son determinados por el modelo lógico y dimensional con el objetivo de obtener la información para llenar cada uno de los campos de las tablas que se pueden obtener desde el dispositivo.

Para el diseño se planteó construir el código por etapas tal como se explica en la figura 15.

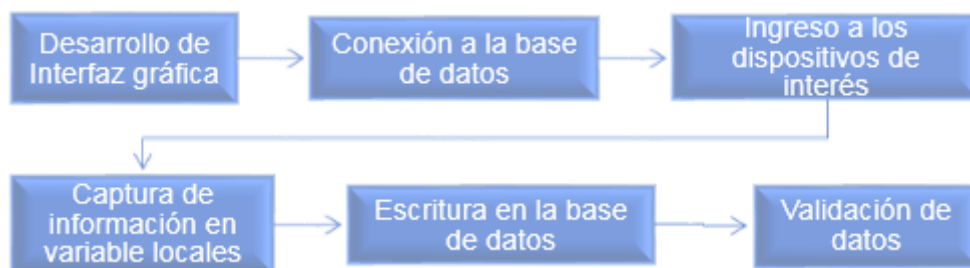


Figura 15. Etapas de construcción del script, autoría propia.

La implementación consiste en el desarrollo de cada uno de estos hitos, para esto se basa en el documento de scripting de SecureCRT el cuál se encuentra en la página web del fabricante.²

6.3.1.1 Desarrollo de la interfaz gráfica

Se desarrolla una interfaz web para que cualquier usuario pueda ejecutar el código. Esta interfaz web se ejecuta con *Microsoft internet explorer* como muestra la figura 16.

² https://www.vandyke.com/support/tips/scripting/scripting_essentials.pdf

Figura 16. Formulario para lanzar scripts, autoría propia.

El ingeniero debe ingresar el usuario y contraseña (usuario y *password*) válidos para ingresar a los dispositivos de red, con los cuales se autentica de forma segura en los equipos del cliente. En segunda instancia se debe ingresar un rango de ejecución, esto es el rango de registros de la TABLA_IP previamente cargada. Si se desea lanzar en más de 50 dispositivos se recomienda usar rangos de 50 y lanzar 10 veces el script en 10 ventanas distintas de SecureCRT.

Los botones creados en la interfaz realizan las siguientes acciones:

- **Guardar Config:** Guarda la configuración en texto.
- **Cancelar:** Cierra el formulario sin ejecutar más acciones.
- **FXO_1:** Inicia el proceso de configuración de puertos FXO para comprobar su estado.
- **FXO_1_VAL:** luego de ejecutar FXO_1 se debe usar este botón en el mismo rango de ejecución para comprobar el estado de los puertos FXO y llevarlos a la base de datos.

- TOMAR_IP: Captura todas las direcciones IPv4 que tenga los dispositivos a los que ingresa y los guarda en la base de datos.
- INV_DB: Entra a todos los equipos de las sedes y captura todos los campos para llevarlos a las diferentes tablas de la base de datos creada durante este proyecto.

6.3.1.2 Conexión a la base de datos

Se realiza la conexión a la base de datos usando un ADO *Connection object*. se muestra el código fuente en la figura 17.

```

-----
DataConn = "D:\Onedrive\RAMON\Job\DATOS_BI.mdb"
Set Conn = CreateObject("ADODB.Connection")
ConStr = "Driver={Microsoft Access Driver (*.mdb)};DBQ=" & DataConn & ";"
Set rs = CreateObject("ADODB.Recordset")
Set rs2 = CreateObject("ADODB.Recordset")
Conn.open(ConStr)
strSQLQuery = "SELECT * FROM TABLA_IP"
rs.Open strSQLQuery, Conn, 3, 3

```

Figura 17. Código de conexión a la base de datos, autoría propia

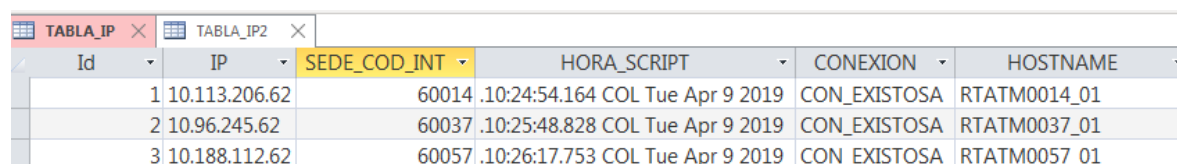
6.3.1.3 Ingreso a los dispositivos de interés

Se realiza la conexión vía SSH a un listado de dispositivos, para esto se crean 2 tablas adicionales en la base de datos.

TABLA_IP, en esta tabla se debe precargar la dirección IP y el campo SEDE_COD_INT que corresponda a dicha IP. Los campos HORA_SCRIPT, CONEXIÓN y HOSTNAME se llenan automáticamente dependiendo del resultado del intento de conexión.

La segunda tabla se llama TABLA_IP2, en esta, todos los campos se llenan de forma automática y permite entrar a todos los dispositivos que estén interconectados en la sede a la cuál ingresa, adicional a los campos ya mencionados en TABLA_IP cuenta con los campos EQUIPO_ORIGEN, PENDIENTE y BUSCAR los cuales se utilizan para validar donde se encontró este equipo, y se marca como pendiente hasta que se ingrese a este dispositivo.

La figura 18 muestra un ejemplo de 3 sitios a los cuales ya pasó el código realizando una captura de información.



Id	IP	SEDE_COD_INT	HORA_SCRIPT	CONEXION	HOSTNAME
1	10.113.206.62	60014	.10:24:54.164 COL Tue Apr 9 2019	CON_EXISTOSA	RTATM0014_01
2	10.96.245.62	60037	.10:25:48.828 COL Tue Apr 9 2019	CON_EXISTOSA	RTATM0037_01
3	10.188.112.62	60057	.10:26:17.753 COL Tue Apr 9 2019	CON_EXISTOSA	RTATM0057_01

Figura 18. Tabla IP, autoría propia.

6.3.1.4 Captura de información en variables locales

Se generó código utilizando comandos show de Cisco para enrutadores y switches, y se almacenaron en variables o arreglos según corresponda, la figura 20 muestra cómo se capturó la versión de sistema operativo de un equipo. La mayor parte del código desarrollado se enfocó en obtener toda la información de interés, darle el tratamiento adecuado y realizar las comprobaciones necesarias para llevarlo a la base de datos. La figura 19 muestra un ejemplo de cómo capturar datos.

```
' OBTENER IOS
enviar("sh ver " & chr(124) & " i imag")
crt.Screen.WaitForString "#"
Fila4=crt.Screen.Get(crt.Screen.CurrentRow-1, 1, crt.Screen.CurrentRow-1,200)
```

Figura 19. Código para capturar el IOS, autoría propia.

6.3.1.5 Escritura en la base de datos

Se escribe en la base de datos mediante sentencias SQL, dependiendo de si existe o no un registro para la información se utilizan sentencias *insert* o *update*. La figura 20 muestra un ejemplo:

```
Conn.execute "INSERT INTO DIM_INV _
& "(ESTADO_ID,HOSTNAME,PID,SERIAL,IP_GESTION,OS,MAC_BASE,POE,INV_NUEVO,SEDE_COD_INT,INV_LU,COMPATIBLE_SCRIPT,SERIAL_CHASIS) VALUES " _
& "('1','" & HOSTNAME & "',''" & PID & "',''" & SN & "',''" & strIP & "',''" & OS_EQUIPO & "',''" & BASEMAC & "',''" & POE & "',''" & SEDE_COD_INT & "'" _
& "',''" & HORA_SCRIPT & "',''" & SI & "',''" & SERIAL_CHASIS & "'"")"
```

Figura 20. Código con sentencias SQL para poblar las dimensiones.

