

TRABAJO FIN DE GRADO

Grado en Ingeniería Informática

UNIVERSIDAD DE ALMERIA

ESCUELA SUPERIOR DE INGENIERÍA

Inteligencia
empresarial
digitalizada en PYME

Curso 2020/2021

Alumno:

Santiago Medina Martínez

Director:

Manuel Torres Gil





Inteligencia empresarial digitalizada en PYME

Curso 2020/2021

Santiago Medina Martínez

Índice

Índice	4
Índice de figuras	6
Índice de tablas	8
Capítulo 1	10
1. Introducción	10
1.1 Objetivos	10
1.2 Infraestructura IT	11
1.3 Planificación	12
Capítulo 2	14
2.1 Creación de una infraestructura IT	14
2.1.1 Elección de herramientas y construcción del esquema	14
2.2 Metodologías y tecnologías utilizadas	16
2.2.1 Hyper-V	16
2.2.2 Active Directory	16
2.2.3 Configuración de redes y NAT	17
2.2.4 VPN	19
2.2.5 Tecnologías similares	21
2.3 Especificación de la configuración	21
2.4 Detalles de configuración	23
2.5 Resumen del capítulo	37
Capítulo 3	38
3.1 DataWarehouse	38
3.1.1 Herramientas y tecnologías utilizadas	38
3.1.2 Datos y esquema	40
3.2 Especificación de la configuración	42
3.3 Detalles de configuración	43
3.4 Resumen del capítulo	53
Capítulo 4	54
4.1 Business Intelligence	54
4.2 Power BI	55
4.3 Especificación de la configuración	56

4.4 Detalles de la implementación	60
4.4.1 Acceso a los recursos de la infraestructura	65
4.5 Resumen del capítulo	70
5. Conclusiones y trabajo futuro	72
5.1 Conclusiones	72
5.2 Trabajo futuro	72
6. Bibliografía	74
Anexo	76
Conexión VPN	76
Scripts para modelado de datos	79
Informes Power BI	85

Índice de figuras

Figura 1.1: Esquema infraestructura IT	12
Figura 2.1: Ordenadores de cada departamento	14
Figura 2.2: Controlador de Dominio y ordenadores unidos al dominio	15
Figura 2.3: Construcción de la infraestructura	15
Figura 2.4: Infraestructura y objetos del dominio	17
Figura 2.5: Configuración NAT y direcciones IP de todos los ordenadores y servidores	18
Figura 2.6: Infraestructura con VPN	21
Figura 2.7: Rol de Servicios de dominio de AD	24
Figura 2.8: Creación del dominio 'TFG.ES'	24
Figura 2.9: Script para crear DC y dominio	24
Figura 2.10: Comprobación DC pertenece al dominio y es controlador de dominio	25
Figura 2.11: Configuración del adaptador de red en el DC	25
Figura 2.12: Unión del SW(NAT) al dominio	26
Figura 2.13: Estructura OU	26
Figura 2.14: Creación de usuarios	27
Figura 2.15: Creación de grupos de usuarios	27
Figura 2.16: Derechos de acceso a usuarios del Departamento TI al DC y ServidorWeb	28
Figura 2.17: GPO para prohibir el acceso al PC	28
Figura 2.18: Establecer horas de inicio de sesión para un usuario	29
Figura 2.19: Instalación del rol RDS	29
Figura 2.20: Instalación servidor de enrutamiento y acceso remoto	30
Figura 2.21: Creación interfaz NAT	30
Figura 2.22: Designación de la dirección del servidor	31
Figura 2.23: Pruebas de red interna y externa	32
Figura 2.24: Reserva de direcciones DHCP para VPN	32
Figura 2.25: Firewall, reglas de entrada	33
Figura 2.26: Firewall, reglas de salida	33
Figura 2.27: Configuración VPN L2TP en el Servidor	34
Figura 2.28: Configuración VPN L2TP en un PC	34
Figura 2.29: Importación de certificado en el servidor	35
Figura 2.30: Exportación del certificado con clave	35
Figura 2.31: Configuración VPN IKEv2 en PC	36
Figura 2.32: Comprobación de conexión VPN IKEv2	36
Figura 3.1: Tablas iniciales de Noaa_Data	41
Figura 3.2: Diagrama en estrella	41
Figura 3.3: Conexión con la BD Noaa_data	44
Figura 3.4: Vista origen de datos	44
Figura 3.5: Selección de la tabla para crear la dimensión fecha	45
Figura 3.6: Selección de los atributos y su tipo en Dimensión Fecha	45
Figura 3.7: Selección atributos de la dimensión lugar	46
Figura 3.8: Jerarquía de fecha	46
Figura 3.9: Eliminar errores duplicidad de claves al procesar la jerarquía	47
Figura 3.10: jerarquía lugar	48
Figura 3.11: Procesado de la dimensión lugar	48
Figura 3.12: Exploración de dimensión lugar	49
Figura 3.13: Asistente para elegir las medidas para la creación del cubo	49
Figura 3.14: Finalización del asistente para crear el cubo	50
Figura 3.15: Implementación del proyecto	50

Figura 3.16: Procesamiento del cubo	51
Figura 3.17: Creación de una medida (Máximo de Prcp)	51
Figura 3.18: Creación columna calculada VelViento	52
Figura 3.19: Consulta MDX para ver columna calculada	52
Figura 3.20: Resumen de columnas calculadas	53
Figura 3.21: Diseño de agregaciones	53
Figura 4.1: Conexión al origen de datos	60
Figura 4.2: Reemplazar '*' por 'o'	61
Figura 4.3: Crear columna elevación en metros	61
Figura 4.4: Modelo de datos del origen BD relacional tras las tareas de limpieza y transformación	61
Figura 4.5: Combinar dos columnas	62
Figura 4.6: Mapa con matriz	62
Figura 4.7: Dashboards de Tª mínima	63
Figura 4.8: Coloreado de datos	63
Figura 4.9: Dashboard de vientos	64
Figura 4.10: Top N estaciones por lluvia	65
Figura 4.11: Compartir carpetas compartidas con los grupos de usuarios	65
Figura 4.12: Asignación de permisos de carpetas por departamento y personales	66
Figura 4.13: Directorio de carpetas compartidas creado	66
Figura 4.14: Vista del directorio compartido a través del usuario Luis	67
Figura 4.15: Directorio antes de compartir el informe y permisos que se aplican	67
Figura 4.16: Informe compartido correctamente	68
Figura 4.17: Compartir un proyecto en Power BI Service	68
Figura 4.18: Crear un área de trabajo	69
Figura 4.19: Vista de informe desde Power BI Mobile	69
Figura 4.20: Código HTML del informe	69
Figura 4.21: Informe publicado en Blogger	70

Índice de tablas

Tabla 1.1: Planificación temporal	13
Tabla 2.1: Especificación VM en Hyper-V	21
Tabla 2.2: Configuración direcciones IPV4	22
Tabla 2.3: Configuración direcciones IPV6	22
Tabla 2.4: Estructura de OU	22
Tabla 2.5: Usuarios del dominio	23
Tabla 2.6: GPO's aplicadas	23
Tabla 2.7: Configuración puertos del router para VPN	23
Tabla 3.1: Descripción de los datos de la tabla de hechos	42
Tabla 3.2: Descripción de los datos de la Dimensión_Fecha	42
Tabla 3.3: Descripción de los datos de la Dimensión_Lugar	43
Tabla 3.4: Descripción de las consultas con nombre creadas en Dimensión_Lugar	43
Tabla 3.5: Jerarquía de la Dimensión Lugar	43
Tabla 3.5: Jerarquía de la Dimensión Fecha	43
Tabla 3.6: Columnas calculadas con expresiones MDX en la tabla de hechos	43
Tabla 4.1: Plantillas para informes de Power BI	59
Tabla 4.2: Iconos utilizados en los informes	60
Tabla 4.3: Color de datos en el gráfico de dispersión de vientes	61
Tabla 4.4: Estructura del directorio y subdirectorios compartido creado en DC	61



Capítulo 1

1. Introducción

Actualmente, la inteligencia empresarial es muy útil para apoyar la toma de decisiones en las empresas, proporcionando una vista completa tanto de los datos antiguos como actuales de la organización.

Por esta razón, se decide crear un proyecto para implementar Business Intelligence en una empresa ficticia, cuya actividad principal es la toma y estudio de mediciones meteorológicas en EEUU.

Para ello se creará un DataWarehouse con datos meteorológicos y se realizarán informes usando la herramienta Power BI que consumirá los datos de un cubo OLAP y de una base de datos relacional creada previamente para desarrollar el DataWarehouse, de esta manera se diversificará el uso de Power BI con dos tipos de orígenes de datos, el primero que no permite la transformación y el segundo que permite la limpieza y transformación de datos.

Para lograr que los empleados puedan trabajar y realizar sus funciones en la organización analizando o creando informes meteorológicos respetando las medidas de seguridad frente al COVID, se necesita crear una infraestructura IT que ofrezca la posibilidad de trabajar desde casa o desde la propia oficina.

Se debe garantizar la confidencialidad e integridad del Sistema de Información creado, para conseguirlo se crea una infraestructura IT cuya seguridad es proporcionada por Active Directory, que permite restringir los recursos, en que periodos de tiempo son accesibles por los distintos empleados, y que además permite el teletrabajo a través de una VPN.

1.1 Objetivos

Los objetivos que se definen para la consecución del proyecto son:

- Creación de un Data Warehouse de medidas meteorológicas usando SQL. [1] [2]
 - Definir un esquema en estrella.
 - Crear dimensiones y jerarquías.
 - Crear columnas calculadas.
 - Procesamiento del cubo.
- Aprender a usar la herramienta para inteligencia empresarial Power BI [3] [4]:
 - Usar múltiples orígenes de datos.
 - Realizar tareas de limpieza y transformación de datos.
 - Crear informes interactivos usando filtros, los distintos elementos visuales y visualizaciones ofrecidas.
 - Creación de columnas, medidas y funciones DAX. [5]

- Publicación de informes.
- Crear una infraestructura con Active Directory que permita teletrabajar y garantizar la confidencialidad del Sistema de Información [6] [7]:
 - Creación de un dominio, usuarios, unidades organizativas y directivas de grupo.
 - Creación de un directorio de carpetas para permitir compartir recursos entre los empleados de la empresa.
 - Diseñar y configurar una red en la que un servidor con NAT permita el acceso a Internet a los dispositivos de la red interna y el intercambio de paquetes autorizados bidireccionalmente entre la red interna y externa. [8] [9]
 - Crear una VPN L2TP/IPSEC y mejora de esta a IKEv2. [10]

1.2 Infraestructura IT

Frente a la necesidad de una infraestructura que simule los distintos departamentos de la empresa ficticia y albergue el Sistema de Información descrito anteriormente, se escoge la posibilidad de crearla usando el programa de virtualización de Microsoft, Hyper-V. [11]

Esta infraestructura se compone de 5 máquinas virtuales, dos con el Sistema Operativo Windows Server 2019 y las tres restantes equipadas con Windows 10 Profesional que actuarán como PC para cada uno de los departamentos existentes: 'TI', 'Gerentes' y 'Analistas'.

En el PC de Departamento TI será donde se cree el Sistema de Información, es decir, donde se crean el DataWarehouse y Dashboards, que posteriormente se comparten a los demás puestos mediante los directorios de carpetas compartidas o el repositorio de Power BI Service. Para poder llevar a cabo estos trabajos se equipará el PC con:

- **SQL Server Management Studio 18** que se utilizará para crear y administrar la base de datos.
- **SQL Server Data Tools** con esta herramienta se crearán los modelos de Analysis Services.
- **SQL Server 2019 Developer**, donde se activarán las características de 'Servicios de Motor de base de datos' y 'Analysis Services'.
- **Power BI** que procesa la información del cubo y la BD relacional convirtiéndola en atractivos Dashboard y gráficos con fácil interpretación para apoyar la toma de decisiones.

Los PCs de los otros dos departamentos son equipados solo con Power BI y Adobe Reader para poder visualizar los informes.

Para centralizar la administración de la infraestructura se usa Active Directory, se crea un dominio y un controlador de dominio que permita gestionar la autenticación de usuarios y permisos, de los usuarios y máquinas pertenecientes al dominio. Para esto se usa una de las máquinas con WS 2019 y se nombra como 'DC'.

La máquina restante nombrada como 'Servidor Web', actuará de **NAT** (Network Address Translation) entre la red externa e interna, la cual tendrá configurada la VPN, también actuará como Firewall.

En la siguiente figura se puede observar un esquema general de la infraestructura creada:

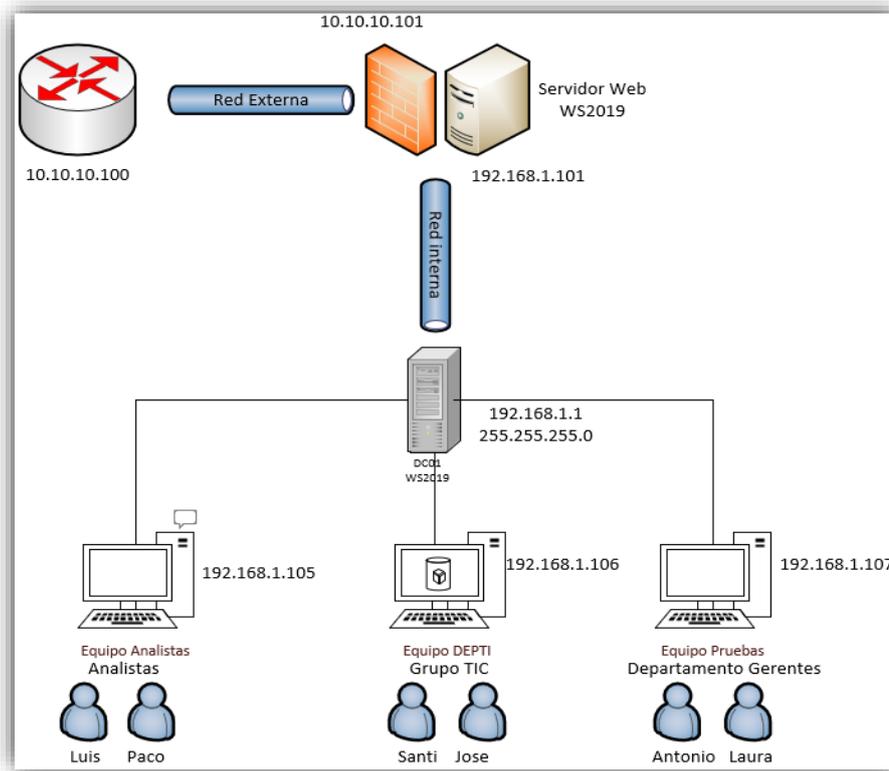


Figura 1.1: Esquema infraestructura IT

1.3 Planificación

Para el desarrollo de este proyecto se han desarrollado las siguientes fases:

- **Primera fase**, consiste en la construcción del esquema de la infraestructura IT, asignando las direcciones de red. Búsqueda de un conjunto de datos para crear el DW.
- **Segunda fase** tiene el propósito de construir y configurar la infraestructura IT con Hyper-V, incluyendo la creación del dominio, unión de los puestos al dominio, creación de los usuarios, grupos de usuarios y directivas de grupo. Instalando también el software de cada PC. Para finalizar esta fase se creará la VPN.
- **Tercera fase** transforma el conjunto de datos encontrados durante la primera fase, construyendo una BD relacional con esquema en estrella y se creará el cubo OLAP. Por último, se realizan las tareas de limpieza y transformación de datos con Power BI de los datos obtenidos de la BD relacional y se crean y publican informes con ambos orígenes de datos.
- **Cuarta fase** pruebas finales para probar el correcto funcionamiento del proyecto realizado y elaboración de la memoria.

A continuación, se mostrará una planificación temporal (Tabla 1.1) y un diagrama de Gantt (Figura 1.2), con una planificación más detallada.

Cada día de trabajo corresponde a 4 horas.

El tiempo de trabajo total para la realización del TFG es de 300 horas aproximadamente.

Fase	Actividad	Duración (días)	Fecha de inicio	Fecha de fin
1	Búsqueda de idea para desarrollar el TFG	3	10/02/2021	12/02/2021
1	Estudio del problema	2	15/02/2021	16/02/2021
1	Realización del esquema empresarial	2	17/02/2021	19/02/2021
1	Búsqueda de un conjunto de datos	2	20/02/2021	21/02/2021
2	Creación de máquinas virtuales y configuración	4	22/02/2021	25/02/2021
2	Creación del dominio y Domain Controller	2	01/03/2021	02/03/2021
2	Unión de todos los dispositivos al dominio	1	03/03/2021	03/03/2021
2	Creación de usuarios y grupos	1	04/03/2021	04/03/2021
2	Creación de directivas y permisos	3	05/03/2021	07/03/2021
2	Instalación de software	2	08/03/2021	09/03/2021
2	Creación de carpetas compartidas	2	10/03/2021	11/03/2021
2	Creación e implementación VPN	5	15/03/2021	19/03/2021
3	Limpieza y transformación de datos	7	22/03/2021	30/03/2021
3	Creación del DataWarehouse	7	31/03/2021	07/04/2021
3	Creación de funciones DAX, informes y gráficas	14	08/04/2021	23/04/2021
3	Publicación de informes	2	26/04/2021	27/04/2021
3	Creación de servicios web y roles de usuario	2	28/04/2021	29/04/2021
4	Pruebas finales	3	31/04/2021	02/05/2021
4	Realización del informe	18	03/05/2021	26/05/2021

Tabla 1.1: Planificación temporal

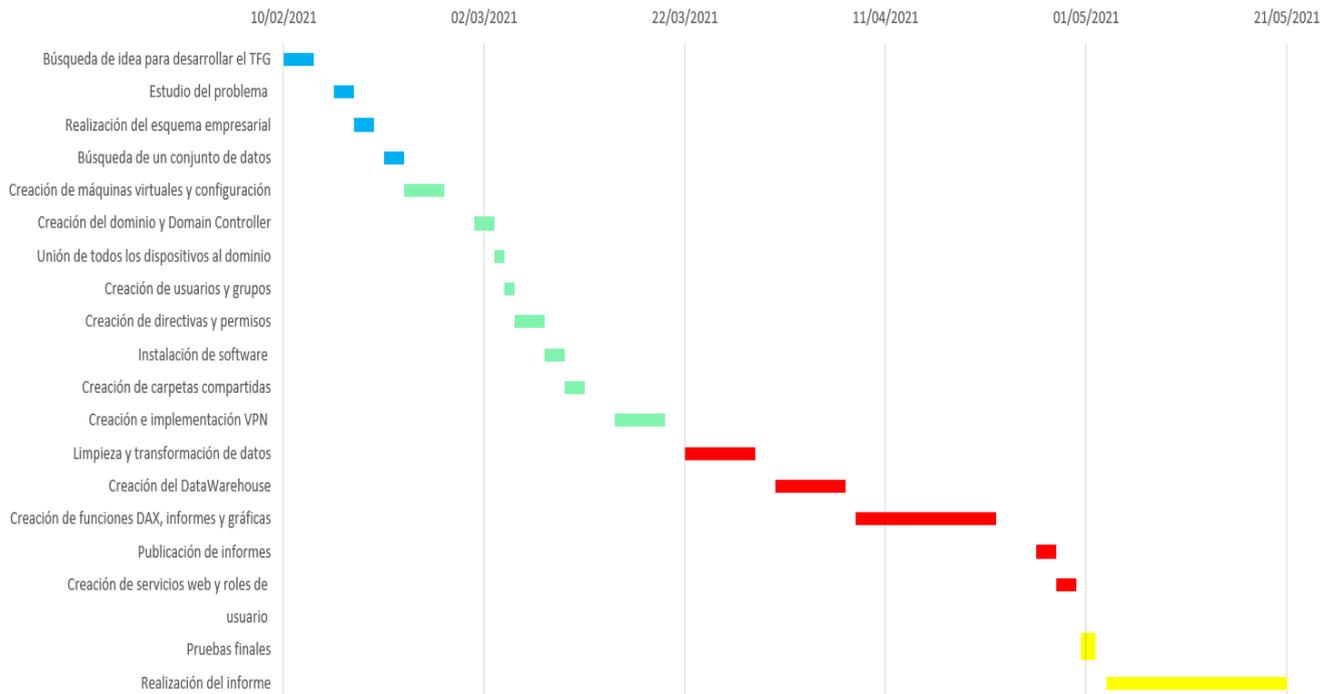


Figura 1.2: Diagrama de Gantt

Capítulo 2

En este capítulo, se creará la infraestructura IT necesaria para simular el entorno empresarial donde se implementará el Sistema de Información de manera que se garantice su seguridad, y permita que los empleados accedan a un PC y los recursos necesarios para desempeñar su trabajo. Además, como se mencionó anteriormente ante la gravedad de la pandemia deben tener la posibilidad de teletrabajar desde casa a través de una VPN.