6.3.1.6 Validación de datos.

Se revisa en las diferentes tablas de la base de datos que los datos almacenados correspondan según lo esperado, que cumplan con los estándares definidos y estén listos para poder ser utilizados por la herramienta de visualización.

6.3.1.7 Verificación.

Durante el desarrollo de este proyecto el código se probó y verificó varias veces con 10 sitios de prueba, esto debido a que no era posible probar de otra manera y cambios debido a errores o mejoras fueron realizados después de evidenciarlos en tiempo de ejecución. Dando cumplimiento al entregable “reporte de pruebas del código” se detalla:

- Se presentaron inconvenientes con algunas cadenas de texto y algunos comandos con caracteres especiales, por lo que se requirió tratamientos especiales a algunas cadenas de texto.
- Inconvenientes en tiempo de ejecución al realizar una acción antes de terminar la anterior, por lo que fue necesario adicionar tiempos de espera en el código. Fue necesario colocar 30 segundos de espera para cuando no se logra conectar por SSH y realizar una validación extra para determinar que no había entrado al dispositivo.
- Se presentaron inconvenientes al enviar los comandos a los equipos Cisco, se perdía sincronía y se saltaban algunas líneas, fue necesario crear una función particular para enviar los comandos a los dispositivos.
- Para lanzar el código se debe maximizar hacia lo ancho del programa SecureCRT para evitar saltos de línea que puedan causar errores en tiempos de ejecución. Esto pasa con salidas de comandos muy largas, también se configuró el comando de cisco "term wid 512" para aprovechar el ancho máximo posible de la consola.

- Se presentaron inconvenientes debido a la diversidad de modelos, tipos de equipos y sistemas operativos en producción por lo que fue necesario generar código para evitar errores por comandos diferentes o inexistentes dependiendo de la versión de sistema operativo. Fue necesario generar algunos comandos específicos para algunos modelos.

Los tiempos promedios de ejecución son de 5 minutos para *Switches*, 4 minutos para *routers* pudiendo este ser hasta de 10 minutos cuando hay enlaces de alto tiempo de respuesta.

El proveedor recomienda usar “`crt.Screen.WaitForString #`” tras cada comando, fue necesario esperar por el HOSTNAME del dispositivo debido a inconvenientes con algunos equipos que tenían carácter “#”.

6.3.1.8 Mantenimiento.

No hace parte del alcance de este proyecto realizar mantenimiento continuo al código desarrollado. Una vez cumplidos los requerimientos y probado en el 100% de los sitios de interés se da por terminado el ciclo de desarrollo.

Para la actualización de los datos en la base de datos de Microsoft Access se propone al interior de la compañía establecer un proceso a ejecutar los últimos 4 días del mes con una dedicación de 20 horas para estar al frente de la ejecución y lanzamiento de los scripts en la totalidad de dispositivos y mantener la información actualizada.

Recorrer la totalidad de dispositivos toma 2 días hábiles con 7 sesiones simultáneas y es una actividad supervisada.

6.4 Análisis de calidad de los datos

Para la realización del análisis de la calidad de datos de la información de este proyecto se utilizó el programa DQ Analyzer 11 de Ataccama,, se esperó a poblar los registros de las dimensiones para realizar esta revisión en los campos de interés.

6.4.1 Análisis de calidad de datos Dimensión Puertos

Se observó que para la dimensión puertos no hay *Nulls*, el campo BUSCAR_PTO tiene 100% únicos. Cobro tiene 17 valores según lo esperado. Se observa que el total de *Switches* son 729. Ver figura 21.

Expression ^	Type	Domain	Non-null	Null	Unique	Distinct	Min	Median	Max
BUSCAR_PTO	STRING	pattern	23,278	0	23,278	23,278	FCQ1520Z...	FOC1832W...	FOT1917S0...
COBRO	STRING	enum pattern	23,278	0	17	17	ACCESS-P...	DATOS	T_VVX400
HOSTNAME	STRING	pattern	23,278	0	1	719	SW_CAP_S...	SWSUC042...	SWSUCVAL...
INACTIVIDAD	INTEGER	enum	23,278	0	0	2	0	0	1
INT_CATEGORIA	STRING		23,278	0	16	43	ARANA	NO_ESTAN...	VOZ
INT_DEAD_VLAN	INTEGER		23,278	0	2	33	0	0	987
INT_DES	STRING		23,278	0	1,582	1,778	ACCESS PO...	SEGURIDAD	WAN-FWC...
INT_MAC_DATOS	STRING	specval	23,278	0	9,815	9,885	0000.cd2a....	No Aplica	No Aplica
INT_MAC_VOICE	STRING	specval	23,278	0	4,023	4,024	0004.f2a1....	No Aplica	No Aplica
INT_MODE	STRING	enum pattern	23,278	0	0	3	access	access	VACIO
INT_NAME	STRING	pattern	23,278	0	128	354	Fa0/1	Fa0/24	Te5/0/2
INT_STATUS	STRING	enum pattern	23,278	0	0	5	down disabled	up connected	up connected
LAST_INPUT	STRING	pattern	23,278	0	263	608	000000	000001	never
LU	STRING	pattern	23,278	0	1	720	.00:05:53....	.12:42:42....	.23:39:15....
PTO_CAMBIO_FAC	STRING	integer enu...	23,278	0	0	1	1	1	1
SEDE_COD_INT	STRING	integer pat...	23,278	0	1	404	10001	10523	66066
SW_SERIAL	STRING	pattern	23,278	0	1	729	FCQ1520Z...	FOC1832W...	FOT1917S03B
VLAN_DATOS	INTEGER		23,278	0	11	74	0	8	1,547
VLAN_VOICE	INTEGER	enum	23,278	0	1	8	0	0	586

Figura 21. Análisis calidad de datos dimensión Puertos, autoría propia.

Según se mencionó en la metodología el código identifica un puerto que no cumple con el estándar en la descripción y se encuentra que hay un 69.75% de descripciones fuera del estándar en los equipos (Ver figura 22).

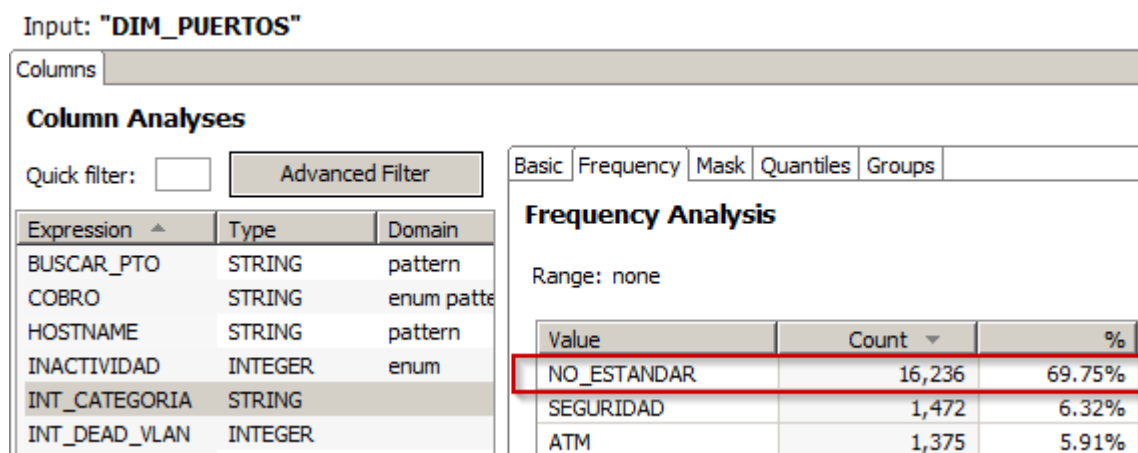


Figura 22. Análisis calidad de datos de frecuencia campo INT_CATEGORIA, autoría propia.

6.4.2 Análisis de calidad de datos de la dimensión RT_PUERTOS

Para la dimensión RT_PUERTOS en la que se espera que no se tengan duplicados, sorprende ver que los valores únicos son bastante menos que la cantidad total por lo que se revisa puntualmente y se observa que la mayoría de direcciones IP está 2 veces, esto es debido a que hay una cantidad importante de *routers* con servicios de telefonía que capturó el script con el mismo nombre de la interfaz *Loopback* 0 por lo que se considera normal hasta 2 repeticiones. Para más de 2 si son duplicadas y representan un inconveniente de calidad de datos y posiblemente de operación del servicio. La figura 23 muestra el resumen general.

Input: "DIM_RT_PUERTOS"

Columns

Column Analyses

Quick filter:

Expression	Type	Domain	Non-null	Null	Unique	Distinct
RT_SERIAL	STRING	pattern	24,114	0	0	1,580
RT_HOSTNAME	STRING	pattern	24,114	0	0	1,579
RT_BUSCAR	STRING		24,114	0	24,114	24,114
RT_INT_NAME	STRING	pattern	24,113	1	851	1,145
RT_INT_DES	STRING		24,114	0	8,590	9,973
RT_DIS_LIST	INTEGER	enum	24,114	0	0	2
RT_INT_IP	STRING	specval pat...	24,114	0	17,731	17,854
RT_INT_MAS	STRING	enum pattern	24,114	0	0	11
RT_LU	STRING	pattern	15,067	9,047	3	1,248
RT_INT_MAS_NUM	INTEGER	enum	24,114	0	0	10

Figura 23. Análisis calidad de datos dimensión RT_Puertos, autoría propia.

En la figura 24 Se observan 3 direcciones ip repetidas más de 2 veces por lo que corresponde a 0.03% de los 24.144 registros.

Frequency Analysis

Range: none

100 most common values:

Value	Count	%
VACIO	6,135	25.44%
192.168.255.254	4	0.02%
172.30.252.1	3	0.01%
172.31.0.2	3	0.01%
10.126.101.65	2	0.01%

Figura 24. Análisis calidad de datos frecuencia de RT_INT_IP, autoría propia.

6.4.3 Análisis de calidad de datos de la dimensión SEDES

La dimensión SEDES cuenta con 731 registros, pero el aplicativo la tomó como si tuviera muchos más, esto no es un inconveniente para analizar, en la figura 24 se observa que hay un

registro que no tiene la ciudad y hay 6 registros que no tienen la latitud y longitud por lo que no va a permitir su visualización en la herramienta de exploración de los datos Qlik.

Se observó que el segundo campo SEDE_COD_INTERNO tiene 731 valores únicos según lo esperado.

Input: "DIM_SEDES"

Columns

Column Analyses

Quick filter:

Expression	Type	Domain	Non-null	Null	Unique	Distinct
CIUDAD_ID	INTEGER		730	1,344	198	241
SEDE_COD_...	STRING	integer	731	1,343	731	731
SEDE_NOMB...	STRING		2,074	0	2,035	2,054
CATEGORIA	STRING	enum pattern	2,074	0	0	6
ADMINISTR...	STRING	enum pattern	2,074	0	1	4
LATITUD	FLOAT		725	1,349	662	666

Figura 24. Análisis calidad de datos dimensión SEDES, autoría propia.

Se revisó puntualmente la latitud y se observa que hay 59 duplicados como muestra la figura 25, dos sitios no repiten latitud exacta por lo que también es un error en la calidad de los datos.

Type	Count	%
Null	1,349	65.04%
Non-null	725	34.96%
Duplicate	59	2.84%
Distinct	666	32.11%
Non-unique	4	0.19%
Uniqate	662	31.92%

Figura 25. Análisis calidad de datos Nulos, únicos y duplicados de Latitud, autoría propia.

Esta dimensión cuenta con 8.89% de registros erróneos para el campo latitud y el mismo valor para el campo longitud.

Un campo no tuvo la ciudad asociada. Representa el 0.13% de los registros.

6.5 Implementar reportes y visualización.

Esta sección da cumplimiento al objetivo específico 4: Implementar reportes y visualización con la base de datos levantada.

6.5.1 Herramienta de visualización

Para la visualización de la información se utilizará Qlik Sense Desktop. A continuación, el paso a paso para cargar la información:

1. Crear una cuenta en Qlik
2. Instalar el programa.
3. Crear una APP
4. Utilizar un conector OLE DB, allí se escoge Microsoft office 12.0 Access Database Engine OLE DB Provide y se indica la ruta de la base de datos.
5. Seleccionar las tablas y campos que se quiere cargar.
6. Realizar las asociaciones tal como fueron diseñadas previamente.

Se verificó que el modelo de datos que ofrece la aplicación Qlik corresponda al modelo lógico diseñado tal como se ve en la figura 26.

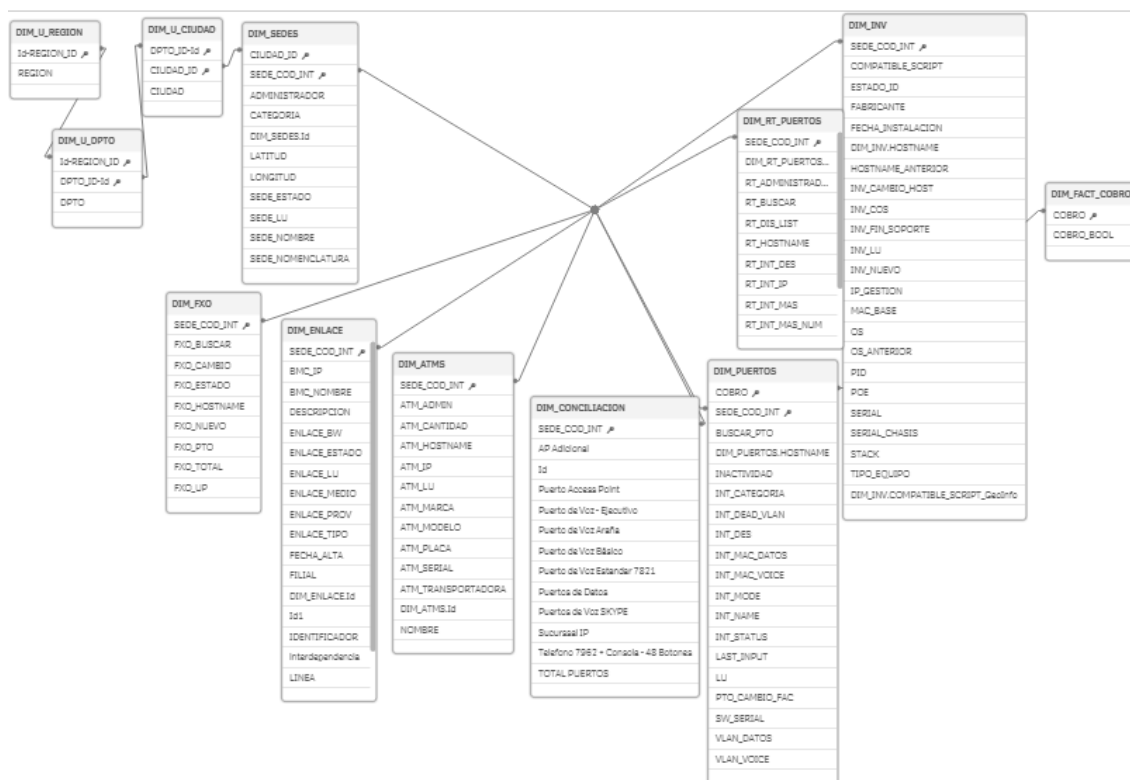


Figura 26. Modelo de datos en Qlik, autoría propia.