A continuación, se procede a describir los materiales usados, así como todo el proceso que se realiza en detalle y por qué se realiza.

Por último, durante este capítulo definiremos los conceptos, especificaciones y detalles de la configuración realizada.

2.1 Creación de una infraestructura IT

En esta sección se describe la infraestructura y las herramientas necesarias para su implementación.

2.1.1 Elección de herramientas y construcción del esquema

Una vez conocida la actividad principal de la empresa ficticia, compuesta por tres departamentos y la necesidad de un computador en cada uno de ellos, se procede a elegir las herramientas software necesarias para implementarla.

En primer lugar, ante la necesidad de tener varios ordenadores, se usará el programa de virtualización Hyper-V, a través del cual podremos crear fácilmente máquinas virtuales resolviendo el problema inicial. La estructura queda de la siguiente forma:



Figura 2.1: Ordenadores de cada departamento

Para poder administrar estos ordenadores se decide usar Active Directory con el que se creará un dominio que será utilizado para la administración de usuarios, grupos y computadoras pertenecientes a él. Esto se consigue mediante otra máquina virtual equipada con WS 2019 y que nombramos DC y actuará como un servidor con la función de controlador de dominio.

Tras esto conseguimos la siguiente infraestructura (Figura 2.2):

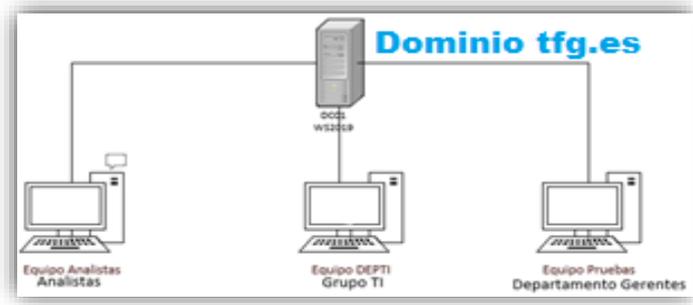


Figura 2.2: Controlador de Dominio y ordenadores unidos al dominio

Una vez creado el dominio y teniendo unidos todos los computadores se necesita realizar la configuración red, donde se encuentra la necesidad de crear otra máquina virtual que actúe como NAT entre la red interna de la infraestructura y la red externa. Además de tener configurada la VPN para permitir el acceso a los recursos y máquinas desde el exterior.

La figura 2.3 representa la infraestructura de las tareas anteriormente descritas.

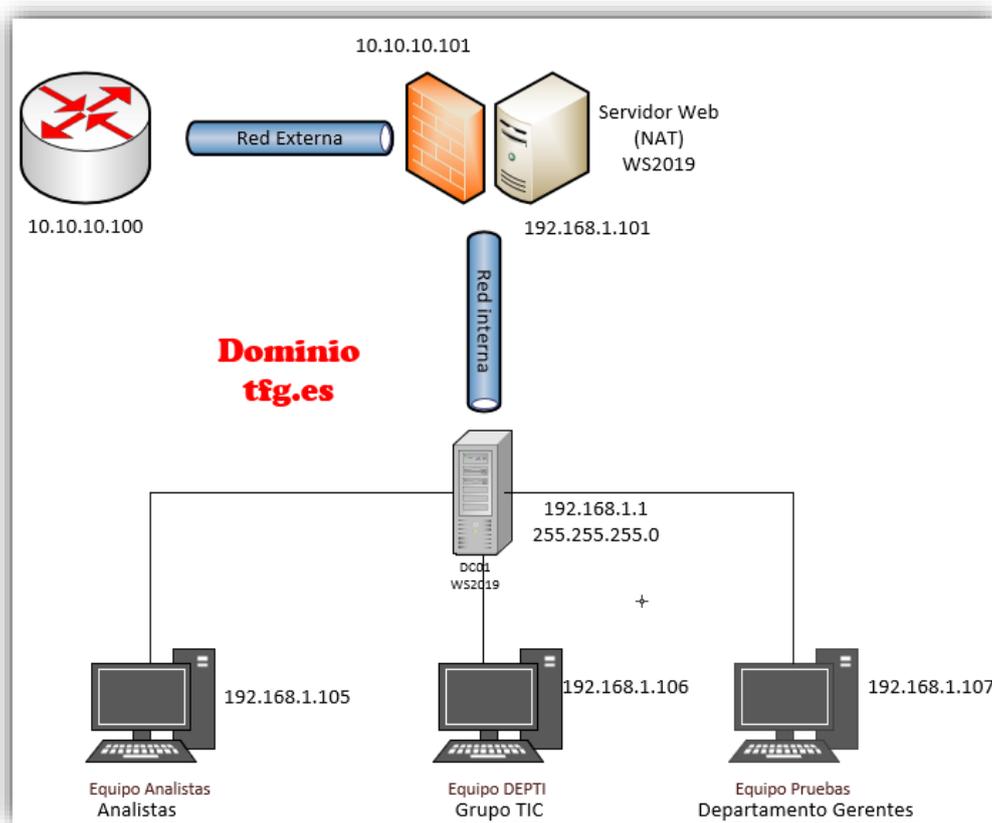


Figura 2.3: Construcción de la infraestructura

Para finalizar, es necesario crear a los empleados de cada departamento y unirlos al grupo de usuarios correspondiente, así como limitar las acciones y recursos a su disposición según el cargo y departamento a los que pertenezcan. Esto se conseguirá hacer gracias a Active Directory como veremos en el apartado [2.4 Detalles de configuración](#).

2.2 Metodologías y tecnologías utilizadas

En esta sección, se explicarán las tecnologías mencionadas previamente y la descripción de los conceptos más relevantes acerca de estas.

2.2.1 Hyper-V

Como se ha introducido previamente de manera breve, Hyper-V es la aplicación de virtualización desarrollada por Microsoft, dicha herramienta está disponible en todos los ordenadores con Sistema Operativo Windows 10 (versión profesional o superior), y sirve para crear y ejecutar máquinas virtuales.

Dentro de la aplicación tenemos el concepto de **conmutador virtual**, que sirve para crear un adaptador de red en la máquina virtual, puede ser de dos tipos:

- Conmutador virtual de red externa: Se usa un adaptador de red físico para dar a la máquina virtual acceso a una red física.
- Conmutador virtual de tipo red interna: Este adaptador solo puede ser usado por las máquinas que se ejecutan en el mismo equipo físico entre él y las máquinas virtuales del mismo.

Los **puntos de control** sirven para guardar un snapshot de la máquina virtual en el momento en el que se realiza, gracias a ellos podemos restaurar una versión antigua de la máquina si surge algún problema. [11]

2.2.2 Active Directory

Es una tecnología que proporciona un servicio ubicado en uno o varios servidores, que administra una base de datos y guarda todos los recursos de la red (usuarios, grupos, unidades organizativas, contraseñas, equipos, impresoras, etc.). También permite administrar todos estos recursos de forma centralizada. Una de las funciones más importantes de Active Directory es administrar las credenciales durante el inicio de sesión de los usuarios, para asignarles un token que será el que le permitirá entrar en los distintos recursos compartidos, usar o no diferentes equipos y dispositivos (impresoras, escáneres...) dentro de la red. [12] [13]

Las **ventajas** de usar AD (Active Directory) son:

- Tiene una gran escalabilidad, puede ser usado desde una pequeña empresa a grandes empresas con varios servidores que dividen la carga. A medida que la empresa crece es muy fácil añadirle servidores y puestos de trabajo.
- Permite administrar y autenticar cuentas de usuario.
- Permite restringir determinadas acciones.
- Consiente el uso de permisos para crear carpetas compartidas en la red de la organización.

Tras una breve descripción de que es AD, se va a hablar sobre una serie de conceptos importantes relacionados con AD que se usan en este proyecto.

El **controlador de dominio**, a partir de ahora **DC**, será el servidor que se encarga de controlar y administrar el acceso a los objetos que forman parte del dominio, almacena la información de la base de

datos del directorio de AD, y administra la seguridad de los objetos como usuarios, grupos de usuarios y dispositivos. Para ello, se debe instalar en este servidor el rol de Servicios y equipos de Active Directory.

Muy relacionado con el anterior se encuentra el concepto de **dominio**, que es un conjunto de ordenadores administrados por un servidor que sea controlador de dominio, que administrará los usuarios, credenciales y objetos del dominio. Algunos objetos que se usan en este proyecto pertenecientes a un dominio son:

- **Unidades organizativas**, son contenedores que establecen jerarquías dentro de un dominio, y sirven para organizar el directorio. Las OU pueden contener otras unidades organizativas, equipos, usuarios, carpetas compartidas, grupos...
Ayudando a aplicar permisos y organizar la estructura de forma eficiente.
- **Equipos** serán los ordenadores de usuarios, que se unan al dominio creado.
- **Controladores de Dominio**, servidor/es con el rol de ‘Servicios de dominio de Active Directory’
- **Usuarios y grupos de usuarios**, las cuentas de usuario sirven para que una entidad, pueda tener acceso o iniciar sesión en un dominio administrado con Windows Server. Por otro lado, los grupos de usuarios son contenedores que contienen usuarios, permitiendo una mejor organización de los mismos y facilita su administración a la hora de aplicar directivas.
- **Directivas de grupo (GPO)**, restringen que un usuario, grupos de usuarios o equipos, tengan acceso a determinadas acciones. Es una de las herramientas más útiles ofrecidas por AD, permiten gestionar los recursos del dominio utilizados por usuarios, grupos o equipos del mismo de manera eficiente. [14]

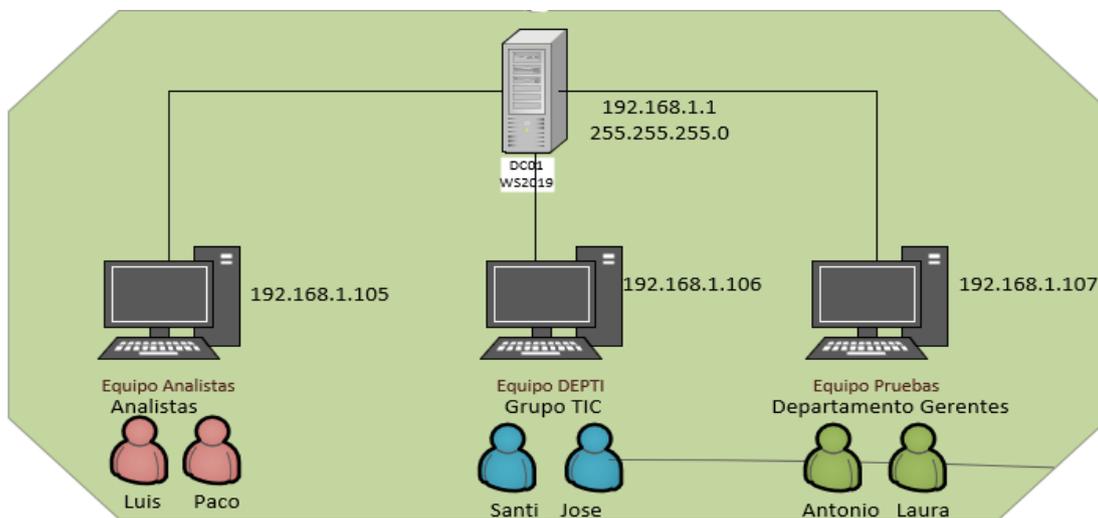


Figura 2.4: Infraestructura y objetos del dominio

2.2.3 Configuración de redes y NAT

Una vez descrito la sección de Active Directory se darán a conocer los conceptos más importantes acerca de la configuración de redes y del NAT.

Ante la necesidad de permitir acceso a los ordenadores del dominio se opta por crear un servidor que se encargue de este trabajo. Para ello, se necesitan dos adaptadores de red uno que conecte a la red interna de la organización y otro externo que conecte a la red física para proveer a la máquina virtual

de conexión a internet. Además, es necesario instalar el rol de Enrutamiento que proporcionará compatibilidad con enrutadores NAT, LAN y Multidifusión. Una vez configurado el NAT se intercambiarán los paquetes entre la interfaz externa y la interna. Los ordenadores a los que se quiere proporcionar acceso a internet se configuraran con direcciones IP estáticas asignándole como puerta de enlace la dirección IP de la red interna del servidor, que actúa como NAT. [9]

A continuación, se definen algunos conceptos importantes:

- **Enrutamiento:** permite la entrada y salida de paquetes entre la red interna y externa de la empresa; hace que redes de ordenadores que utilizan direcciones IP privadas, se conecten a Internet usando una única dirección IP pública. En cuanto a la seguridad, las máquinas conectadas a la red mediante NAT no son visibles desde el exterior, por lo que un atacante externo no podría averiguar si una máquina está conectada o no a la red. [8]
- **Panel Enrutamiento y acceso remoto:** permite configurar el NAT y acceso VPN.
- **DNS** o Sistema de nombres de dominio, sirve para simplificar el acceso a otros dispositivos IP, asignando nombres con los que resolver la dirección. También permite la autenticación de usuarios y dispositivos en el dominio. [15]
- **Puerta de enlace** su función es enviar los paquetes del dispositivo que no se conozca por qué interfaz debe enviarse por una ruta determinada, en este caso de uso la interfaz interna del Servidor Web

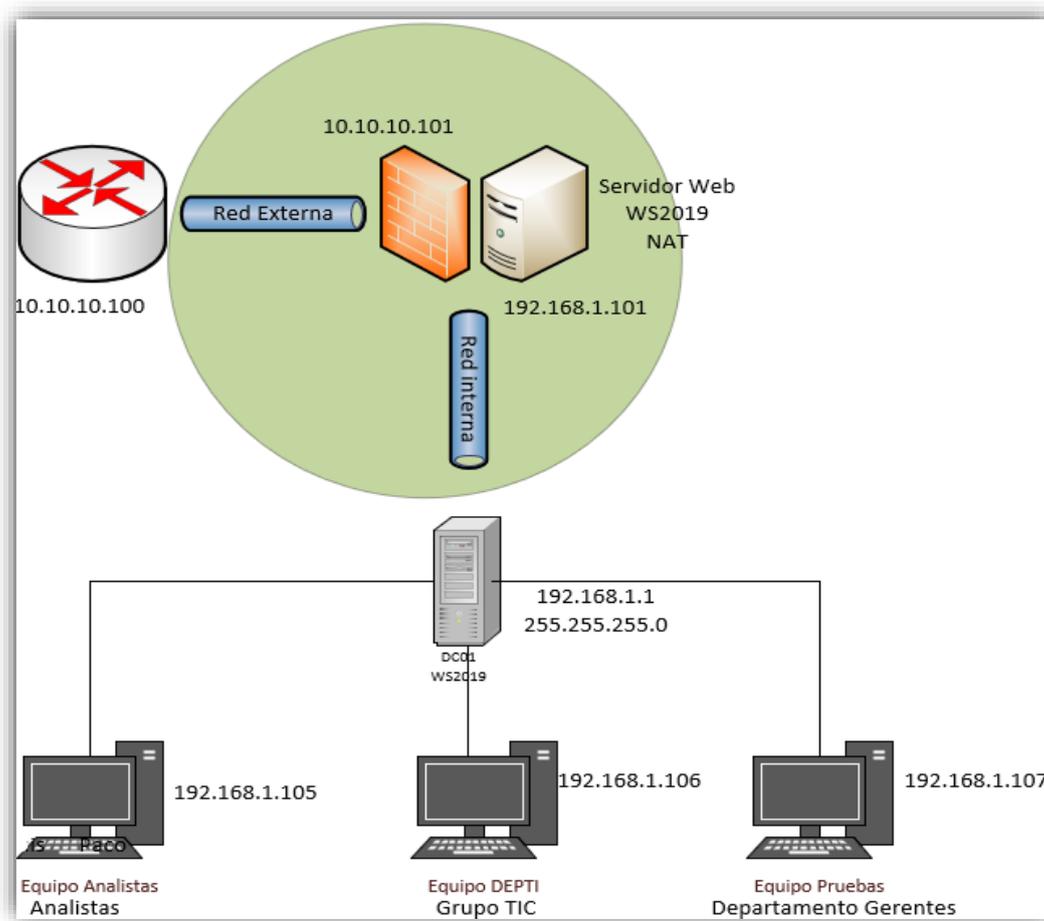


Figura 2.5: Configuración NAT y direcciones IP de todos los ordenadores y servidores

2.2.4 VPN

Una vez configurados los adaptadores de red de la infraestructura IT y el servidor que actúa como NAT, todavía se requiere una solución que permita a los usuarios, acceder a sus puestos de trabajo o recursos teletrabajando. Para esto, se configura en el servidor que actúa como NAT, los servicios de acceso remoto, que permiten crear una VPN o Red Privada Virtual. Esta crea una LAN (Local Area Network) que permite conectar los dispositivos de forma segura por internet, sin necesidad de que los equipos que la forman estén físicamente conectados entre sí, usando como soporte Internet. De esta forma se crea una VPN para acceder desde el exterior, a la infraestructura IT creada para la empresa, lo que permite que un usuario del dominio acceda desde el exterior de la red empresarial, a un puesto por Terminal Server, ejecutando una Conexión de Escritorio Remoto o a un recurso compartido escribiendo en el explorador de archivos la dirección donde se encuentra en la red interna y teniendo activa la conexión VPN. [10]

Sintetizando lo anterior, los empleados podrán teletrabajar en sus PCs gracias a la VPN y Escritorio Remoto desde fuera de la empresa. Para acceder, deberán autenticarse usando unas credenciales de dominio válidas, además de tener configurada una conexión VPN con la contraseña compartida o un certificado.

Las **ventajas** de implantar la VPN en la organización son:

- Los empleados tendrán acceso a sus puestos de trabajo en cualquier lugar.
- Los archivos y recursos de la organización tendrán una mayor seguridad y no estarán expuestos a cualquier persona en la red.
- Permite que los empleados tengan acceso a recursos corporativos.
- Ubicación anónima el usuario saldrá a internet como si estuviera en el mismo sitio donde se encuentre el ServidorWeb.

A continuación, una serie de conceptos básicos que se usarán en este capítulo:

- **Protocolo de autenticación del túnel:** es el que se utiliza en el momento en el que se crea o establece el túnel. Hay que procurar que sea seguro para evitar ataques (Man in the Middle). [16]
- **Protocolo de seguridad de túnel:** es el que se utiliza una vez que el túnel ha sido creado.
- **Certificado:** sirve para identificar y establecer de forma segura comunicaciones vía VPN.
- **IPSEC** es un conjunto de protocolos que se encargan de asegurar las comunicaciones, autenticando y cifrando cada paquete IP en un flujo de datos.
- **Protocolo EAP-MSCHAPv2** protocolo de autenticación por desafío mutuo de Microsoft. Se caracteriza por:
 - Permitir la autenticación mediante el uso de usuario y contraseña.
 - Credenciales de Winlogon que permiten especificar la autenticación mediante las credenciales de inicio de sesión del equipo.
 - Autenticación bidireccional de la conexión entre cliente y servidor.

- **Protocolo PEAP** es un protocolo de autenticación que destaca por:
 - Su funcionamiento es similar al protocolo EAP usado en VPN L2TP/IPSEC con clave compartida.
 - Usa un canal seguro (TLS) que protege el proceso de autenticación del ataque conocido como 'Man in the Middle Attack'.
 - Proporciona privacidad e integridad a los datos.

Se van a describir dos tipos de VPN que se implementarán en este proyecto, la primera versión VPN L2TP/IPSEC y una segunda versión mejorando las prestaciones y seguridad de la primera VPN IKEv2.

VPN L2TP/IPSEC

El protocolo de autenticación túnel usado en esta versión será el protocolo EAP, es importante que tanto el Servidor de la infraestructura donde está configurada la VPN como los dispositivos que tengan configurada la VPN soporten el mismo tipo de seguridad EAP.

La autenticación se realiza de la siguiente manera; el cliente envía un desafío, el cual se encuentra cifrado con la contraseña compartida, al recibirlo el servidor cifra su desafío con la contraseña compartida y comprueba que ha generado lo mismo que ha recibido del cliente. Si la comprobación es correcta se envía al cliente una respuesta autenticada que se basa en el desafío enviado previamente.

El protocolo de seguridad IPSEC es usado para cifrar los paquetes intercambiados asegurando conexiones seguras. Evitando hackeos de los datos enviados a través de la VPN o de que una tercera persona pueda capturarlos.

VPN IKEV2

El protocolo **IKEv2** es un protocolo de cifrado de peticiones y respuestas que fue el resultado de la colaboración entre Microsoft y CISCO.

Este protocolo tiene la ventaja de que proporciona una conexión más rápida, segura y estable, que los demás protocolos, gracias al soporte de MOBIKE, en caso de romperse la conexión, es capaz de volver a reconectar obviando el proceso de autenticación local.

Además, incluyen su uso en teléfonos móviles, mediante Mobile VPN con IKEv2 ya que ofrece un protocolo de movilidad y multiproveedor muy flexible para cambiar de redes.

El protocolo de autenticación usado en esta versión será el protocolo PEAP. [17]

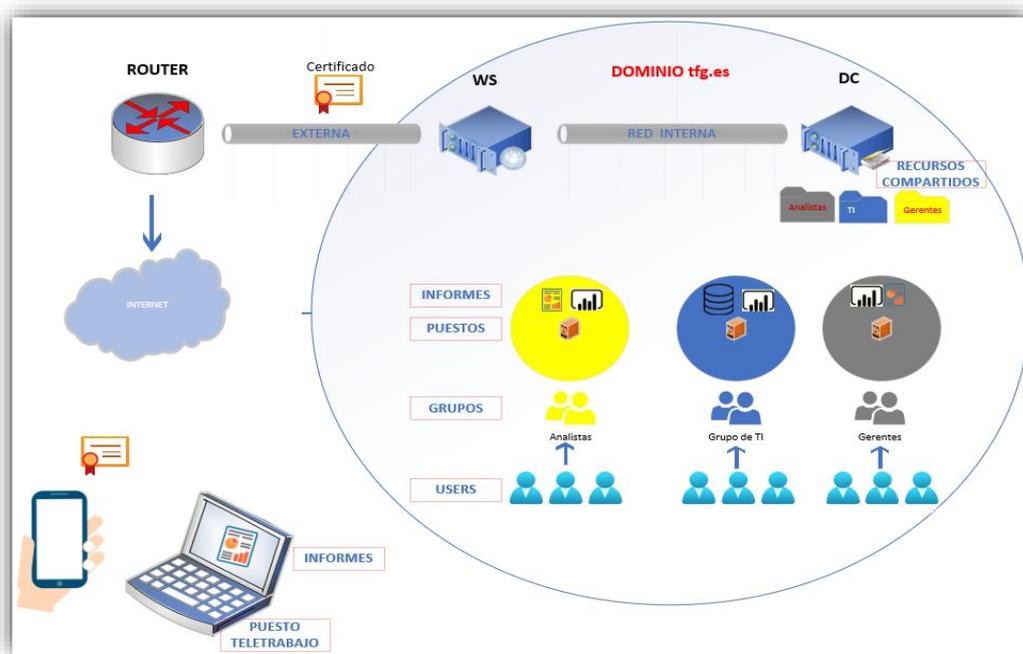


Figura 2.6: Infraestructura con VPN

2.2.5 Tecnologías similares

VirtualBox, es el virtualizador desarrollado por Oracle, ofrece la ejecución de máquinas virtuales de manera remota, a través de Remote Desktop Protocol.

VMware, es un hipervisor que soporta gran cantidad de Sistemas Operativos, reduce el tiempo de inactividad y simplifica la gestión del centro de datos. [18]

OpenVPN, esta herramienta Open Source permite la creación de conexiones VPN con cifrado SSL y TLS de forma segura. También es una aplicación que realiza las tareas de servidor y cliente en los equipos implicados. [19]

WireGuard, creado por Linux, es un VPN que mejora el rendimiento de los protocolos IPsec y OpenVPN. Utilizan un código de numerosas líneas para conseguir un túnel más seguro y de fácil administración. [20]

2.3 Especificación de la configuración

En esta sección se especificarán los detalles de los elementos que conforman la infraestructura IT.

En la siguiente tabla se verán la configuración de las máquinas virtuales en Hyper-V.

Nombre	SO	Tamaño de disco	Memoria RAM	Conmutadores
Controlador Dominio	WS2019	128 GB	3096 MB	Interno
ServWeb (NAT)	WS2019	128 GB	3096 MB	Interno y externo
Puesto TI	W10	128 GB	3096 MB	Interno
Puesto Analistas	W10	128 GB	3096 MB	Interno
Puesto Pruebas	W10	128 GB	1024 MB	Interno

Tabla 2.1: Especificación VM en Hyper-V

Una vez creadas las máquinas virtuales y unidas al dominio, hay que realizar la configuración de direcciones de red IPV4 como se muestra en la tabla 2.2 e IPV6 en la tabla 2.3.

Nombre	Dir. IPV4	Máscara	DNS	P. ENLACE
DC	192.168.1.1	255.255.255.0	192.168.1.1	192.168.1.101
Servidor web (int)	192.168.1.101	255.255.255.0	192.168.1.1	-
Servidor web (ext)	10.10.10.101	255.255.255.0	8.8.8.8	10.10.10.100
Equipo Analistas	192.168.1.105	255.255.255.0	192.168.1.1	192.168.1.101
Puesto TI	192.168.1.106	255.255.255.0	192.168.1.1	192.168.1.101
Equipo Pruebas	192.168.1.107	255.255.255.0	192.168.1.1	192.168.1.101

Tabla 2.2: Configuración direcciones IPV4

Nombre	Dir. IPV6	Long prefijo	DNS	P. ENLACE
DC	Fd::2	64	Fd::2	Fd::3
Servidor web (int)	Fd::3	64	Fd::2	-
Servidor web (ext)	-	-	-	-
Equipo Analistas	Fd::4	64	Fd::2	Fd::3
Puesto TI	Fd::5	64	Fd::2	Fd::3
Equipo Pruebas	Fd::6	64	Fd::2	Fd::3

Tabla 2.3: Configuración direcciones IPV6

Tras la creación y configuración de las máquinas del dominio, el siguiente paso será la creación de unidades organizativas para organizar los elementos del dominio. Se definen 3 OU, una para cada departamento, cada una de las OU contiene a su vez dos nuevas OU (Equipos y Usuarios) que contendrán los equipos de ese departamento y los usuarios respectivamente. De esta forma tendremos los equipos y usuarios pertenecientes a un departamento con una buena clasificación, en la tabla 2.4 se muestra la estructura de OU.

TFG.ES	Analistas	Equipos
		Usuarios
	Departamento TI	Equipos
		Usuarios
	Gerentes	Equipos
		Usuarios

Tabla 2.4: Estructura de OU

Ahora, se une cada equipo a la OU equipos del departamento al que pertenece, es decir, Analistas – Equipo Analistas, Departamento TI – Puesto TI y Gerentes – Equipo Pruebas.

También se crean los usuarios y se unen a la OU Usuarios del departamento al que pertenecen, además de a un Grupo de usuarios lo que facilitará aplicar GPO's:

Administrador	Administrador (Permisos para todo)	
Grupo de seguridad	Nombre del grupo	Nombre usuarios
	Analistas	luis@tfg.es
		Paco@tfg.es
	Departamento TI	smm@tfg.es
	jose@tfg.es	

	Gerentes	Antonio02@tfg.es
		l@tfg.es

Tabla 2.5: Usuarios del dominio

Para finalizar en la tabla 2.6 se especifican las Directivas de Grupo aplicadas a cada grupo de usuarios.

Grupo Analistas	<ul style="list-style-type: none"> ➤ Panel de control deshabilitado. ➤ Quitar el acceso a comandos Apagar, reiniciar y suspender. ➤ Ocultar el icono Ubicaciones de red del escritorio. ➤ Impedir el acceso al símbolo del sistema. ➤ Deshabilitada la cuenta de invitado ➤ Bloquear pantalla tras inactividad ➤ Acceso remoto y local del grupo de usuarios Analistas a su PC
Gerentes	<ul style="list-style-type: none"> ➤ Deshabilitada la cuenta de invitado ➤ Bloquear pantalla tras inactividad ➤ Acceso remoto y local del grupo Gerentes a su PC
Departamento TI	<ul style="list-style-type: none"> ➤ Deshabilitada la cuenta de invitado ➤ Bloquear pantalla tras inactividad ➤ Acceso remoto y local a los PC 's de todos los departamentos.

Tabla 2.6: GPO's aplicadas

Para la configuración de la VPN es necesario abrir los siguientes puertos en el router doméstico y dirigir el tráfico entrante por estos a la máquina Servidor Web [21], como se especifica en la siguiente tabla:

Protocolo de seguridad	Certificado	Protocolo de autenticación	Puertos
L2TP/IPSEC	Clave compartida o certificado	MS-CHAP-V2	TCP 1724 47 (GRE)
IKEv2	Certificado	IPSEC	UDP 500 UDP 4500

Tabla 2.7: Configuración puertos del router para VPN

En el siguiente apartado se explicará cómo se realiza toda la creación y configuración de toda la infraestructura.

2.4 Detalles de configuración

Una vez expuesta toda la teoría y especificado los objetos que conforman la infraestructura se procede a la creación de esta en el desarrollo de este apartado.

Para comenzar con la configuración como se indicó en el apartado [2.2.2 Active Directory](#), una vez creadas las máquinas, en la máquina **DC** se agrega el rol 'Servicios de dominio de Active Directory'. Para ello, es necesario Abrir el Administrador del servidor → Agregar roles y características, se activa la casilla correspondiente al rol, como se muestra en la figura 2.7.

Tras la instalación, el Administrador del servidor muestra una advertencia informando que se debe promover la máquina a controlador de dominio. Al hacer click sobre ella, se abre el Asistente para configuración de Servicios de dominio de Active Directory, se procede a crear el dominio '**tfg.es**', figura 2.8.

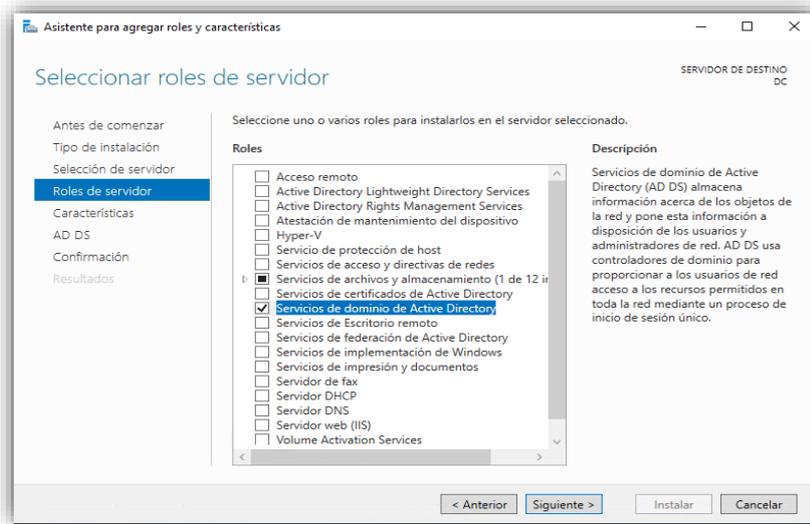


Figura 2.7: Rol de Servicios de dominio de AD

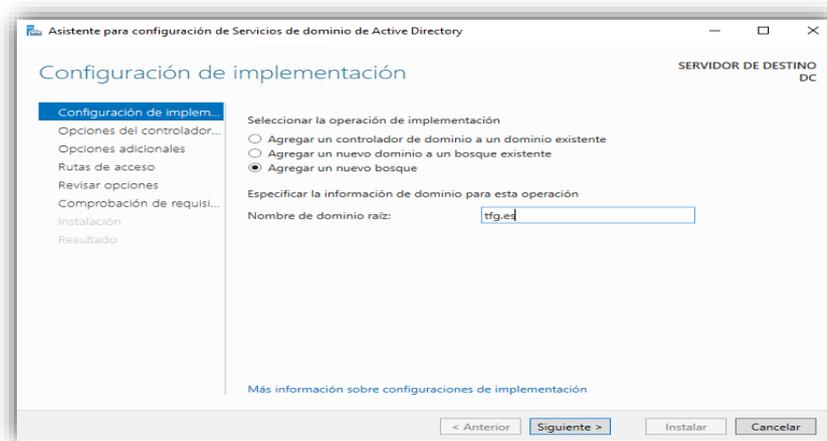


Figura 2.8: Creación del dominio 'TFG.ES'

Al finalizar el asistente se genera el script con él se puede realizar la instalación de forma desatendida.

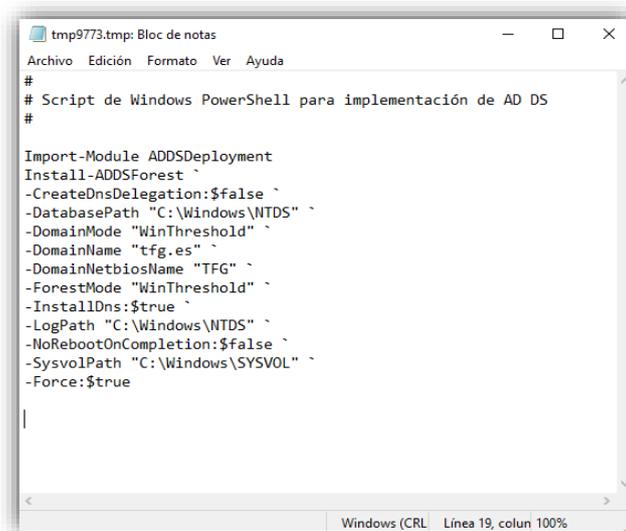


Figura 2.9: Script para crear DC y dominio

Tras finalizar la instalación se reinicia la máquina, al iniciar de nuevo en ella en las credenciales de usuario debe usar el nombre de dominio seguido de \nombreUsuario y la contraseña (tfg\Administrador). Para comprobar que la máquina es ahora controlador de dominio, abrir el Administrador del servidor -> Usuarios y equipos de Active Directory -> Desplegar el nodo del dominio 'tfg.es' -> Comprobar que en la Unidad Organizativa 'Domain Controllers' se encuentra la máquina DC como se muestra en (Figura 2.10).

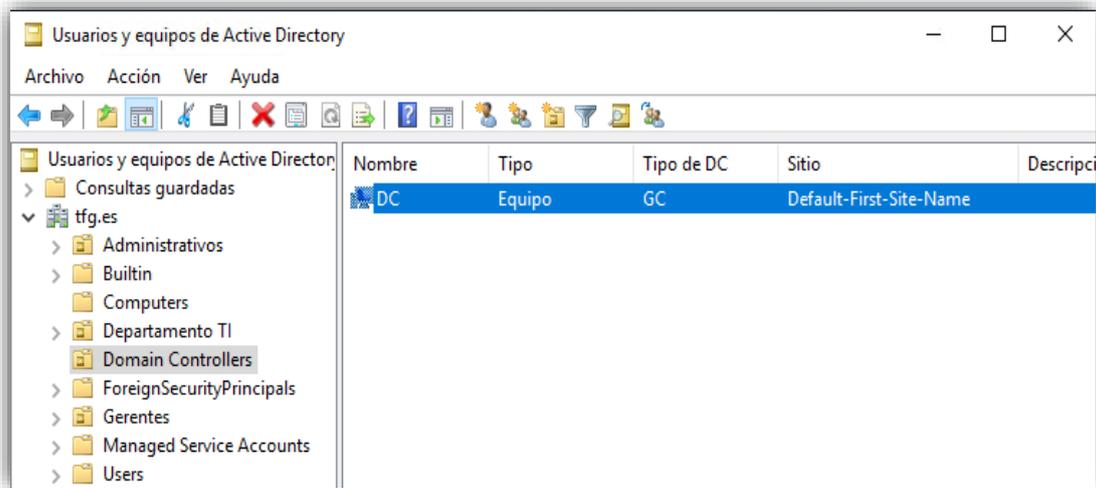


Figura 2.10: Comprobación DC pertenece al dominio y es controlador de dominio

A continuación, revisaremos la configuración del adaptador de red de la máquina DC y reconfiguraremos de nuevo la dirección del servidor DNS preferido, debido a que cuando se eleva la máquina a controlador de dominio, por defecto asigna al DNS el valor del loopback 127.0.0.1 en IPV4 y ::1 en IPV6. Hay que cambiarlos a 192.168.1.1 en IPV4 y FD::2 en IPV6.