6.5.2 Creación de reportes y visualización.

Se crearon diferentes visualizaciones para dar cumplimiento a los requerimientos de información. Se elaboró un reporte para Inventario, uno para Operación, uno para Facturación y uno para ATMs.

La figura 27 muestra el reporte de inventario. Se detallan los elementos.

- Mapa: Un mapa con diferentes casas de área para los departamentos, puntos para los municipios de Colombia y puntos, pero de otro color para las sedes del cliente (1).
- Indicador de sede: Debajo del mapa se encuentra una tabla que muestra las sedes con el código seleccionado (2).

- Tipo Dispositivo: Muestra el inventario de la selección según el tipo de dispositivo (3). PID Dispositivo: Muestra el Product ID, que es lo mismo que el modelo del dispositivo según el fabricante (4).
- Tabla OS: Tabla que muestra el sistema operativo que tienen actualmente los dispositivos (5).
- Panel de filtrado: Panel para realizar filtros y cambiar los datos presentados según lo deseado por el usuario (6).
- Diagramas de distribución geográfica: Gráficas de torta y de barras para según la selección mostrar la cantidad por ciudad, departamento y región (7).

The screenshot displays the Qlik Sense Desktop interface for an inventory dashboard. The dashboard is titled "Inventario" and features several key components:

- Map (1):** A map of Colombia with a red box highlighting the "Medellin" area. The map is titled "Selección para filtrar" and includes a legend for "SEDE_COD_INT" with layers for "Point layer" and "Area layer".
- Filters (2):** A search bar for "SEDE_COD_INT" with the value "10375" and a search bar for "SEDE_NOMBRE" with the value "LA AGUACATALA".
- Tipo Dispositivo (3):** A horizontal bar chart showing the distribution of device types. The x-axis represents the count (0 to 4), and the y-axis lists device types: FXO (4), SWITCH (2), and FUENTE (1).
- PID Dispositivos (4):** A horizontal bar chart showing the distribution of Product IDs. The x-axis represents the count (0 to 4), and the y-axis lists Product IDs: VIC2-4FXO (4), CISCO2901K9 (1), and ISM-SRE-300-K9 (1).
- Tabla OS (5):** A table showing the operating systems (OS) for the selected devices. The table has columns for "TIPO" and "OS".
- Right Panel (6):** A vertical list of filters for "SEDE_NOMBRE", "SEDE_ADMINISTRADOR", "CIUDAD", "DEPARTAMENTO", "REGION", "CODIGO", "CATEGORIA", and "PID".

TIPO	OS
CISCO2901K9	c2900-universalk9-mz.SPA.156-3.M4.bin
WS-C2960+24PC-L	c2960-lanbasek9-mz.150-2.SE9.bin
WS-C2960+24TC-L	c2960-lanbasek9-mz.150-2.SE9.bin

Dist Geográfica

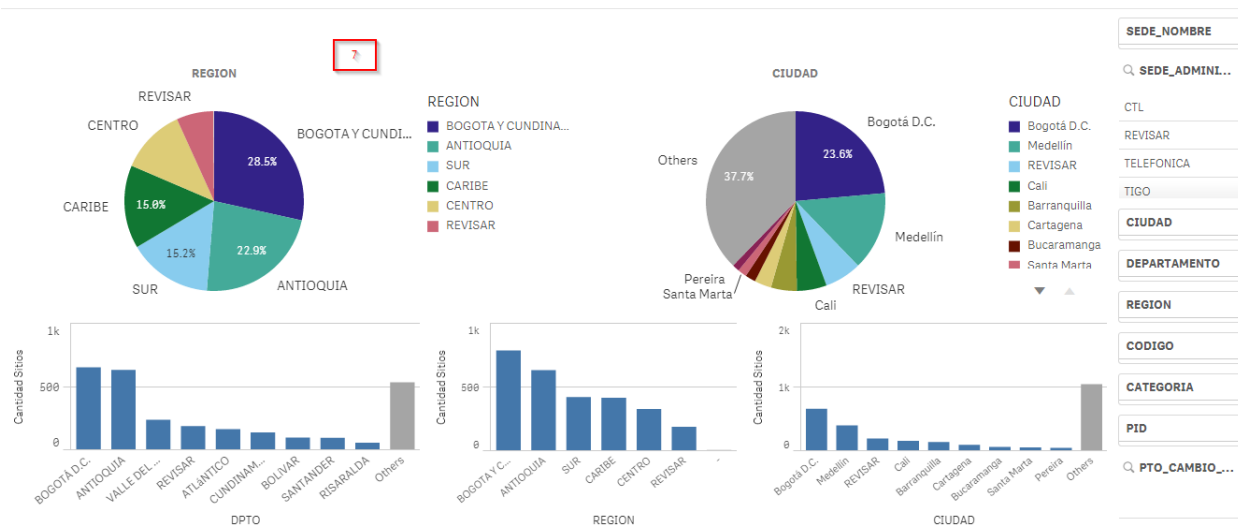


Figura 27. Reporte de Inventario, autoría propia.

La figura 28 muestra el reporte de operación, se repite el mapa y el panel de filtrado del reporte anterior. Se detalla:

- Tabla Cobro: Muestra el tipo de puerto según el cobro y la cantidad (1).
- Búsqueda de MAC: Muestra 2 campos para introducir la dirección mac de datos o voz, abajo se encuentra la tabla en la cual se muestra que sede, equipo, interfaz y descripción tiene la dirección mac escrita (2).
- Indicadores de puertos: Muestra todos los indicadores deseados (3).
- Enlace principal: Muestra detalles del enlace principal de las sedes seleccionadas (4).
- Enlace *Backup*: Muestra detalles del enlace *backup* de las sedes seleccionadas (5).
- Tabla Sede: Muestra las sedes y su código de la selección (6).

- Gestión de la selección: Muestra las direcciones IP de los equipos en operación de las sedes seleccionadas (7).
- Subredes de la selección: Muestra todas las subredes de las sedes seleccionadas (8).