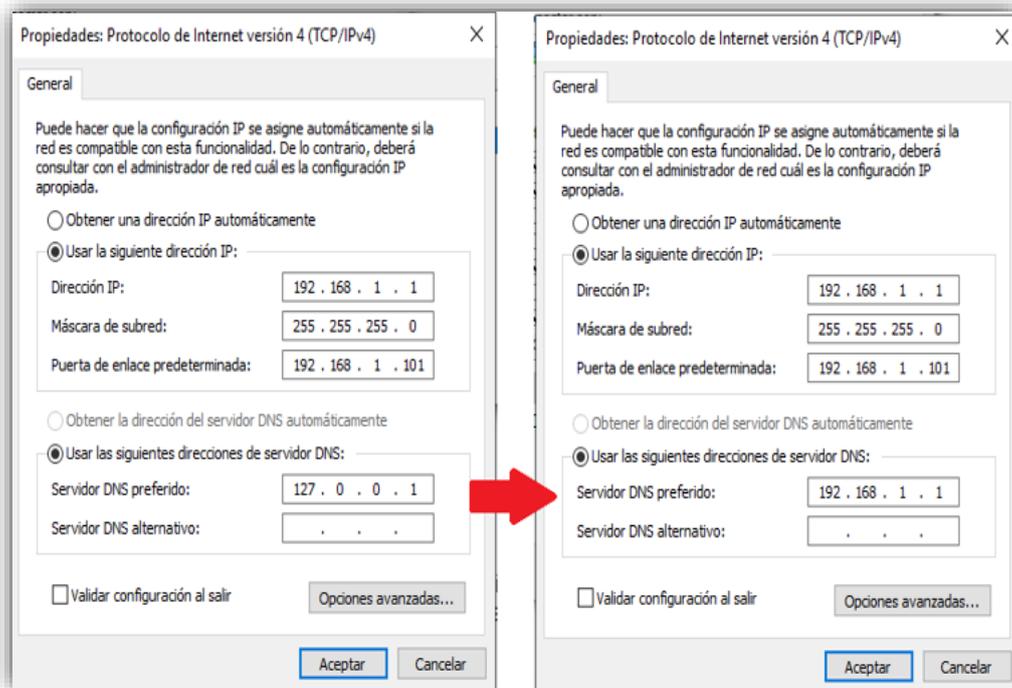


Figura 2.11: Configuración del adaptador de red en el DC

Continuaremos uniendo el Servidor Web al dominio (Figura 2.12). Para ello, en el administrador del servidor, seleccionar la pestaña de Servidor local e introduce el nombre del dominio “tfg.es” tras reiniciar el ServidorWeb, se escriben las credenciales del usuario tfg\Administrador para iniciar sesión como usuario del dominio.

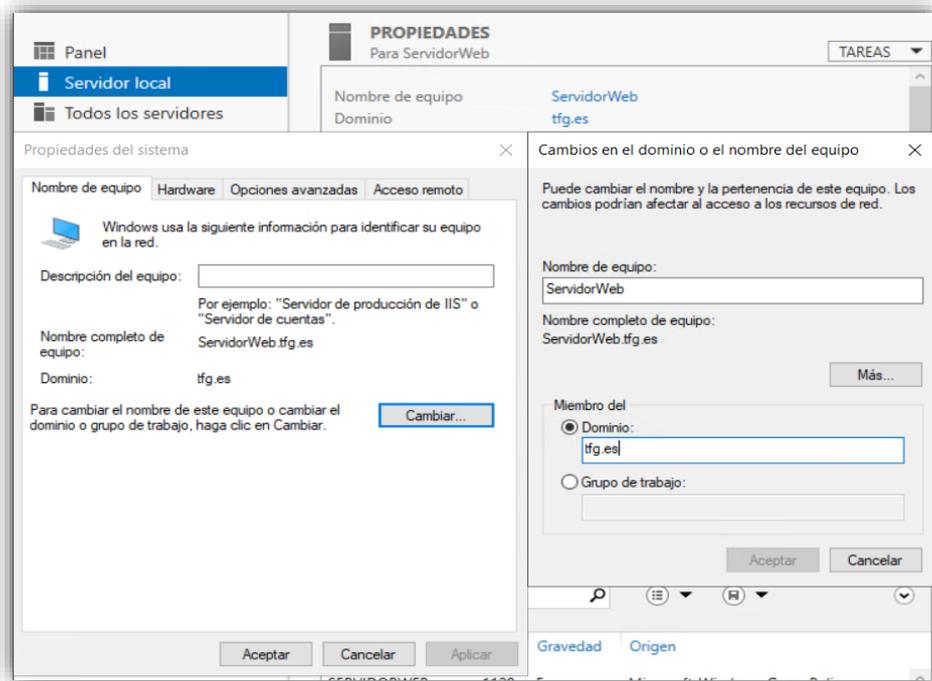


Figura 2.12: Unión del SW(NAT) al dominio

De la misma forma se unen todas las máquinas virtuales al dominio y se configuran sus adaptadores de red con los datos mostrados en [2.3 Especificación de la configuración](#).

A continuación, se muestra cómo se realiza la **creación de OUs**, se debe desplegar el nodo Usuarios y equipos de Active Directory → sobre el nombre del dominio → Acción → Nuevo → Unidad Organizativa.

Como resultado se tiene la estructura mostrada en la figura 2.13:

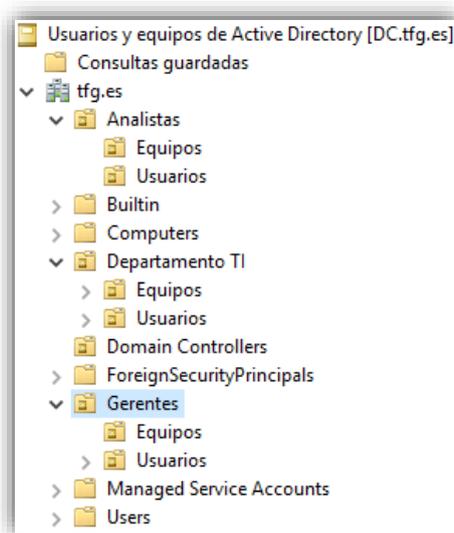


Figura 2.13: Estructura OU

Seguidamente se crearán los usuarios de cada departamento (Figura 2.14), dentro de cada departamento se crea un grupo de usuarios, a los que se agregan los usuarios de dichos departamentos. En este proyecto, al tratarse de una pequeña empresa con pocos usuarios esto puede parecer redundante, pero en el caso de que la empresa crezca, una buena organización inicial facilita mucho el trabajo.

La figura 2.15 muestra como está organizado el departamento de Analistas con dos OU: equipos y usuarios.

- Equipos: contiene los ordenadores que usan los analistas.
- Usuarios: que contiene las personas que forman el departamento.

El Grupo de usuarios Analistas afecta a los usuarios de dicho departamento, facilitando en el futuro añadir nuevos componentes al equipo o aplicar directivas de grupo.

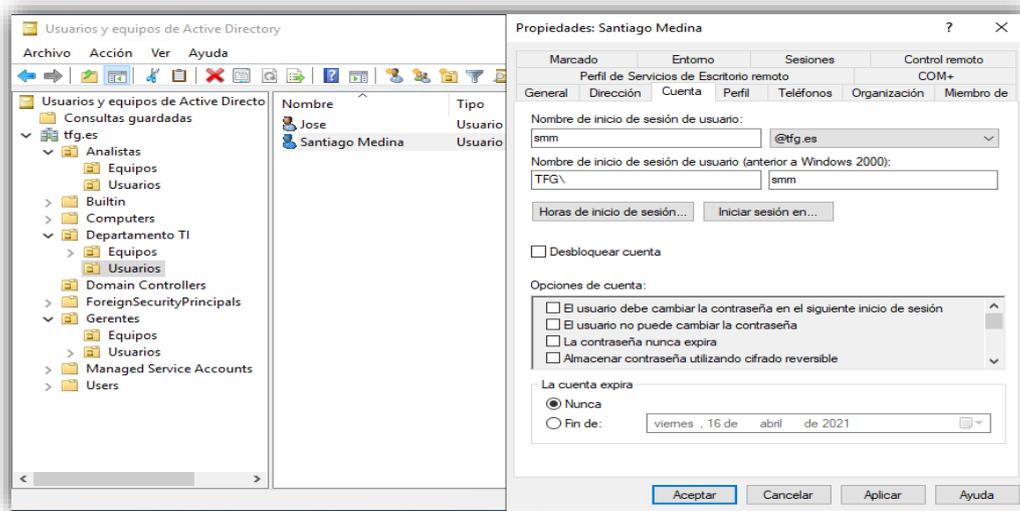


Figura 2.14: Creación de usuarios

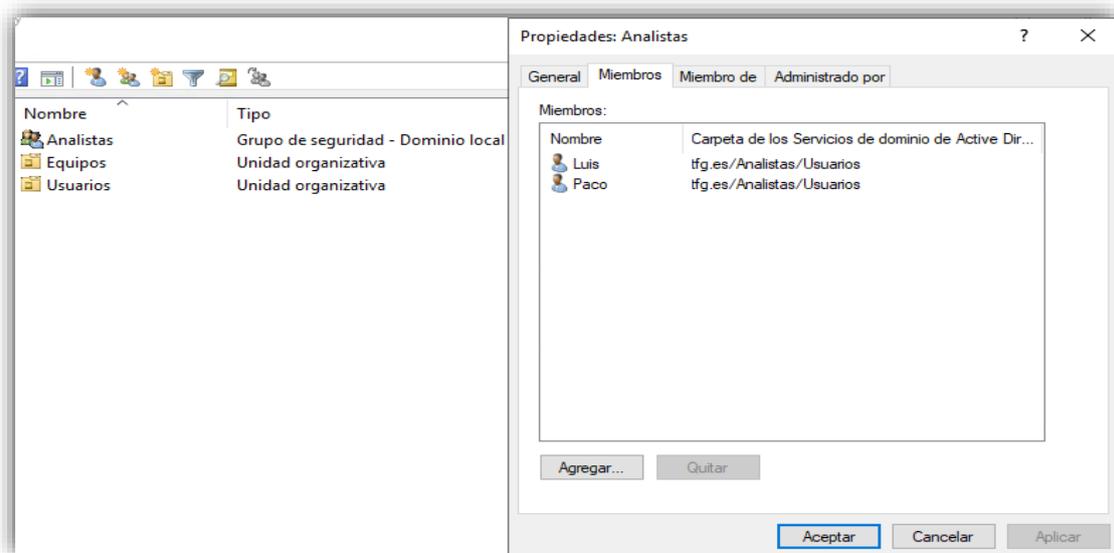


Figura 2.15: Creación de grupos de usuarios

Para finalizar con el apartado de Active Directory, se aplican las directivas de grupo especificadas en la sección anterior, en caso de tener políticas que se contradigan, se aplicará la más específica, es decir, en caso de tener una a nivel de dominio que contradiga a otra a nivel de usuario, la última será la

política aplicada para ese usuario. Si no se contradicen, se unen ambas políticas aplicándose la suma de ambas. En la figura 2.16 se muestra cómo se dan permisos de acceso al controlador de dominio y al servidor web para poder realizar tareas de mantenimiento a los usuarios del grupo Departamento TI.

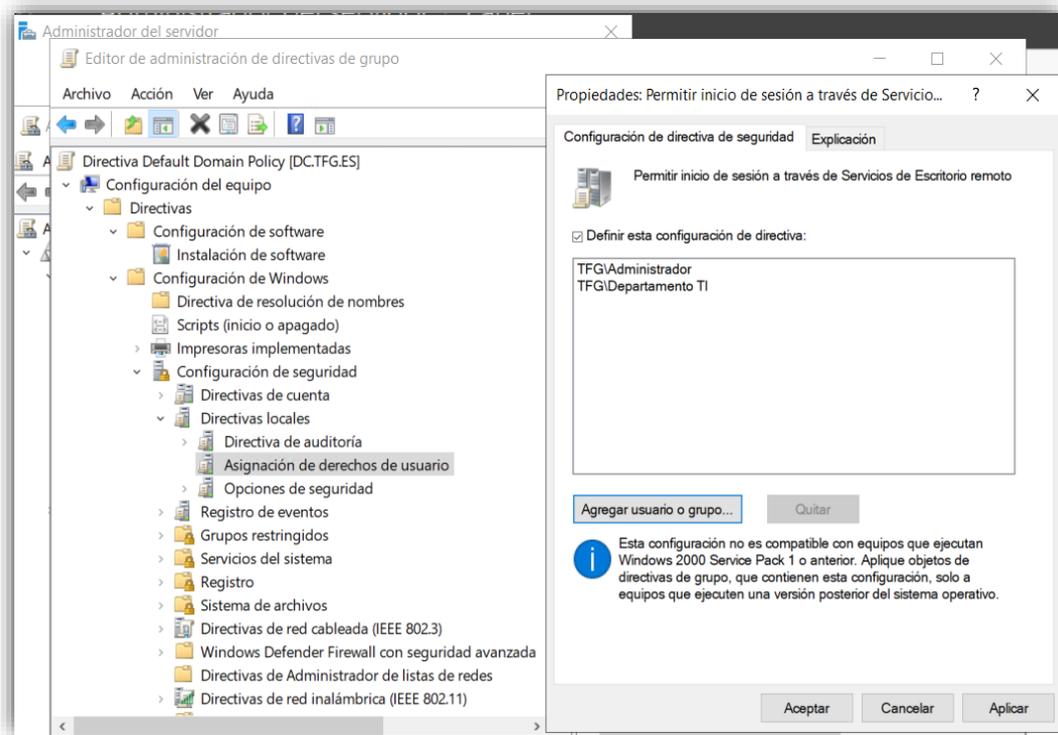


Figura 2.16: Derechos de acceso a usuarios del Departamento TI al DC y ServidorWeb

La figura 2.17, ilustra como se crea una nueva GPO (Group Police Objects) para los usuarios que componen el grupo de Analistas, a los que se prohíbe el acceso al panel de control.

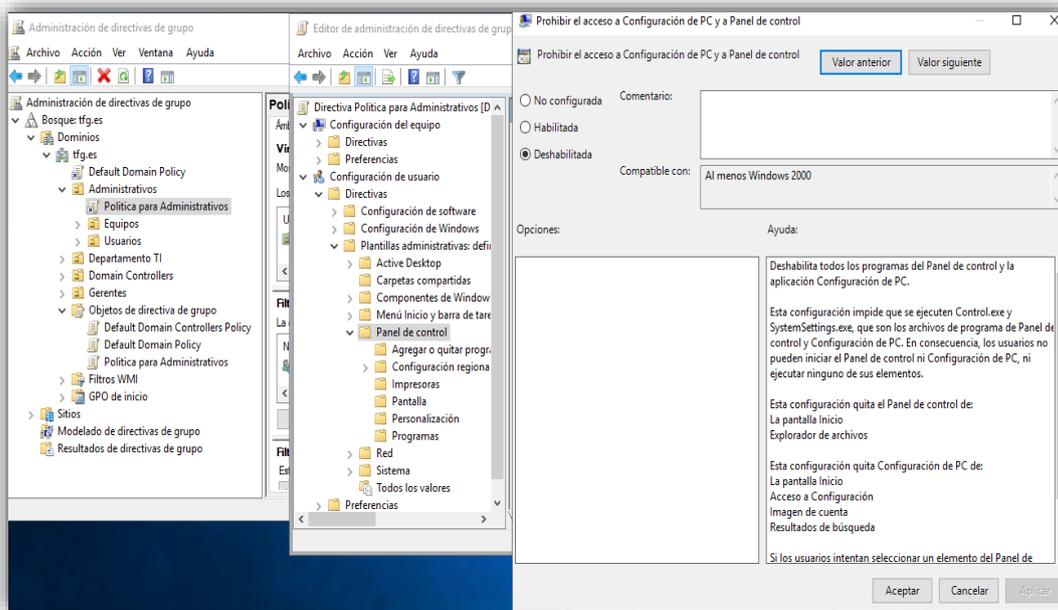


Figura 2.17: GPO para prohibir el acceso al PC

En figura 2.18 se muestra la forma de establecer a los usuarios los horarios de inicio de sesión.

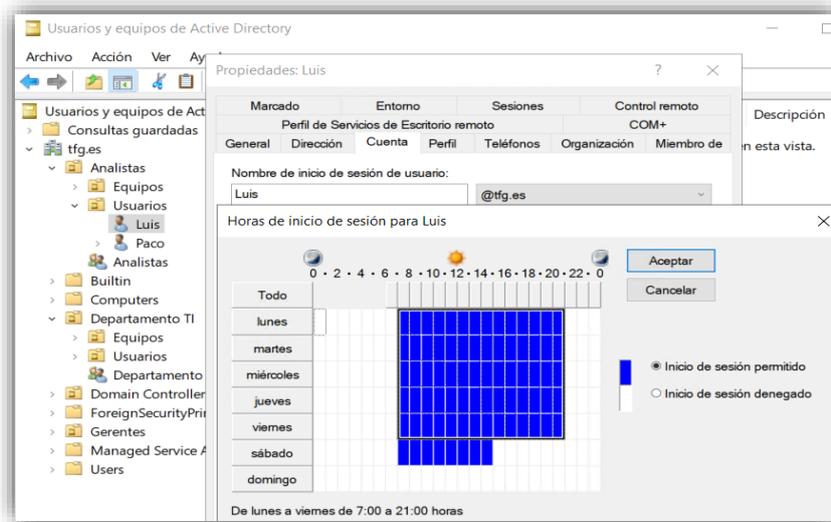


Figura 2.18: Establecer horas de inicio de sesión para un usuario

Una vez configurados los puestos y todos los elementos del dominio, es la hora de **configurar el NAT y la VPN**. Para ello, la máquina ‘Servidor Web’ se debe configurar como un router, para que el controlador de dominio y los demás ordenadores de la organización puedan acceder a Internet. Para conseguirlo se le agrega el rol de Acceso remoto (RDS), que permite a los usuarios gestionar los equipos de la organización de forma no presencial, mediante el uso de una VPN y enrutar los paquetes.

Al instalar este rol, se seleccionan los servicios de ‘**Enrutamiento**’, ‘**DirectAccess y VPN**’.

Lo que permitirá como se mencionó anteriormente; la compatibilidad del ServidorWeb con NAT, para permitir la entrada y salida de paquetes entre la red interna y externa de la empresa; hace que redes de ordenadores que utilizan direcciones IP privadas, se conecten a Internet usando una única dirección IP pública. En cuanto a la seguridad, las máquinas conectadas a la red mediante NAT no son visibles desde el exterior, por lo que un atacante externo no podría averiguar si una máquina está conectada o no a la red. Además, la VPN permite conectar dispositivos de forma segura por internet, sin necesidad de que estén conectados físicamente entre sí. En la figura 2.19 se muestra su instalación.

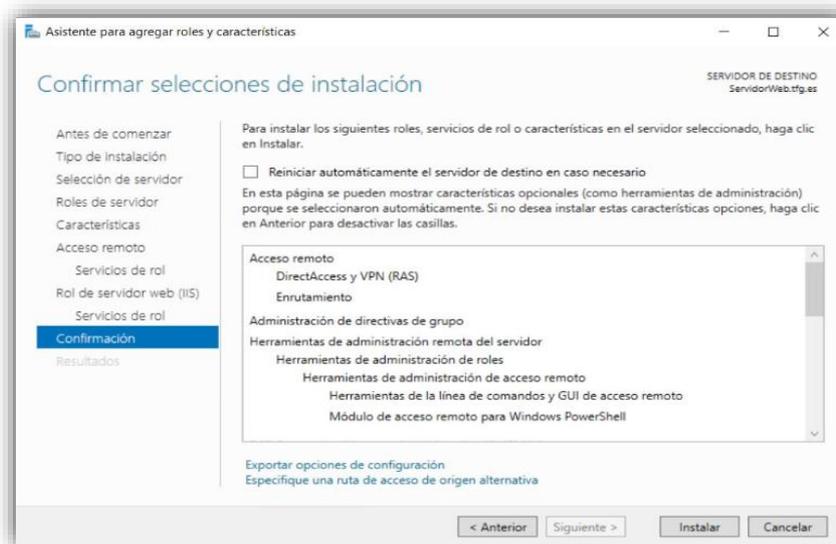


Figura 2.19: Instalación del rol RDS

Una vez instalado el rol, el Administrador del servidor muestra una advertencia informando que debemos configurar el “acceso remoto”. Continuamos eligiendo la opción “implementar sólo VPN”.

El siguiente paso es activar los servicios de “acceso remoto” que se van a usar, seleccionando las casillas ‘Acceso a VPN’ y ‘NAT’ (figura 2.20).

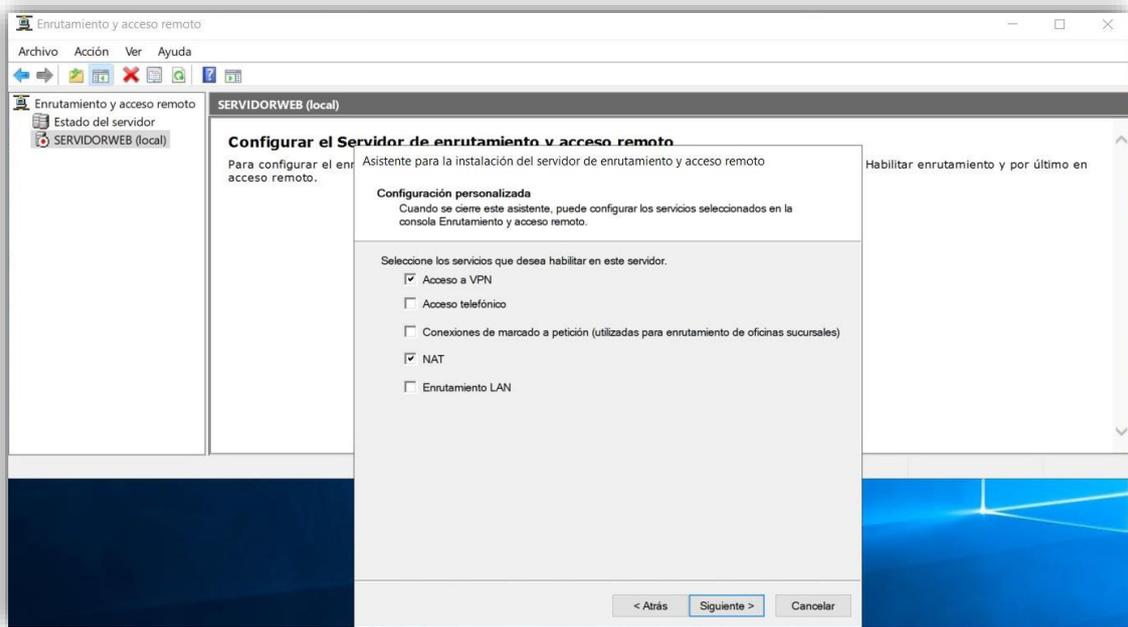


Figura 2.20: Instalación servidor de enrutamiento y acceso remoto

Tras completar la configuración, se reinicia el servidor. El icono ahora será verde, esto nos indica que el servidor está funcionando.

Comenzaremos configurando el NAT en el ‘ServidorWeb’, para que funcione como router entre la red Interna y Externa.

Se despliega el árbol del SERVIDORWEB, en el nodo IPV4 se agregan dos nuevas interfaces; una para la red externa y otra para la red interna de la infraestructura (figura 2.21).

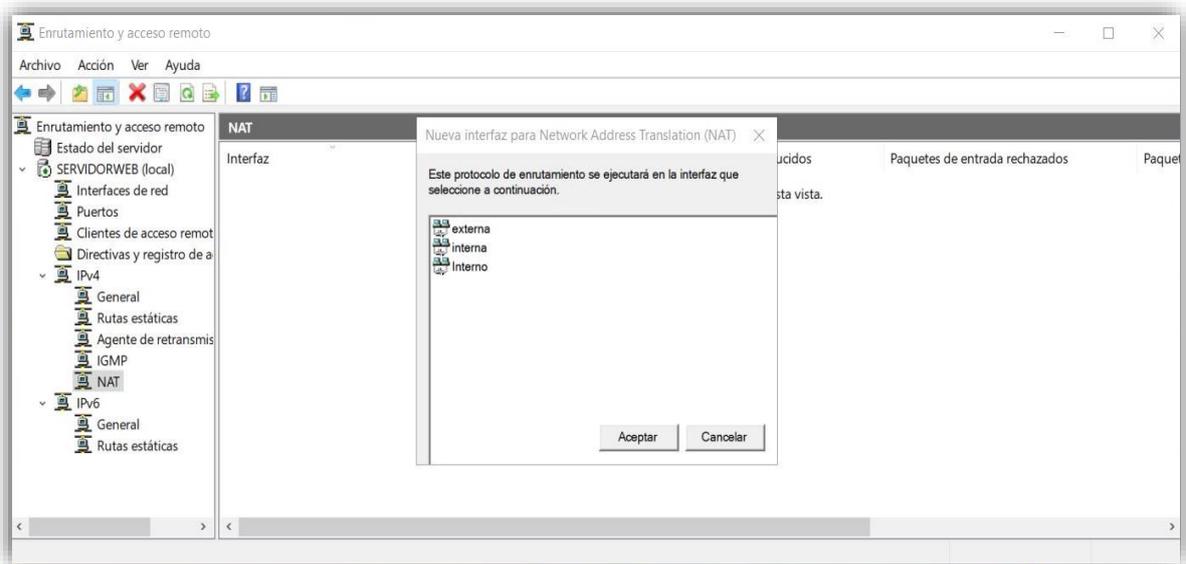


Figura 2.21: Creación interfaz NAT

En la red externa se marca la opción de ‘Habilitar NAT en esta interfaz’, que habilita a los clientes en la red, el envío y recepción de información desde internet. Además, en servicios y puertos se habilita el ‘Servidor web (HTTP)’ asignándole la dirección IP del ServidorWeb en la red interna (figura 2.22).

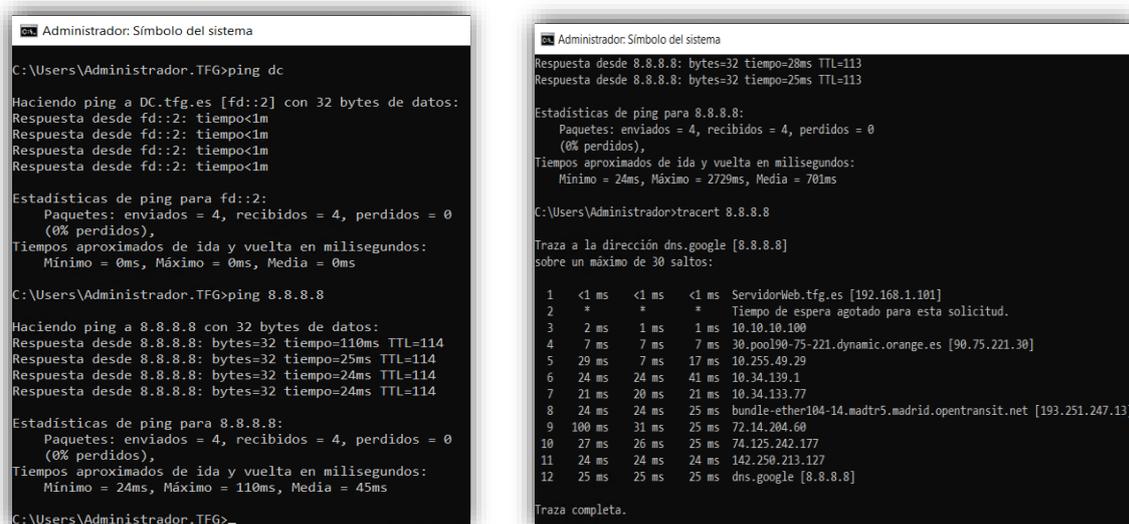


Figura 2.22: Designación de la dirección del servidor

Una vez realizada la configuración de redes en el DC y servidor web, se habilita la regla en el firewall ‘Archivos e impresoras compartidos (petición eco: ICMPv4 de entrada)’, para permitir realizar ping entre los dos equipos y comprobar el correcto funcionamiento de la red interna.

La red externa se comprueba desde el servidor web haciendo, en primer lugar, un ping al router y un segundo ping al DNS de Google. También hacemos ping para comprobar el correcto funcionamiento de la red interna a la máquina DC. Para finalizar la configuración del NAT, debemos comprobar que la máquina DC puede conectar con Internet a través del ServidorWeb, haciendo un ping al DNS de Google (8.8.8.8).

Se hacen pruebas de conectividad entre dispositivos de la red interna y dispositivos saliendo a internet a través del Servidor Web (Figura 2.23).



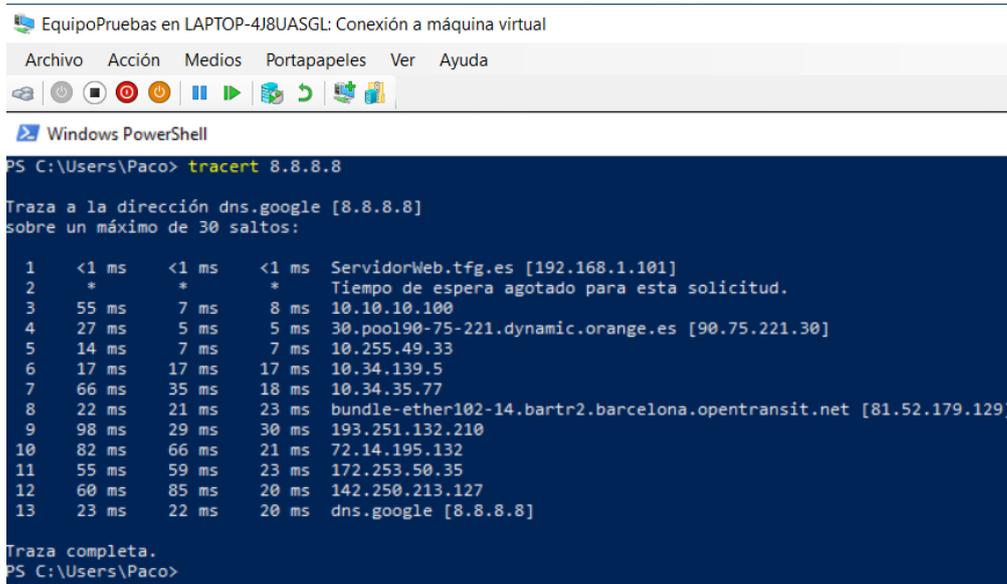


Figura 2.23: Pruebas de red interna y externa

A partir de este momento, podemos considerar que la infraestructura creada está funcionando correctamente, y con conexión a Internet. Por último, en este capítulo se configura la VPN para poder acceder desde el exterior y poder teletrabaja.

A continuación, se muestran los pasos a seguir para configurar la primera versión de la VPN **L2TP/IPSEC**.

En primer lugar, será necesario reservar direcciones IPV4 para que los usuarios de la empresa puedan acceder a la VPN, se reservan 11 direcciones estáticas desde 192.168.1.110 hasta 192.168.1.120, seleccionando propiedades en la consola de Enrutamiento y acceso remoto como se muestra en la figura 2.24, las cuales se asignan a estos usuarios de la empresa cuando se conectan a la VPN, la primera es reservada por el propio servidor.

De esta manera se garantiza que no se asignen estas direcciones a ningún otro dispositivo.

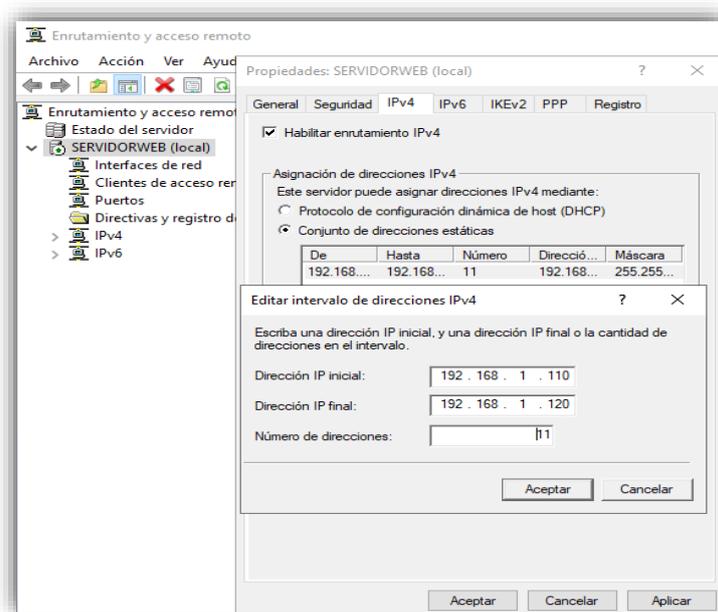


Figura 2.24: Reserva de direcciones DHCP para VPN

En segundo lugar, se deben de **habilitar los puertos de entrada** (figura 2.25) y salida (figura 2.26) ‘Enrutamiento y acceso remoto (L2TP)’, en el Firewall de Windows Defender.

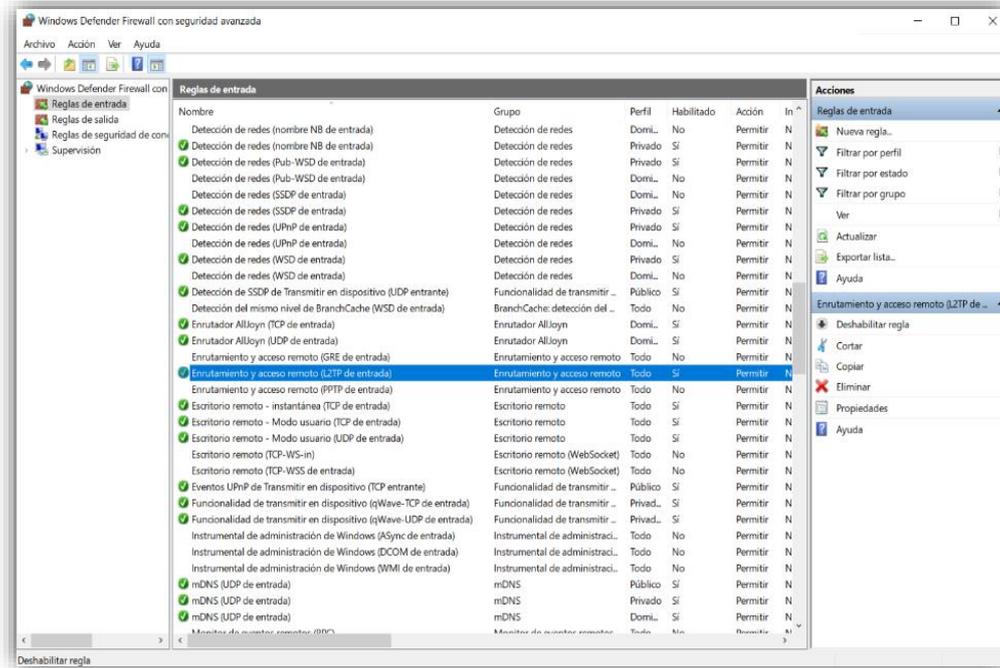


Figura 2.25: Firewall, reglas de entrada

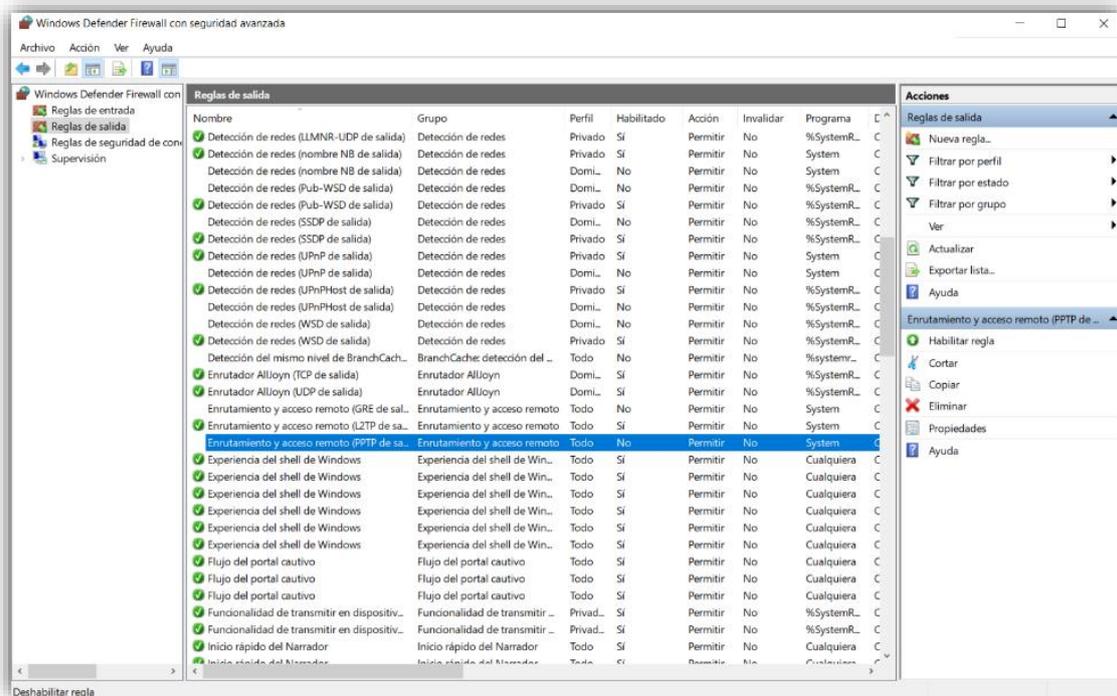


Figura 2.26: Firewall, reglas de salida

Además, se desactivará el ‘Enrutamiento y acceso remoto (GRE)’ y ‘Enrutamiento y acceso remoto (PPTP)’ que se encuentran activados por defecto. El protocolo L2TP utiliza el puerto TCP 1701, estas reglas sirven para permitir el tráfico de Protocolo de túnel de nivel 2.