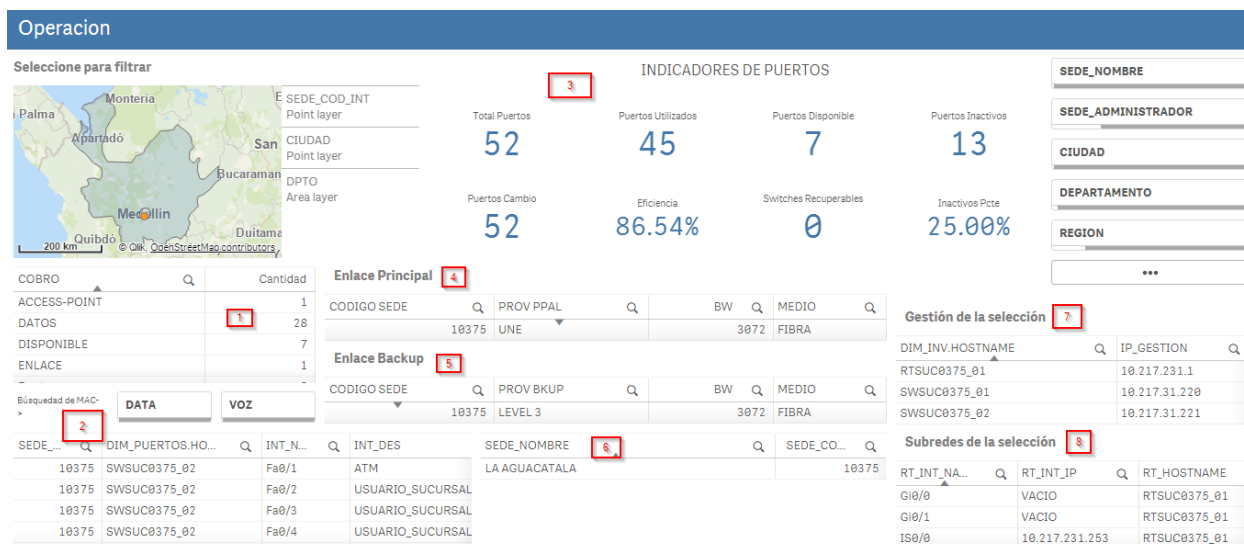


Figura 28. Reporte Operación, autoría propia.

La figura 29 muestra el reporte de facturación, para este se decidió construir una gran tabla que combinara los datos del archivo de conciliación con los datos obtenidos por el script para su comprobación, se trata cada dato en 3 columnas; la primera es el valor que está en facturación (marcado con F), la segunda columna es de los datos obtenidos de los equipos y la tercera columna (Marcado con D) indica la diferencia entre los dos. Si la diferencia es negativa, ésta es a favor del prestador del servicio, si es positiva es a favor del cliente. También se incluyen paneles de filtrado, mapas, y una sección con indicadores deseados, se destaca en el panel de filtrado un campo para identificar cuáles son los puertos que generan cambio de facturación.

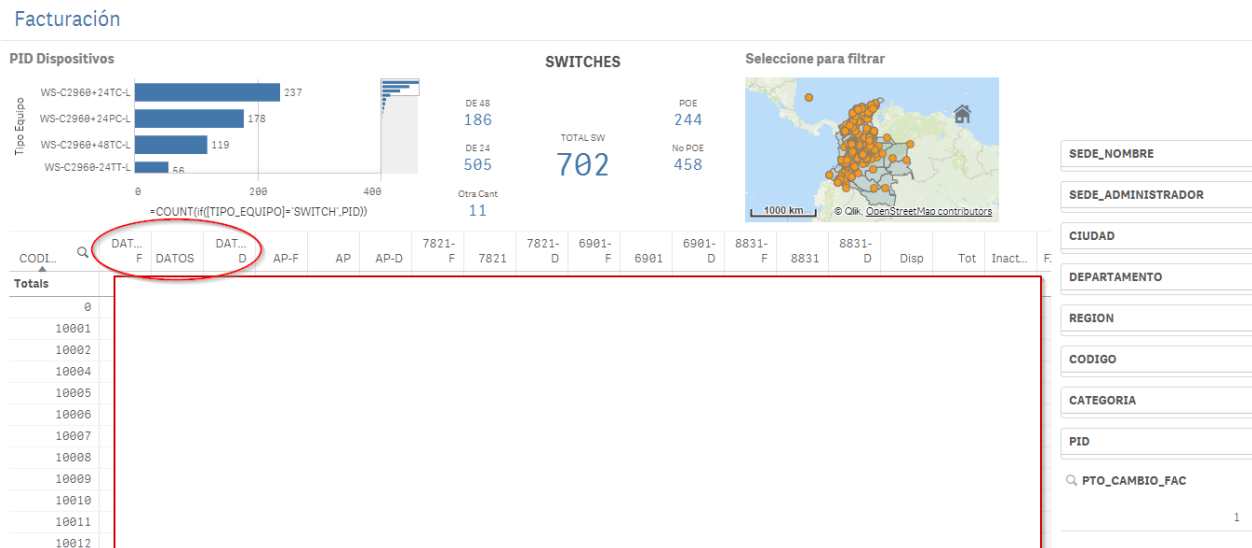


Figura 29. Reporte Facturación. Autoría propia.

La figura 30 muestra el reporte de ATMs, se mantienen los elementos de mapa y de filtrado para navegar por el reporte. Se detalla

- Nombre Sede: Indica el nombre de las sedes seleccionadas (1).
- Cajeros por sitio: Muestra la cantidad de cajeros de un sitio según el código (2).
- Cajeros x Región: Muestra la distribución de cajeros en las regiones (3).
- Distribución por Marca: Muestra la distribución de cajeros según la marca de este (4).
- Distribución por transportadora: Muestra la distribución según la transportadora de valores que atiende la selección (5).
- Indicadores: Muestra los indicadores según la selección para esta sección (6).

CAJEROS

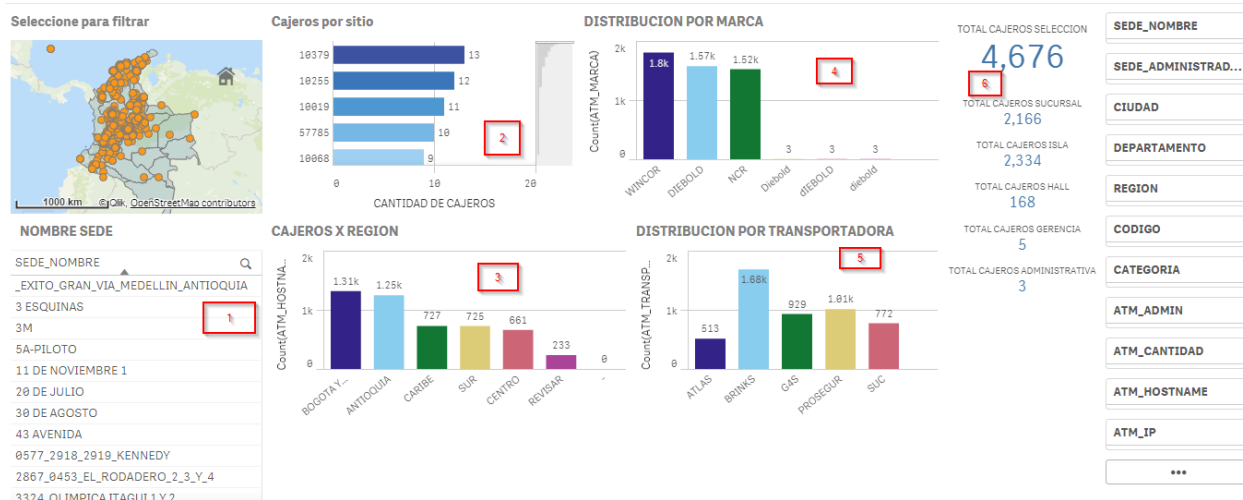


Figura 30. Reporte Cajeros, autoría propia.