En el ServidorWeb se debe **configurar la VPN**, para ello en ‘Enrutamiento y acceso remoto’ → Propiedades → Seguridad, se utilizará el Protocolo de autenticación extensible (EAP) y una clave compartida (los PC que se configuren la VPN deben usar la misma) como ilustra la figura 2.27. Por último, se reinicia el Servicio de enrutamiento y acceso remoto en ServidorWeb.

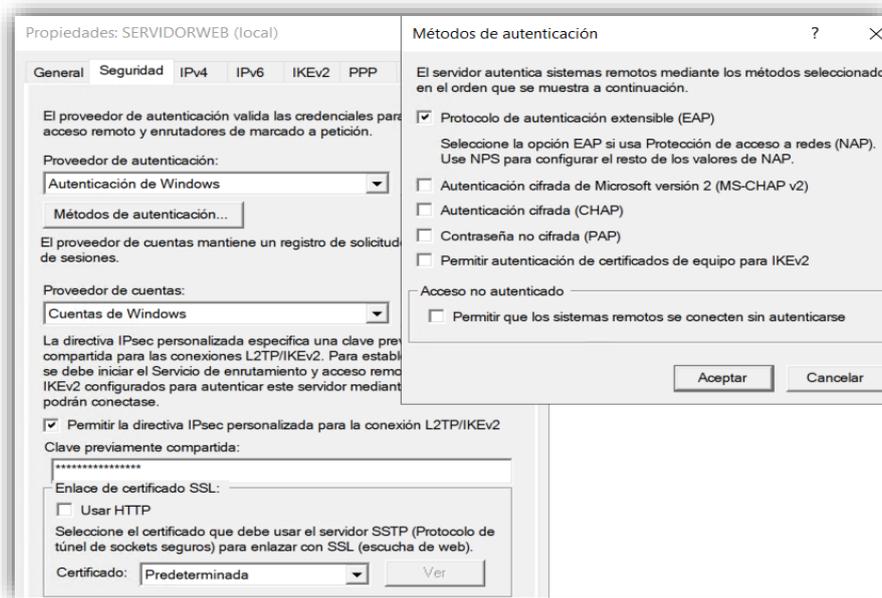


Figura 2.27: Configuración VPN L2TP en el Servidor

Una vez configurada en el servidor, se configura la conexión VPN en un PC y probaremos su funcionamiento desde otra red externa a la red de la infraestructura mediante la VPN. Para ello, en centro de redes y recursos compartidos se configura una nueva conexión o red → Conectarse a un área de trabajo (Figura 2.28 izquierda). Por último, se configuran las propiedades de la VPN en el PC, poniendo la misma clave compartida que en el servidor web (Figura 2.28 derecha)

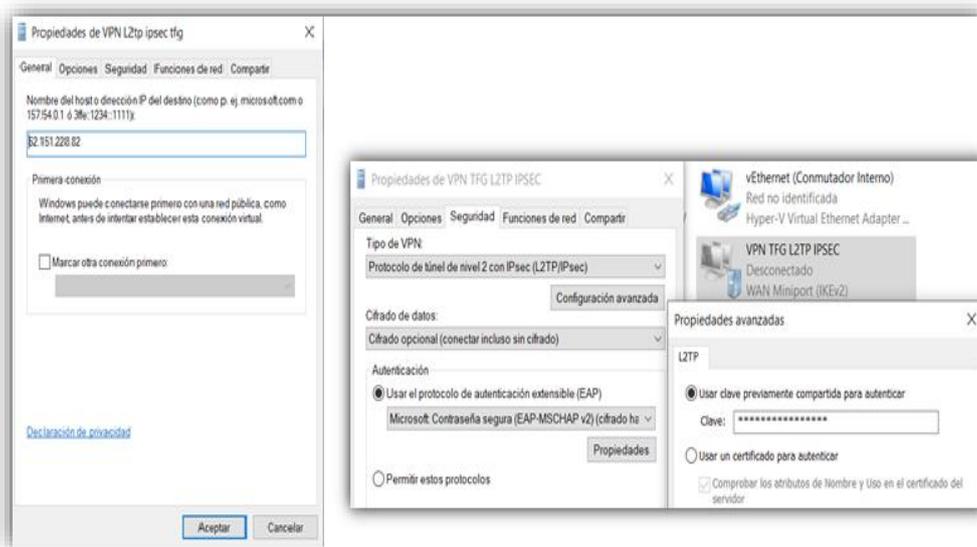


Figura 2.28: Configuración VPN L2TP en un PC

Una vez expuesta la configuración de la primera versión, se mostrará la mejora a la segunda versión con **VPN IKEv2**.

El primer paso será agregar el certificado al ‘ServidorWeb’, para ello, debemos abrir una consola (mmc), Microsoft Management Console y agregar el complemento ‘Certificados’. Que administrará certificados de la Cuenta de equipo → Equipo local.

Se importará el certificado en los nodos ‘Personal’ y ‘Entidades de certificación raíz de confianza’ (Figura 2.29).

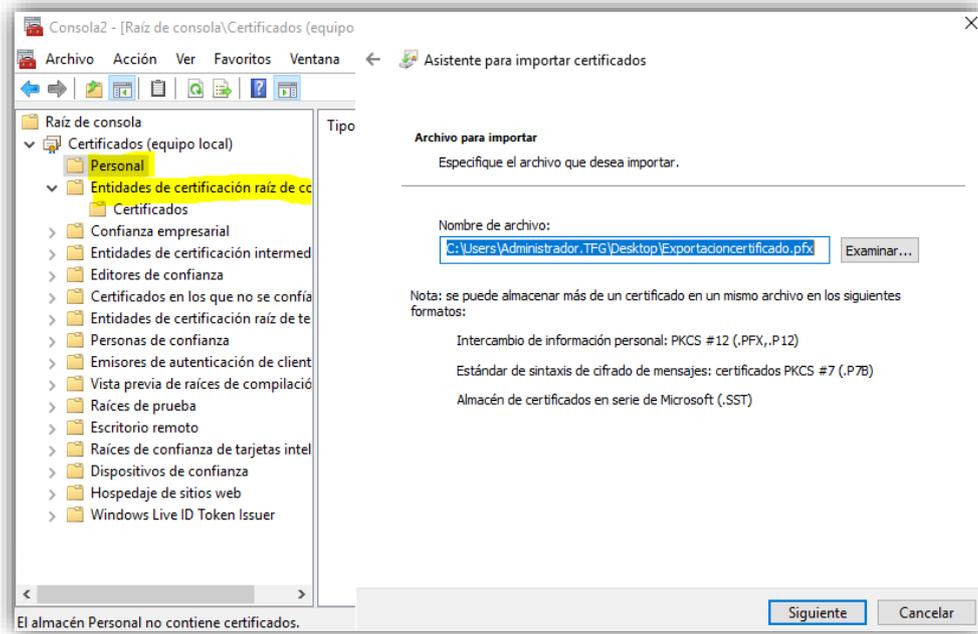


Figura 2.29: Importación de certificado en el servidor

Una vez importado el certificado se exportará para instalarlo en los demás dispositivos de la infraestructura (Figura 2.30). Se importa el certificado de la misma manera que se ha realizado previamente en el servidor, en cada uno de los puestos a través de los que se quiera teletrabajar, usando la VPN, con la diferencia de que la clave se marcará como no exportable, para que no puedan exportarlo desde esos dispositivos a otros.

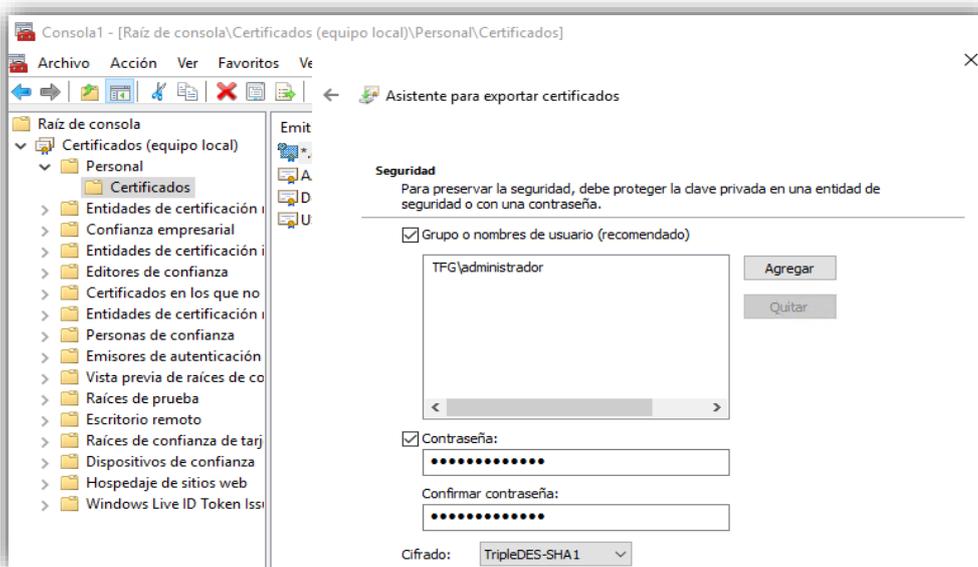


Figura 2.30: Exportación del certificado con clave

La siguiente tarea es configurar el ‘ServidorWeb’, para ello en Enrutamiento y acceso remoto → Ver propiedades ‘SERVIDORWEB (local)’ → Pestaña Seguridad, se hacen los siguientes cambios:

- Métodos de autenticación: Permitir autenticación de certificados de equipo para IKEv2.
- Deshabilitar la directiva IPsec
- Seleccionar el certificado que se ha importado previamente.

Finalizaremos creando las conexiones VPN en los PC desde los que se accede (Figura 2.31) y se realizan las pruebas de conectividad (Figura 2.32).

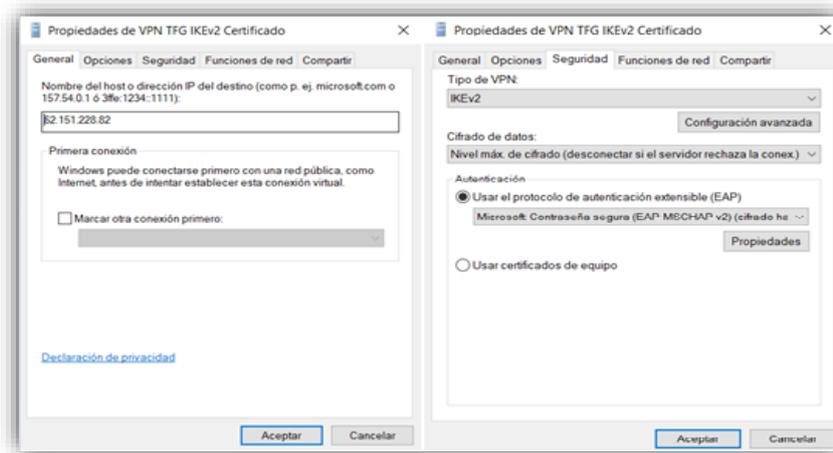


Figura 2.31: Configuración VPN IKEv2 en PC

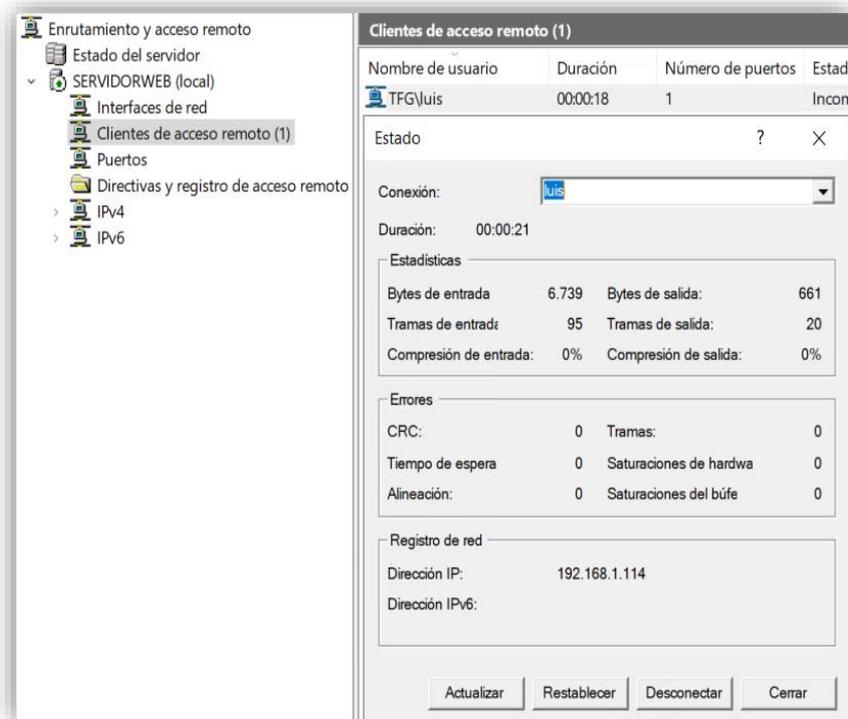


Figura 2.32: Comprobación de conexión VPN IKEv2

En el [anexo Conexión VPN](#) se muestran pruebas de conectividad de ambas versiones.

2.5 Resumen del capítulo

En este capítulo se ha visto la necesidad de construir una infraestructura de máquinas virtuales administradas en un dominio con Active Directory, esta infraestructura tiene un servidor que actúa como NAT para permitir el intercambio de paquetes entre los dispositivos del dominio e Internet. Además, de tener configurada una VPN que permita el teletrabajo de forma segura.

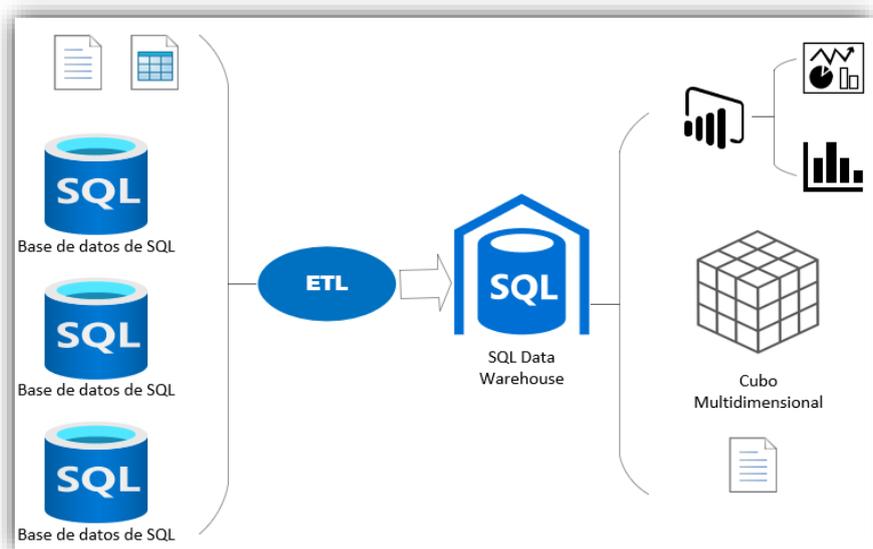
El objetivo de esta parte del proyecto era conseguir una estructura segura y útil para albergar el Sistema de Información donde se realizará la actividad principal de la empresa ficticia, el análisis de datos meteorológicos para ello, la implementación del DataWarehouse y creación de informes con Power BI se verán en los siguientes capítulos.

Palabras clave: *Hyper-V, Active Directory, OU, GPO, Dominio, NAT, VPN.*

Capítulo 3

En este capítulo, se mostrará el conjunto de datos elegido y la creación del DataWarehouse que permita apoyar la toma de decisiones en la empresa creada, incluyendo como se transforman dichos datos para conseguir una Base de Datos relacional con un diagrama en estrella para seguidamente crear el cubo multidimensional. Este apartado es fundamental para la realización del proyecto, comenzaremos creando el Sistema de Información que proveerá a la infraestructura de las herramientas necesarias, para en el siguiente capítulo poder analizar los datos y crear informes.

A continuación, se procede a describir el software necesario, así como todo el proceso realizado en detalle y las razones de por qué se realiza. También se definen conceptos y especificaciones de la configuración realizada.



3.1 DataWarehouse

En este apartado, se describe el conjunto de datos y herramientas con los que se crea el DW.

3.1.1 Herramientas y tecnologías utilizadas

En este apartado se habla en primer lugar del software necesario para la creación del DW, que será el proporcionado por Microsoft, debido a que la compatibilidad e integración de todas las herramientas que componen el caso de uso es total y durante el desarrollo del grado han sido utilizadas en numerosas asignaturas. También existe gran cantidad de documentación y cursos gratuitos de estas herramientas en la web.

En segundo lugar, se darán a conocer conceptos importantes que se tratan en este capítulo.

En la máquina virtual perteneciente al Departamento TI, donde reside el Sistema de Información será necesario la instalación del software descrito a continuación para crear el DataWarehouse.

- **Microsoft SQL Server Management Studio**, necesario para realizar el modelado de datos mencionado en el apartado anterior.
- **SQL Server 2019** con las características de motor SQL y Analysis Services, el primero es un sistema de administración de bases de datos, mientras que el segundo es un motor de datos analíticos en línea para la ayuda de toma de decisiones.
- **SQL Server Data Tools**, permite integrar Visual Studio con Analysis Services.

Otras herramientas Open Source que podrían haber sido utilizadas son:

- **Spago BI**, es una multiplataforma de Business Intelligence, que permite tanto la gestión como el análisis de datos. Ofrece la posibilidad de crear informes, OLAP, minería de datos...
- **PENTAHO**, es una plataforma basada en Java que permite la creación de cubos que se componen de archivos XML, para análisis multidimensional.

A continuación, una serie de **conceptos** básicos que se usarán en este capítulo:

- **Data Warehouse**: es un almacén que contiene gran cantidad de datos modelados y no volátiles pertenecientes a una empresa u organización variantes en el tiempo, los cuáles, se pueden obtener desde diversas fuentes con el fin de analizar datos y facilitar la toma de decisiones. Los datos son almacenados de forma segura, además, deben tener una extracción, procesamiento y presentación eficientes para llevar a cabo la toma de decisiones. Los principales **objetivos** de un DW son [1] [22]:
 - Acceso a los datos de la organización.
 - Los datos almacenados deben ser consistentes, deben tener una estructura, eliminando errores e inconsistencias.
 - Permitir a los usuarios tener varios puntos de vista (consultas y combinaciones desde cualquier dimensión), con datos fácilmente comprensibles.
 - Ofrecer alto rendimiento de consultas.
 - Los datos deben tener diferentes niveles de detalle y almacenar muchos datos históricos. Además, una característica del DW es la desnormalización, es decir, existe información que se repite.

Un DW se compone de dos tipos de tablas, las tablas de hechos y las dimensiones.

- **Tablas de hechos**, cada fila es una transacción o evento, los hechos incluyen los identificadores que los relacionan con las dimensiones y las medidas que son propiedades de los hechos que deben de tener la misma granularidad.
 - **Dimensiones**: Cada dimensión está formada por tuplas de atributos que caracterizan los hechos, es decir, dan más información sobre los hechos y se pueden organizar en jerarquías.
- Cada fila de la tabla de hechos será un evento en este caso de uso, corresponderá con la toma de una medida meteorológica, se denomina **hecho**, el cual, tendrá asociadas **medidas** que son atributos numéricos que dan información que analizar, en este caso, temperatura, lluvia, nieve...
 - **Dimensión**: proporciona más información relacionada con los hechos. Agrupa los datos con mayor nivel de detalle, una dimensión admite la posibilidad de formar jerarquías, estas tienen diferentes niveles de detalle, las representaciones jerárquicas se conocen como exploración ascendente

(**roll-up**) y exploración descendente (**drill-down**). Por ejemplo, en exploración ascendente desplaza la jerarquía hacia arriba, agrupándola en unidades mayores a través de una dimensión. En este caso agrupando, los días, en meses, trimestres o años

- **Diagrama en estrella**, es uno de los diagramas más utilizados, se compone de una tabla central que representa la tabla de hechos que está relacionada con todas las tablas que la rodean (dimensiones) que se encuentran desnormalizadas.
- **ROLAP**, Procesamiento Analítico On-Line Relacional, son sistemas y herramientas OLAP que se construyen sobre BD relacionales.

Dentro de Visual Studio Data Tools, tenemos los siguientes conceptos:

- **Origen de datos**: será la fuente desde donde se obtienen los datos, en el proyecto será la BD Noaa_Data.
- **Vista de origen de datos**, proporciona una vista del diagrama especificado como origen de datos. Permitiendo modificarla o unir múltiples esquemas para formar el modelo multidimensional e integrar datos desde distintos orígenes. En esta vista se pueden agregar o quitar tablas, crear consultas con nombre y nuevas relaciones.
- **Cubos** esta perspectiva será utilizada para crear una estructura multidimensional o cubo con el fin de integrar la información para su análisis. Aquí es donde se eligen las tablas de la vista de origen de datos que se usaran como tablas de hechos, sus medidas y las dimensiones. Una vez creado el cubo permite realizar las siguientes acciones sobre él:
 - Crear **medida y grupos de medida**, son agregaciones de datos numéricos como sumatorias, recuentos o expresiones MDX, mientras que los grupos de medida son las que contienen las medidas que se crean.
 - **Particiones** se encarga de segmentar los datos del cubo produciendo un mejor rendimiento y procesamiento de este.
 - **Agregaciones**, reduce el tiempo de respuesta del cubo pre-calculando y almacenando consultas.

3.1.2 Datos y esquema

La fuente de datos del proyecto, se obtiene de NOAA, National Oceanic And Atmospheric Administration, posee información de gran variedad de estaciones meteorológicas pertenecientes a Estados Unidos. [23] [24]

El clima puede llegar a afectar a todos los sectores de la sociedad. Desde paralizar una ciudad durante días con una nevada con el consecuente aumento de consumo de electricidad y falta de importaciones de alimentos, hasta un clima extremadamente caluroso con el riesgo de incendios o sequías. Por esto, el estudio de estos datos es de especial interés para poder comprender por qué ocurre y poder estar preparados para cuando suceda de nuevo un hecho similar, disminuyendo los daños causados en la sociedad o para estudiar el cambio climático e intentar tomar mejores decisiones para evitarlo.

Inicialmente, hay 5 archivos de texto cada uno corresponde a un estado, y están compuestos por estaciones, datos geográficos de estas, fecha y medidas meteorológicas (Figura 3.1).

Se trabaja sobre estos datos para conseguir un esquema estrella formado por una tabla de hechos rodeada de las tablas de dimensión. Este proceso se visualiza paso a paso en el **Anexo: [Scripts para modelado de datos](#)**, en la figura 3.2 se muestra el esquema en estrella obtenido.

NUEVA YORK noaa_station_id latitud_grad_dec longitud_grad_dec elevacion fecha dapr mdpr prcp profNieve nieve tMedia tmax tmin tobs wesd wesf vMediaViento dirRachaMasRapida2min dirRachaMasRapida5seg velRachaMasRapida2min velRachaMasRapida5seg	ILLINOIS noaa_station_id latitud_grad_dec longitud_grad_dec elevacion fecha dapr mdpr prcp nieve profNieve tmax tmin tobs wesd TEXAS noaa_station_id elevacion latitud_grad_dec longitud_grad_dec fecha mdpr dapr prcp nieve pNieve tmax tmin tobs dirViento velViento	FLORIDA noaa_station_id elevacion latitud_grad_dec longitud_grad_dec fecha mdpr dapr prcp nieve pNieve tmax tmin tobs dirViento velViento CALIFORNIA noaa_station_id elevacion latitud_grad_dec longitud_grad_dec fecha mdpr dapr prcp nieve pNieve tmax tmin tobs dirViento velViento
---	--	---

Figura 3.1: Tablas iniciales de Noaa_Data

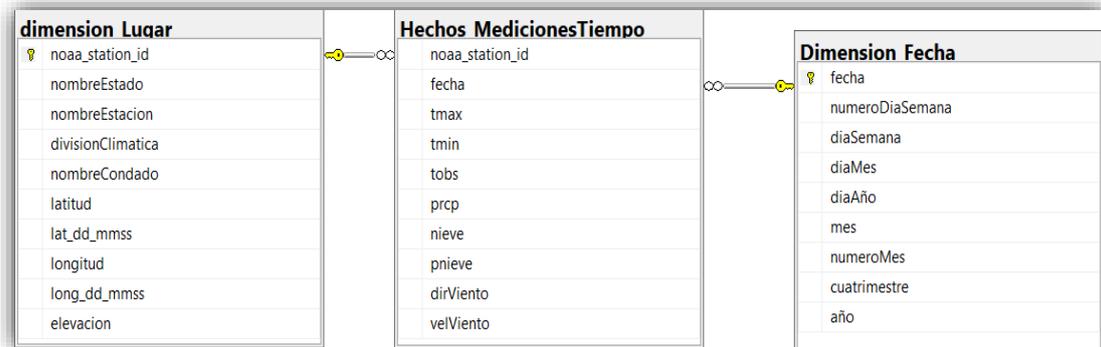


Figura 3.2: Diagrama en estrella

La tabla central del esquema ‘**Hechos_MedicionesTiempo**’ se compone de los identificadores (FK) noaa_station_id y fecha. Como medidas tiene las temperaturas máxima, mínima y observada, lluvia, nieve, profundidad de la nieve, dirección y velocidad del viento. Esta tabla, almacenará en cada fila un hecho, en este caso un evento meteorológico con las medidas mencionadas.

En la tabla ‘**dimension_Lugar**’, se indican los datos demográficos de las estaciones de la tabla de hechos, su Primary Key es noaa_station_id y se compone de los campos: estado donde se ubica la estación, nombre de la estación, división climática, condado, latitud, longitud y elevación.

Por último, la tabla ‘**Dimension_Fecha**’ almacena los datos de las fechas que comprenden los hechos, es decir, desde 01-01-2016 hasta 31-12-2019. Tiene como PK fecha y como atributos el número de día de la semana, día de la semana en letra, día del mes, día del año, mes, numero mes y año.

3.2 Especificación de la configuración

En esta sección se especificarán los detalles de las partes del cubo creado.

En primer lugar, se comenzará con una descripción de los datos que conforman la tabla de hechos (Tabla 3.1), la dimensión fecha (Tabla 3.2) y la dimensión lugar (Tabla 3.3).

Además, para la dimensión lugar se crean las consultas con nombre especificadas en la tabla 3.4, que crean dos columnas nuevas en la dimensión.

Una vez descritos estos datos se procede a mostrar las jerarquías que se deciden usar en cada dimensión (tabla 3.5 y tabla 3.6), así como, las relaciones entre ellas.

Hechos_MedicionesTiempo	
Nombre	Descripción
Noaa_station_id	Identificador de estaciones noaa
Fecha	Identificador de fecha(aaaa/mm/dd)
Tmax	Temperatura máxima (°F)
Tmin	Temperatura mínima (°F)
Tobs	Temperatura observada (°F)
Prcp	Cantidad de lluvia (100 inch)
Nieve	Cantidad de nieve (10 inch)
Pnieve	Profundidad de la nieve (10 inch)
DirViento	Dirección del viento (°)
VelViento	Velocidad viento (mph)

Tabla 3.1: Descripción de los datos de la tabla de hechos

Dimension Fecha	
Nombre	Descripción
Fecha	Identificador de fecha(aaaa/mm/dd)
numeroDiaSemana	Número del día de la semana (1 → 7)
diaSemana	Día de la semana (L → D)
diaMes	Número del día en el mes
diaAño	Número del día en el año
Mes	Nombre del mes
numeroMes	Número del mes
Trimestre	Trimestre del año
Año	Número del año de la medición

Tabla 3.2: Descripción de los datos de la Dimensión_Fecha

Dimensión Lugar	
Nombre	Descripción
Noaa_station_id	Identificador de estaciones noaa
nombreEstado	Nombre del estado donde está la estación
nombreEstacion	Nombre de la estación
DivisiónClimática	División a la que pertenece la estación
Latitud	Latitud de la estación
Longitud	Longitud de la estación
Lat_dd_mmss	Latitud en días_minutos_seg
Long_dd_mmss	Longitud en días_minutos_seg
Elevación	Elevación de la estación sobre el nivel del mar (10 x inch)
Pais	País donde se ubica la estación

Tabla 3.3: Descripción de los datos de la Dimensión_Lugar

Dimensión Lugar	
Nombre de la consulta	Descripción
Estado_Condado	Agrupar en uno los campos Condado y Estado
Estado_condado_país	Agrupar en uno los campos Condado, Estado y País

Tabla 3.4: Descripción de las consultas con nombre creadas en Dimensión_Lugar

Jerarquías:

Jerarquía País-Estado-Condado-NombreEstación
País
Nombre Estado
Nombre Condado
Nombre Estación
Noaa Station Id
Relaciones de atributo
Noaa Station Id → Nombre Estación → Nombre Condado → Nombre Estado → País

Tabla 3.5: Jerarquía de la Dimensión Lugar

Jerarquía Año-Trimestre-Mes-Día
Año
Trimestre
Mes
Día Mes
Fecha
Relaciones de atributo + Key Columns para resolver duplicidades
Fecha → Día Mes (+ Año + Mes) → Mes (+Año) → Trimestre (+Año) → Año

Tabla 3.5: Jerarquía de la Dimensión Fecha

Los cálculos especificados en la tabla 3.6 son realizados para transformar las medidas al sistema métrico utilizado en España para estudiar los datos más fácilmente.

Nueva columna calculada	Expresión
Velocidad viento (km/h)	[Measures].[Vel Viento]*1.60934
Prpc (L x m^2)	[Measures].[Prpc]*28.32
Nieve (L x m^2)	[Measures].[Nieve]*2.832
T max (°C)	([Measures].[Tmax]-32)*5/9
Tmin (°C)	([Measures].[Tmin]-32)*5/9
Tobs(°C)	([Measures].[Tobs]-32)*5/9
Elevacion (m)	[Dimension Lugar].[Elevacion]*0.3048
Pnieve (cm)	[Measures].[Pnieve]*3.048

Tabla 3.6: Columnas calculadas con expresiones MDX en la tabla de hechos

3.3 Detalles de configuración

Una vez conocida la teoría, especificados los datos del diagrama en estrella que se va a crear y jerarquías que conforman el cubo, a lo largo del apartado se verán los pasos necesarios para crearlo.

En primer lugar, se transforman los archivos de datos meteorológicos descargados, las tareas de modelado de datos realizadas en SQL Management Studio se ven en el anexo, [Scripts para modelado de datos](#).

Una vez modelados los datos, obtenemos la base de datos relacional que usa el Proyecto multidimensional y de minería de datos creado en SQL Server Data Tools. Su origen de datos es la BD 'noaa_data', para conectar con ella, utilizaremos el proveedor SQL Server Native Client 11.0' y el nombre de servidor apuntará a 'DEPTI' como ilustra la figura 3.3, que es donde se encuentra el origen de datos.

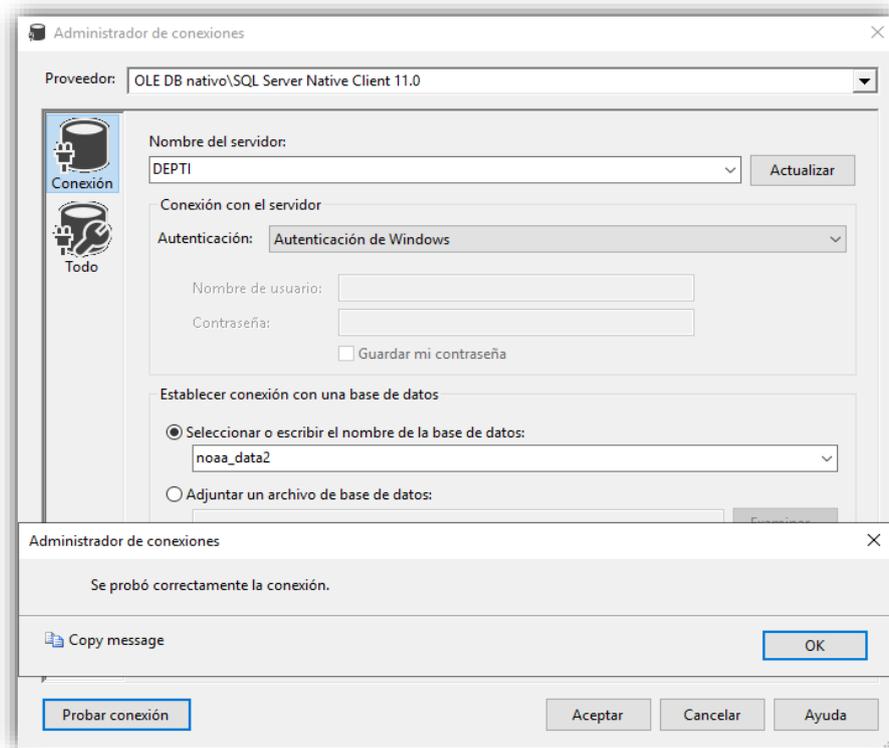


Figura 3.3: Conexión con la BD Noaa_data

Para **crear la vista del origen** de datos se utiliza el asistente. Para ello, seleccionar en el Explorador de soluciones la opción, 'Vistas del origen de datos' una vez iniciado el 'Asistente para vistas del origen de datos' se escoge el origen de datos creado previamente y las tablas de la base de datos relacional noaa_data que queremos incluir en la vista (Figura 3.4):

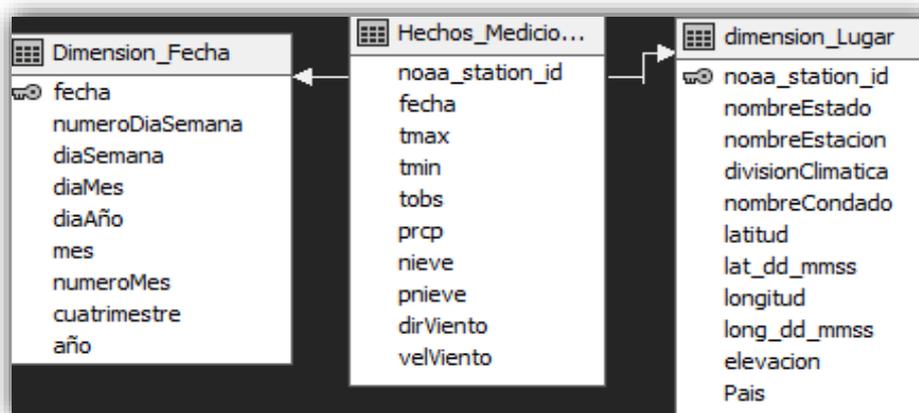


Figura 3.4: Vista origen de datos

Para **crear las dimensiones** seleccione en el explorador de archivos la categoría 'Dimensiones', click en el botón derecho y elegir la opción 'Nueva dimensión', una vez iniciado el asistente para dimensiones, que ayuda a crear la nueva dimensión, se procede a crear dos dimensiones 'Dimension_Fecha' y 'Dimension_Lugar' para ello, se utilizarán tablas existentes en el origen de datos:

- **Dimensión Fecha:** En la figura 3.5 se selecciona la información para especificar la vista origen de datos, la tabla y las columnas que son clave, para crear la dimensión. En el siguiente paso (Figura 3.6) se seleccionan todos los atributos de la tabla que resultan interesantes para crear la dimensión y se especifica el tipo del atributo. Por último, se introduce el nombre de la dimensión y se finaliza la creación.

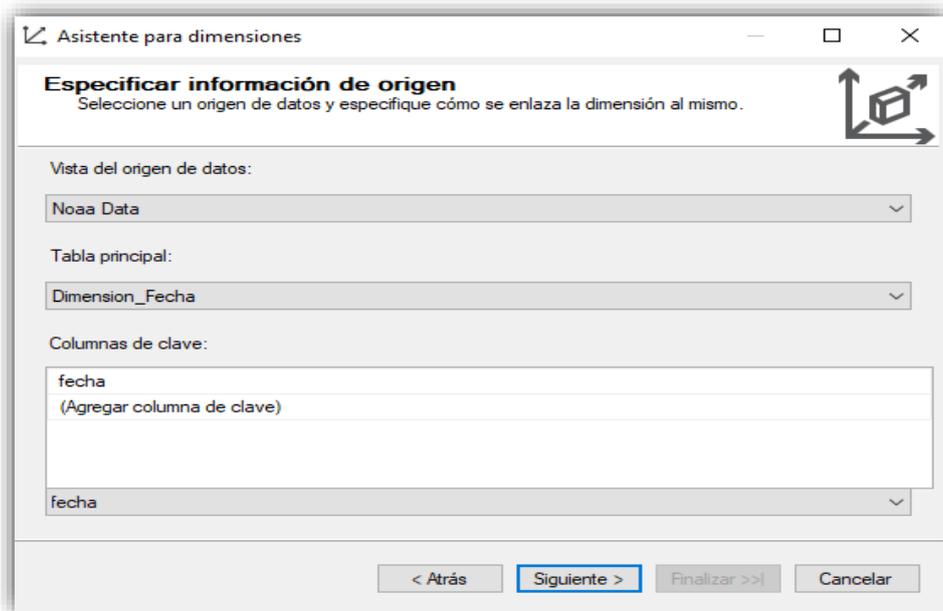


Figura 3.5: Selección de la tabla para crear la dimensión fecha

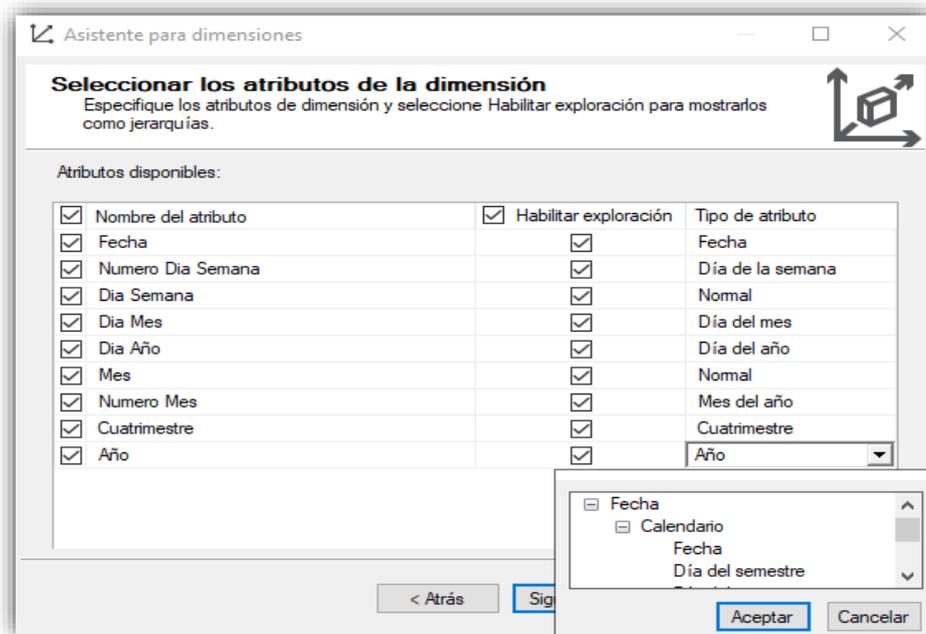


Figura 3.6: Selección de los atributos y su tipo en Dimensión Fecha

- **Dimensión Lugar:** Para crear la dimensión lugar se utiliza la tabla 'dimension_Lugar' se seleccionan todos sus atributos y se establece el tipo de atributo (Figura 3.7). Para finalizar el asistente, se escribe el nombre deseado para la dimensión.