6.5.3 Actualización de los datos.

Para realizar la actualización de la información visualizada en Qlik en la sección de Data Manager se presiona el icono de actualización y Qlik actualizará los registros que contiene la base de datos de MS Access (Ver figura 31).

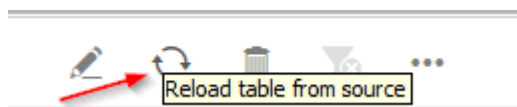


Figura 31. Actualizar data de la visualización. Autoría Propia.

Los registros de la base de datos de MS Access se actualizan con nuevas ejecuciones de los scripts o cambios en las tablas por un usuario.

6.5.4 Impacto corporativo de la solución de inteligencia de negocios.

El conocimiento de este tipo de soluciones aún es muy bajo en la compañía y en la operación particular, la analítica sigue siendo un tema novedoso y cautivador. Las diferentes áreas se interesan al ver el potencial y los resultados e inmediatamente se disparan diversas actividades para mejorar el proceso y ajustar lo que se evidencia de la solución de BI.

La figura 32 muestra un ejemplo de esto, (se ocultan nombres y fotos por privacidad), uno de los objetivos de este proyecto era apoyar la facturación y se puede ver como en la primera iteración ya se generan actividades al interior.

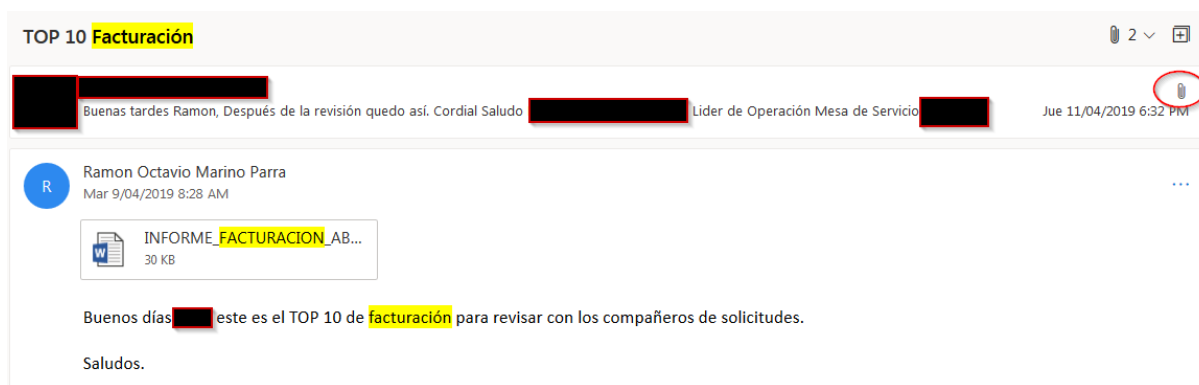


Figura 32. Actividad de revisión generada por resultados de este proyecto, autoría propia.

Se observa en la figura 33 un 77.63% de ocupación de puertos lo que implica una oportunidad de optimización para retomar equipos y redistribuir estos en las bodegas como parte del Stock y evitar una compra o bien para distribuir a otros clientes. Este valor fue muy llamativo para los

líderes de la operación e incluso a nivel de supervisor para cuestionarse sobre la eficiencia en otros proyectos.



Figura 33. Indicadores de puertos

7. CONCLUSIONES

7.1 Conclusiones de la ejecución del proyecto de BI.

La utilización de *frameworks* y metodologías establecidas, si bien tiene mucho tiempo siguen siendo totalmente vigentes y facilitaron el desarrollo de este proyecto. Teniendo muy claro el paso a paso y los hitos a obtener durante el mismo para su completitud se logra cumplir con los objetivos planteados en el tiempo estipulado, así mismo se revalida que el interés y el apoyo de los líderes de procesos o áreas es crucial.

La fase de requerimientos de información del negocio fue crucial para obtener una visión global del proyecto y saber con precisión como afrontar y que información es necesario incluir en el modelo dimensional para llevarlo a cabo con éxito.

Se logró desarrollar un código funcional para la captura de datos directo desde la fuente superando obstáculos relacionados con el software de captura y las diferentes versiones de la línea de comandos de los equipos Cisco que varía según modelo y versión de sistema operativo, obteniendo así información valiosa sin realizar inversión alguna en herramientas y con la posibilidad de obtener a la medida lo que se requiere y pueda extraer.

La posibilidad de obtener información directa de los equipos tiene un potencial grandísimo para su posterior análisis, no solamente para obtener información sino para realizar configuraciones masivas, *checklist* automáticos y procedimientos, contar con un *skill* programador en áreas de *Networking* no es lo más usual pero esto disminuye la necesidad de contar con múltiples herramientas para satisfacer las necesidades y va de la mano con la proyección tecnológica de migrar hacia entornos más orientados a software y virtualización como SDWAN y SDLAN en la que ejecución de programas o APIS con Python o JSON toman fuerza.

Se cumplen todos los requerimientos de visualización en la herramienta Qlik excepto la de visualizar que puerto tiene una extensión en particular debido a que no se contó con los permisos necesarios para obtener e incluir la información que permitiera realizar la visualización, este servicio es prestado parcialmente por otros proveedores y se prefirió no realizar para no mostrar información bastante incompleta. También para buscar la dirección MAC se hace la salvedad que puede que al momento de tomar los datos en los equipos Cisco un dispositivo no esté conectado y no pueda entonces capturar la MAC.

Los conocimientos adquiridos durante los cursos recibidos en la línea de inteligencia de negocios fueron determinantes para el desarrollo de este proyecto, se destacan los cursos Introducción de BI, Bodega de Datos y cubos, Gobernabilidad de los datos, Fundamentos base de datos, Indicadores, Visualización de información y Gerencia de proyectos de BI, comprobando el alto nivel académico del programa y su utilidad para la creación de soluciones de analítica.

7.2 Conclusiones del conocimiento adquirido con la ejecución del proyecto.

El análisis de calidad de datos arroja que se tienen un número muy alto de descripciones de puertos fuera del estándar por lo que es un punto importante para mejorar, desarrollar un código para corregirlo además de limitar al estándar lo que un operador puede ingresar herramientas como Cisco ISE o Cisco ACS ayudaría a que el error humano fuese menor y tener así un mejor indicador.

El anterior campo no afectó la información que se puede obtener tras el desarrollo de este proyecto por tratarse de un campo descriptivo. Los demás campos analizados no tienen un impacto importante, pero queda la responsabilidad para la operación de corregir con la ventaja de saber qué corregir y dónde.

Con este proyecto la compañía logra tener un punto único de información confiable para prestar un mejor servicio al poder dar respuesta más rápido a informes y reportes de operación, de incidentes y de factibilidad de nuevos servicios.

Este trabajo entrega como beneficio para la compañía los siguiente:

601 puertos de datos que no se están facturando.

29 teléfonos de asesor que no se están facturando.

55 teléfonos básicos que se están facturando demás al cliente.

20 puertos de *Access Points* que no se están facturando.

Si bien económicamente los 84 teléfonos dan una diferencia cercana a cero los otros 621 puertos a favor ofrecen un aumento importante en el ingreso mensual superior a 20 millones de pesos.

El indicador *Switches* recuperables se crea como aporte para la compañía por parte de este proyecto en conjunto con el gerente del proyecto y se le da visibilidad al interior. Su fórmula es $\text{Puertos disponibles}/(24*2)$, fórmula que sale al realizar un pequeño muestreo y concluir que se pudiese recuperar la mitad de los *Switches* que tienen puertos disponibles

8. TRABAJOS FUTUROS

Con la ejecución del proyecto se genera al interior de la compañía la semilla para futuros trabajos que potencializan lo realizado en este proyecto; por ejemplo, la inclusión de datos de tráfico operación de enlaces WAN extrayéndola de herramientas actuales. Integraciones con diferentes herramientas de monitoreo o de seguimiento de los servicios; Inclusión de inventario de otros clientes, más las bodegas con las que se cuenta en el país. Volver esta solución multi cliente y no solo dedicada para un único cliente es de por sí un trabajo retador que con seguridad agregaría mucho valor a la compañía.

Lo anterior implicaría también la implementación de un motor de base de datos más robusto que MS Access que permita darle la solidez necesaria al almacén de datos para ser un servicio confiable y escalable en el tiempo y un equipo de personas dedicado para esta labor.

El dar a conocer soluciones de analítica al interior de la compañía y de los clientes abre la posibilidad a futuros desarrollos y negocios, ya las preguntas van más allá de la analítica descriptiva y se empieza a explorar en el negocio de las telecomunicaciones a nivel operativo que beneficios nos puede traer la analítica predictiva, *machine learning*, entre otros.

9. REFERENCIAS

- Chaudhuri, S., Dayal, U., & Narasayya, V. (2011). An overview of business intelligence technology. *Communications of the ACM*, 54(8), 88. <https://doi.org/10.1145/1978542.1978562>
- Chen, M., Mao, S. W., & Liu, Y. H. (2014). Big Data: A Survey. *Mobile Networks & Applications*, 19(2), 171–209. <https://doi.org/10.1007/s11036-013-0489-0>
- Cisco Systems. (2007). Network Management System : Best Practices White Paper, 1–18.
- Completo, J., Cruz, R. S., Coheur, L., & Delgado, M. (2012). Design and Implementation of a Data Warehouse for Benchmarking in Clinical Rehabilitation. *Procedia Technology*, 5, 885–894. <https://doi.org/10.1016/j.protcy.2012.09.098>
- Dama. (2008). *The DAMA Dictionary of Data Management*.
- DAMA. (2013). The six primary dimensions for data quality assessment: defining data quality dimensions, 1–16.
- Davenport, T. H., & Harris, J. G. (2007). *Competing on Analytics: The New Science of Winning* (1st ed.). Boston, Mass: Harvard Business School Press.
- Diirr, T., & Santos, G. (2014). Improvement of IT service processes: a study of critical success factors. *Journal of Software Engineering Research and ...*, 2(1), 4. <https://doi.org/10.1186/2195-1721-2-4>
- Esteves, R., & Alves, P. (2013). Implementation of an Information Technology Infrastructure Library Process – the Resistance to Change. *Procedia Technology*, 9, 505–510. <https://doi.org/10.1016/j.protcy.2013.12.056>
- Few, S. (2014). Display Platforms for Quantitative Information. *Visual Business Intelligence Newsletter*, (October/November/December 2014), 1–7. Retrieved from

http://www.perceptualedge.com/articles/visual_business_intelligence/display_platforms_for_quantitative_information.pdf

- Financiera, S. (2007). Circular externa 0052 de 2007. Retrieved from https://www.superfinanciera.gov.co/descargas?com=institucional&name=pubFile21560&downloadname=ance052_07.doc
- Finch, G., Davidson, S., Kirschniak, C., Wickersheimer, M., Rodenbeck Reese, C., & Schockley, R. (2014). Analytics : The speed advantage. *Executive Report*, 23.
- Holsapple, C., Lee-Post, A., & Pakath, R. (2014). A unified foundation for business analytics. *Decision Support Systems*, 64, 130–141. <https://doi.org/10.1016/j.dss.2014.05.013>
- Inmon, W. H. (2011). A tale of two Architectures(2). *Orchids*, 69(8), 758–760. <https://doi.org/10.1042/BC20070109>
- Jones, M. E. M. E., & Song, I. Y. (2008). Dimensional modeling: Identification, classification, and evaluation of patterns. *Decision Support Systems*, 45(1), 59–76. <https://doi.org/10.1016/j.dss.2006.12.004>
- Juran, J. M., & Godfrey, A. B. (1998). *Juran's Quality Control Handbook*. McGrawHill. <https://doi.org/10.1108/09684879310045286>
- Karkouch, A., Mousannif, H., Al Moatassime, H., & Noel, T. (2016). Data quality in internet of things: A state-of-the-art survey. *Journal of Network and Computer Applications*, 73, 57–81. <https://doi.org/10.1016/j.jnca.2016.08.002>
- Khalilian, A. T., Ibrahim, O., & Nilashi, M. (2017). Integrated feedback control reporting for improving quality of technical service reporting in IT service management. *Telematics and Informatics*, 34(8), 1736–1771. <https://doi.org/10.1016/j.tele.2017.08.007>
- Kimball, R., & Ross, M. (2011). *The data warehouse toolkit: the complete guide to dimensional modelling*. Nachdr.]. New York [ua]: Wiley. <https://doi.org/10.1145/945721.945741>
- Koutsoukos, N. (2018). Riverbed a Gartner NPMD “Leader” and an APM “Challenger.” Retrieved September 17, 2018, from <https://www.riverbed.com/newsletter/riverbed-a->

gartner-npmd-leader-for-the-fifth-year-and-an-apm-cha.html

- Larson, D., & Chang, V. (2016). A review and future direction of agile, business intelligence, analytics and data science. *International Journal of Information Management*, 36(5), 700–710. <https://doi.org/10.1016/j.ijinfomgt.2016.04.013>
- Lee, M. L., Hsu, W., & Kothari, V. (2004). Cleaning the Spurious Links in Data. *IEEE Intelligent Systems*, 19(2), 28–33. <https://doi.org/10.1109/MIS.2004.1274908>
- Li, S. T., Shue, L. Y., & Lee, S. F. (2008). Business intelligence approach to supporting strategy-making of ISP service management. *Expert Systems with Applications*, 35(3), 739–754. <https://doi.org/10.1016/j.eswa.2007.07.049>
- Liu, D. (2009). SSH Features. In I. Elsevier (Ed.), *Next Generation SSH2 Implementation* (Vol. 6, pp. 65–86). Burlington, MA 01803: Syngress Publishing, Inc.
- Merino, J., Caballero, I., Rivas, B., Serrano, M., & Piattini, M. (2016). A Data Quality in Use model for Big Data. *Future Generation Computer Systems*, 63, 123–130. <https://doi.org/10.1016/j.future.2015.11.024>
- Mezzanzanica, M., Boselli, R., Cesarini, M., & Mercorio, F. (2015). A model-based evaluation of data quality activities in KDD. *Information Processing and Management*, 51(2), 144–166. <https://doi.org/10.1016/j.ipm.2014.07.007>
- Mosley, M., Brackett, M., Earley, S., & Henderson, D. (2009). *The DAMA Guide to The Data Management Body of Knowledge*. Technics Publications, LLC Post.
- Paul Lomax, Matt Childs, R. P. (2003). *VBS Script in a nutshell*. (R. Petrusa, Ed.) (2nd ed.). O'Reilly.
- PRTG - Price List. (n.d.). Retrieved from <https://www.es.paessler.com/prtg/pricing>
- R. Kalakota. (2011). Gartner says – BI and Analytics a \$12.2 Bln market. Retrieved from <https://practicalanalytics.co/2011/04/24/gartner-says-bi-and-analytics-a-10-5-bln-market/>
- Rai, A., Dubey, V., Chaturvedi, K. K., & Malhotra, P. K. (2008). Design and development of data mart for animal resources. *Computers and Electronics in Agriculture*, 64(2), 111–119.

<https://doi.org/10.1016/j.compag.2008.04.009>

- Rikhardsson, P., & Yigitbasioglu, O. (2018). Business intelligence & analytics in management accounting research: Status and future focus. *International Journal of Accounting Information Systems*, 29(February), 37–58. <https://doi.org/10.1016/j.accinf.2018.03.001>
- Rivadera, G. (2010). La metodología de Kimball para el diseño de almacenes de datos (Data warehouses). *Cuadernos de La Facultad, Universidad Catolica de Salta*, (5), 56–71. Retrieved from <http://www.ucasal.edu.ar/htm/ingenieria/cuadernos/archivos/5-p56-rivadera-formateado.pdf>
- Rud, O., Parr, O., Rud, O., Parr, O., Rud, O., Parr, O., & Rud, O. (2009). *Business Intelligence Success Factors. Business* (Vol. 1). <https://doi.org/10.1109/MSR.2009.5069473>
- S. Pressman, R. (2010). Ingeniería del software: Un enfoque práctico. In *Ingeniería del software: Un enfoque práctico*.
- Stanton, J. (2012). An Introduction to Data Science. *Syracuse University*, 1–157. https://doi.org/10.1007/978-3-319-50017-1_1
- Sulaiman, N. S., & Yahaya, J. H. (2013). Development of Dashboard Visualization for Cardiovascular Disease based on Star Scheme. *Procedia Technology*, 11(Iceei), 455–462. <https://doi.org/10.1016/j.protcy.2013.12.215>
- Sullivan, B. (2015). 2012 Bower Award for Business Leadership presented to John T. Chambers, Cisco Systems, Inc. | San Jose, California. *Journal of the Franklin Institute*, 352(7), 2585–2586. <https://doi.org/10.1016/j.jfranklin.2015.02.016>
- Terplan, K. (1992). *Communication networks management*. (P. Hall, Ed.) (2nd ed.). Englewood Cliffs, N.J.
- Tokola, H., Gröger, C., Järvenpää, E., & Niemi, E. (2016). Designing Manufacturing Dashboards on the Basis of a Key Performance Indicator Survey. *Procedia CIRP*, 57, 619–624. <https://doi.org/10.1016/j.procir.2016.11.107>
- Yamada, H., Yada, T., & Nomura, H. (2011). Developing network configuration management

database system and its application—data federation for network management. *Telecommunication Systems*, 993–1000. <https://doi.org/10.1007/s11235-011-9607-0>

ANEXO 1

PREGUNTAS REALIZADAS – REQUERIMIENTOS DE NEGOCIO

- ¿Actualmente cuál es el mayor inconveniente al momento de realizar la facturación?

No se tiene precisión en la cantidad de puertos, no se tiene un sistema que nos permita contar fácilmente para todas las sedes los puertos activos discriminados según su tipo.

- ¿Cuántos tipos de puerto existen, y cuál es su impacto en la facturación?

Los puertos se facturan dependiendo de su uso, si es un teléfono, un *Access point* o un puerto de usuario se cobra diferente, si es un enlace o un troncal entre equipos nuestros no se cobra, también depende de tipo de sede varía su costo.

- Cómo ejercen el control actualmente y se llevan el conteo de puertos

Se tiene una base del área comercial y de facturación y cuando llegan una orden de servicio se traduce en una orden de cambio y una vez ejecutada por correo electrónico informamos para que se lleve el control, es un proceso manual.

- ¿La operación lleva un control de los puertos, o solo informa de los cambios?

Solo informa de los cambios al área encargada es su responsabilidad actualizarla y llevarlo a la conciliación.

- ¿La operación lleva un registro de los cambios realizados?

Si, en una tabla de Excel, con su respectiva orden de cambio generada por el cliente.

- ¿Qué tan actualizada está la información de inventario de puertos y que porcentaje está incluida en las herramientas de gestión como IM Spectrum o Cisco Prime?

No se tiene CMBD y actualmente llevamos en Excel hojas de vida de los dispositivos, cada sede tiene un Excel con toda la información, pero no es información confiable prefiero que se obtenga de los equipos como indicas que se puede hacer con un script.

- ¿Quién configura cambios en los puertos?

La gestión de solicitudes ejecuta el proceso a cargo de dos personas. Pueden existir casos de soporte en los que se muevan puertos de ubicación o incluso daños de equipos que genere un cambio importante.

- ¿Existe un estándar de marcación de puertos que me permita determinar que tipo de puerto es?

Si pero al ser texto descriptivo puede no ser confiable.

- ¿Todos los *Access Points* o teléfonos cuestan lo mismo sin importar el modelo?

Los *Access Points* cuestan lo mismo, los teléfonos según el tipo se hace un cobro diferente.

- ¿Los enlaces de comunicaciones se cobran?

Los enlaces provistos por nosotros no se cobran como puerto adicional, pero los del enlace de respaldo si se factura como si fuese un enlace de datos.

- ¿Llevan estadísticas de cambios en los puertos por órdenes de cambio, daño o incidente?

No y nos gustaría que así fuese.

- ¿Dónde encuentro la información más actualizada de inventario de puertos y de dispositivos instalados??

En las hojas de vida de los equipos, también en un reporte de IM Spectrum, un informe del Prime hasta donde permita la licencia, hay que buscar la información en varias partes para generar reportes, usualmente toma tiempo. Para facturación se puede tomar como base el archivo de conciliación con el cliente.

- ¿Spectrum tiene una CMBD?

No, es solamente monitoreo.

- ¿Se tiene alguna otra herramienta como CMBD?

No. Está en proceso de construcción y la información de este proyecto de BI podríamos aprovecharla para su construcción.

- ¿Se tiene distribución geográfica de los equipos instalados describiendo modelo y cantidad?

No. Sería muy útil para la planificación de inventario.

También se realizan preguntas tomadas del cuestionario propuesto por el Ingeniero Javier Bermudez aprendidas durante el curso de Bodega de Datos.

- ¿Cómo analizas normalmente tu gestión?

Manualmente, en las conciliaciones y mirando sabanas de datos con cambios e incidentes

- ¿Qué data usas?

Ordenes de cambio, sedes, alta o baja e puertos, equipos reemplazados y sitios nuevos

- ¿Cómo la obtienes?

Se encuentra en la carpeta compartida o por correo,

- ¿Qué reportes utilizas normalmente?

Ninguno, todo es manual

- ¿Qué capacidades analíticas te gustaría tener?

Distribución geográfica, sedes de más valor, sedes que hayan crecido, puertos sin usar por más de 5 semanas.

- ¿Hay cuellos de botella para obtener la información?

Si, todo es manual y a veces hay que poner el grupo de ingenieros para reportes específicos.

- ¿Cuánta información histórica es requerida?

Ninguna

- ¿Qué oportunidades de mejora existen en lo referente al análisis de información para toma de decisiones?

Revisar distribución geográfica, *switches* que podríamos optimizar por tener muchos puertos disponibles, apagar puertos con mucho tiempo inactivos.