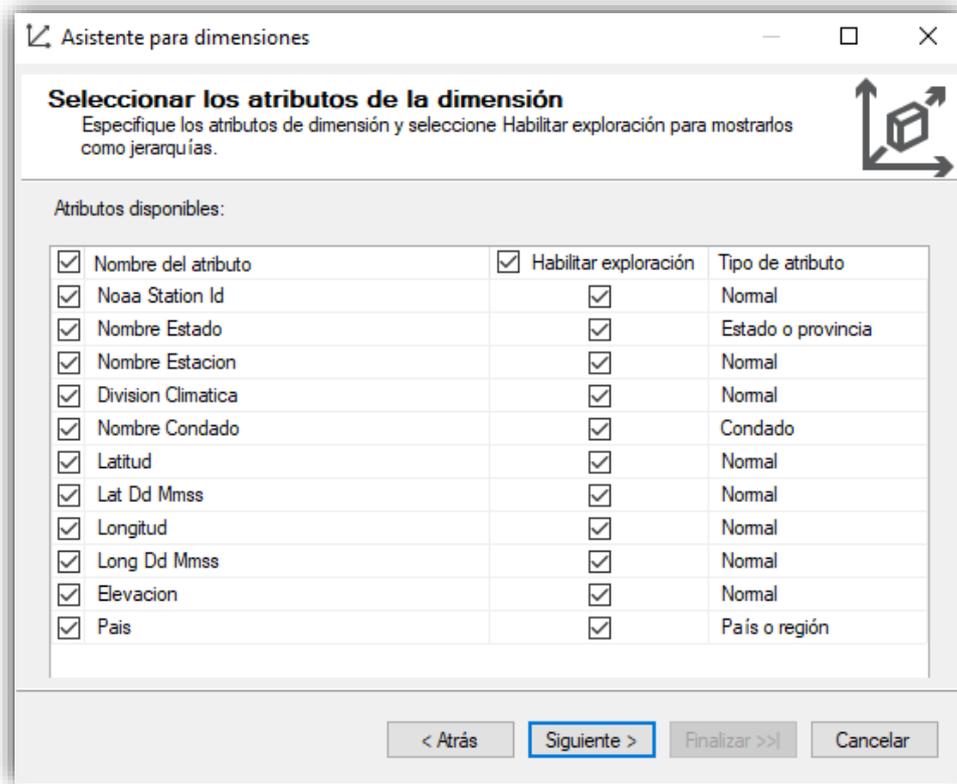


Figura 3.7: Selección atributos de la dimensión lugar

Una vez creadas las dimensiones, tenemos la posibilidad de crear jerarquías con sus atributos.

En este proyecto como se especificó en el apartado [3.2 Especificación de la configuración](#), se crearán dos jerarquías una para la 'dimensión fecha' y otra para 'dimensión lugar':

- **Jerarquía fecha:** Permite la exploración por Año-Trimestre-Mes-Día del mes (Figura 3.8).

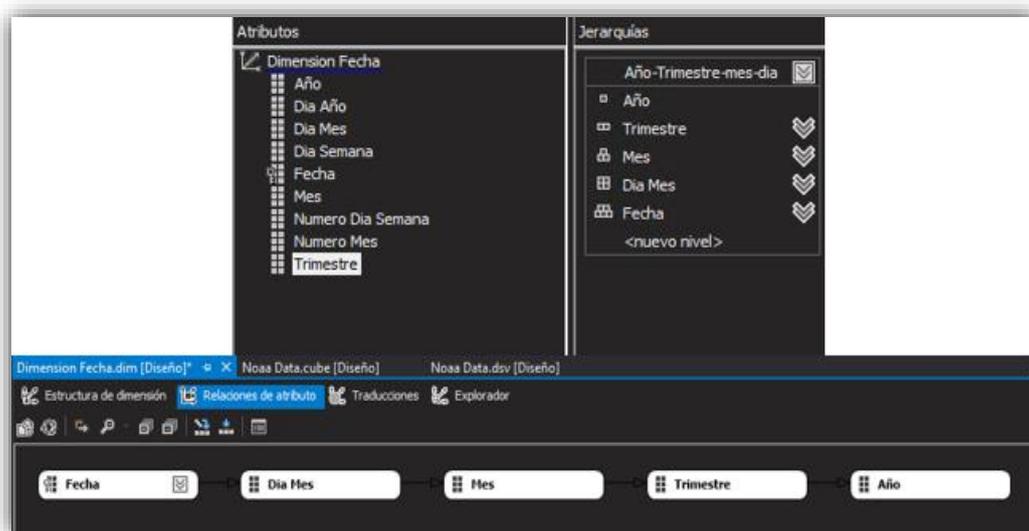


Figura 3.8: Jerarquía de fecha

Para eliminar los errores de **duplicidad de claves** de los atributos al procesar la dimensión hay que añadirle en propiedades 'Key columns': Año y atributo y en 'Name column' y 'Value column' el nombre del atributo (Figura 3.9). Añadiendo el año evitaremos la duplicidad de claves.

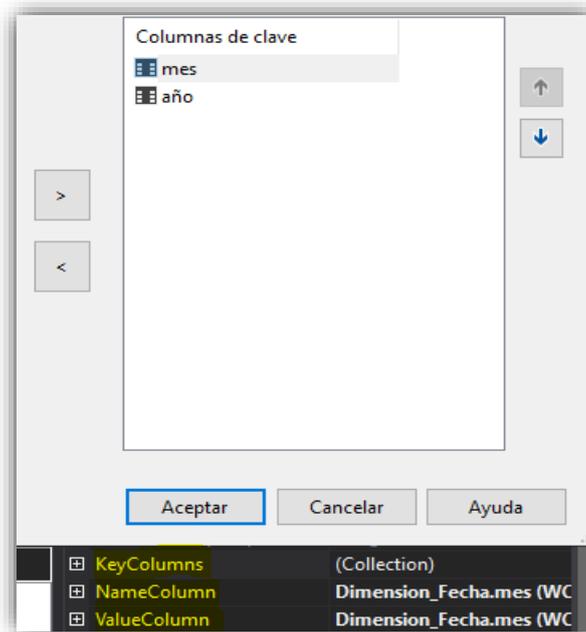
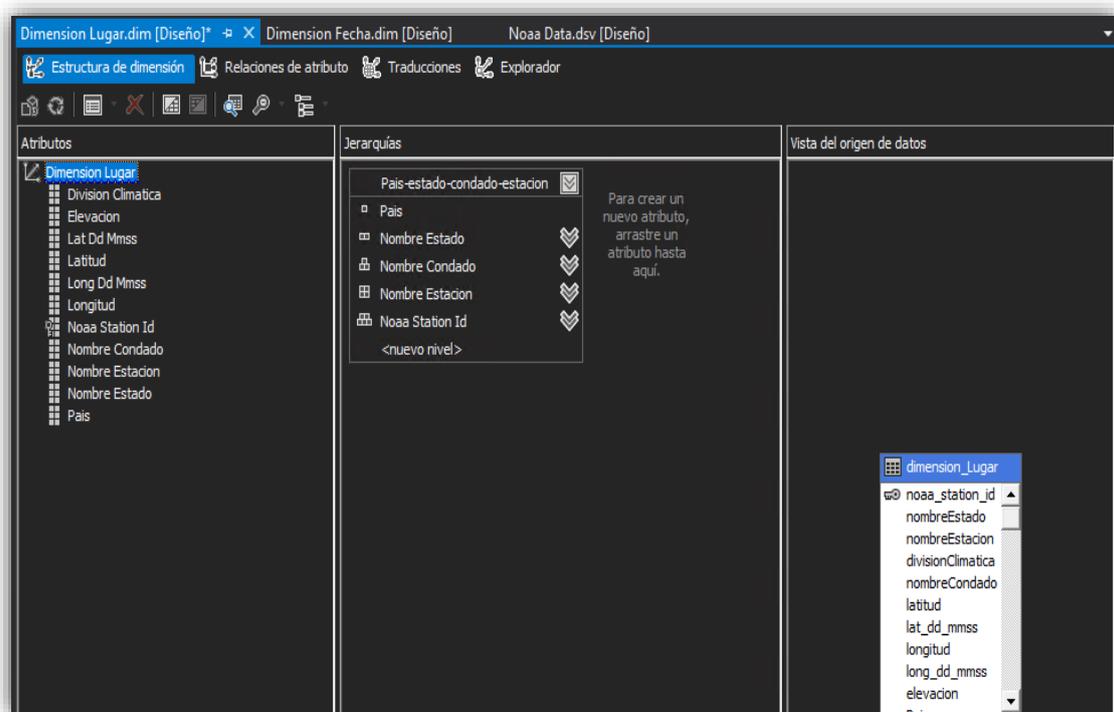


Figura 3.9: Eliminar errores duplicidad de claves al procesar la jerarquía

- **Jerarquía lugar:** Permite la exploración a través de los siguientes atributos: País - Nombre Estado - Nombre Condado - Nombre Estación (Figura 3.10).



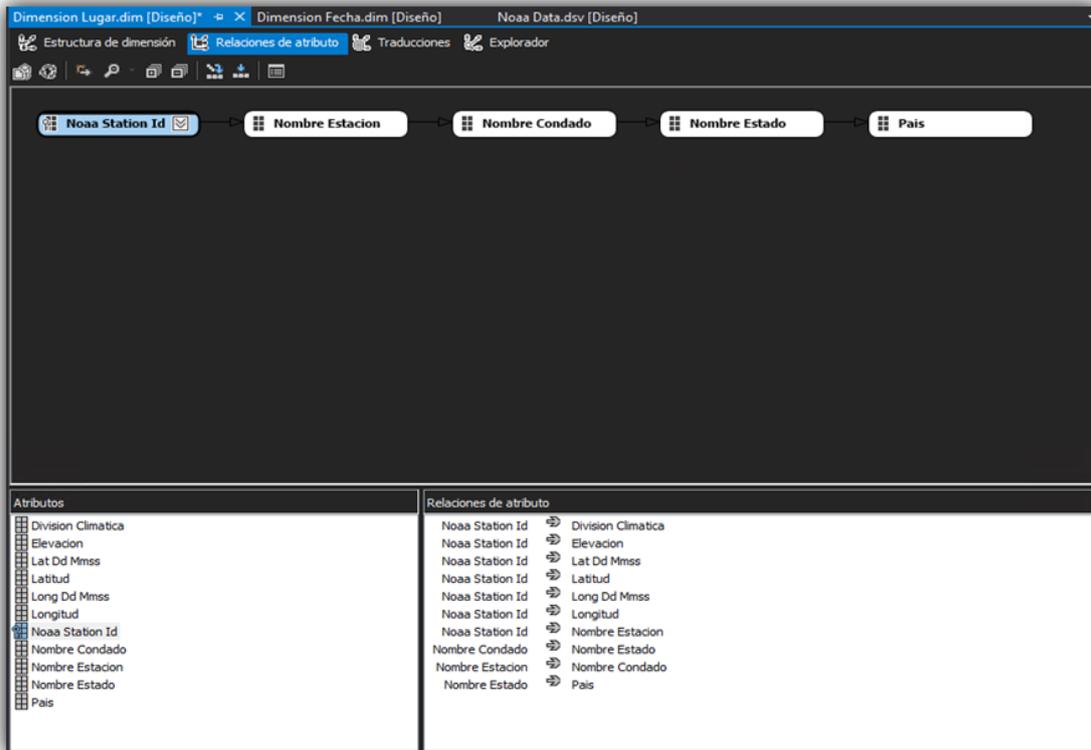


Figura 3.10: jerarquía lugar

En la figura 3.11 podemos ver como se **procesa la dimensión lugar** una vez creada la jerarquía. Para ello, haga click derecho sobre ella en el explorador de soluciones → Procesar.

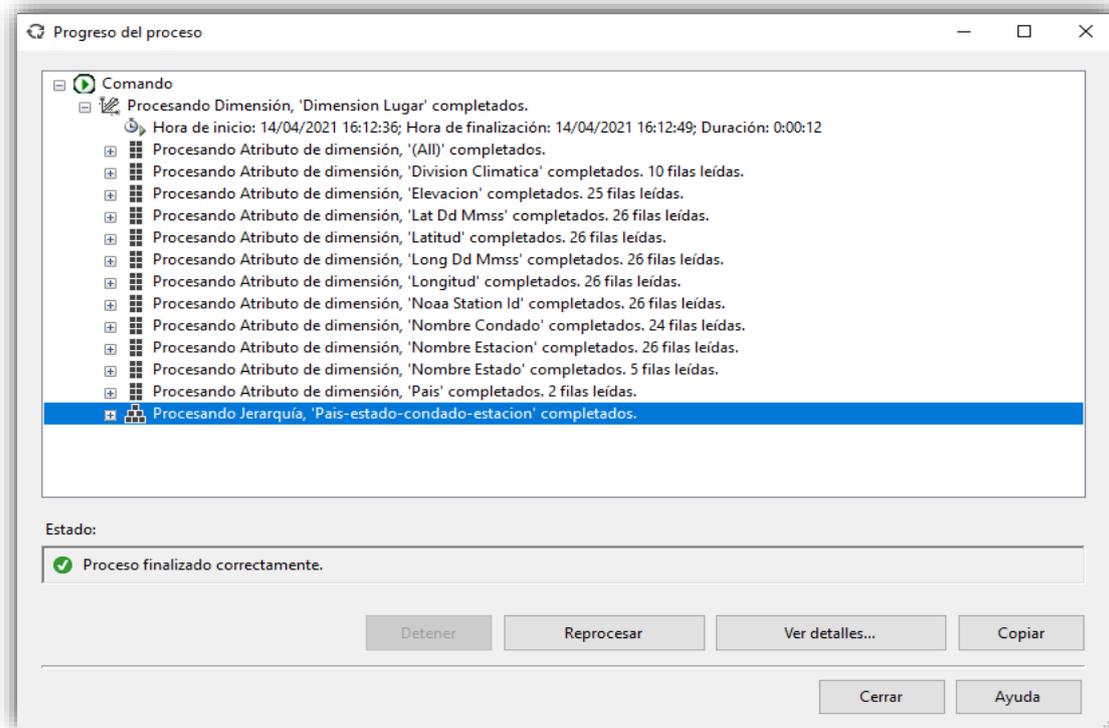


Figura 3.11: Procesado de la dimensión lugar

La figura 3.12, ilustra como navegar a través de los diferentes niveles de la jerarquía creada.

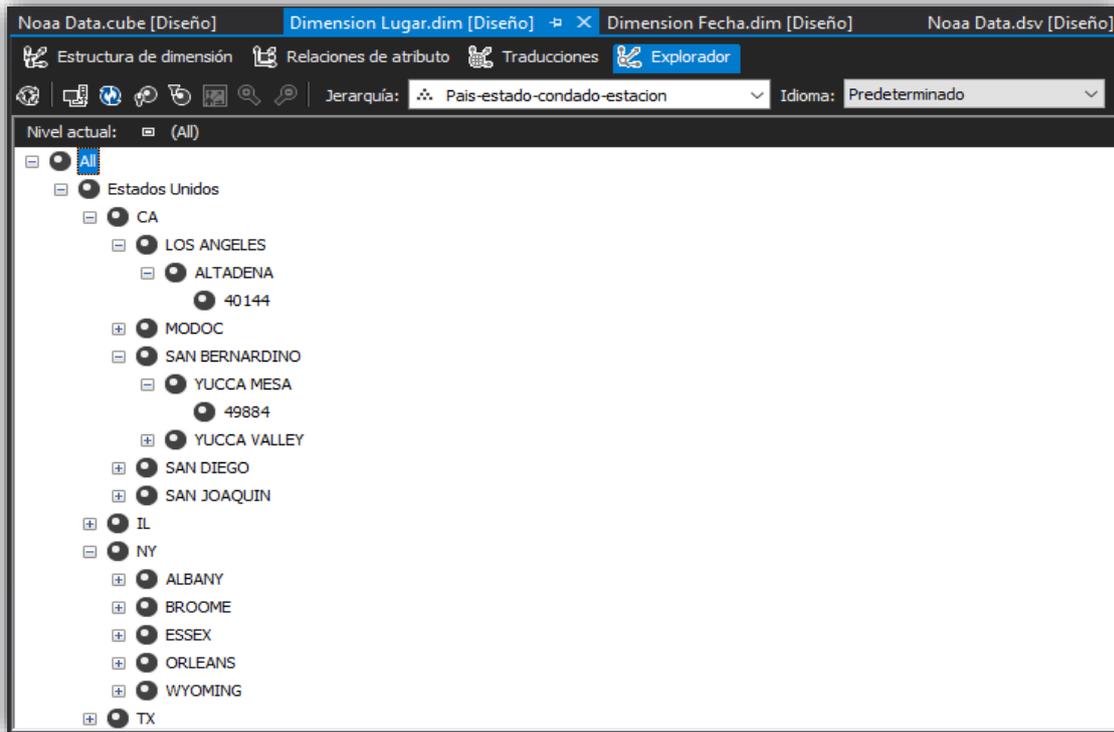


Figura 3.12: Exploración de dimensión lugar

Una vez creado el origen de datos, la vista del origen de datos y las dimensiones con sus jerarquías, se procede a **crear el cubo**. En el explorador de soluciones seleccionamos la característica ‘Cubos’ y ‘Nuevo cubo’, el asistente indica que se usarán las tablas existentes, la figura 3.13, ilustra la elección de la tabla de hechos y sus medidas.

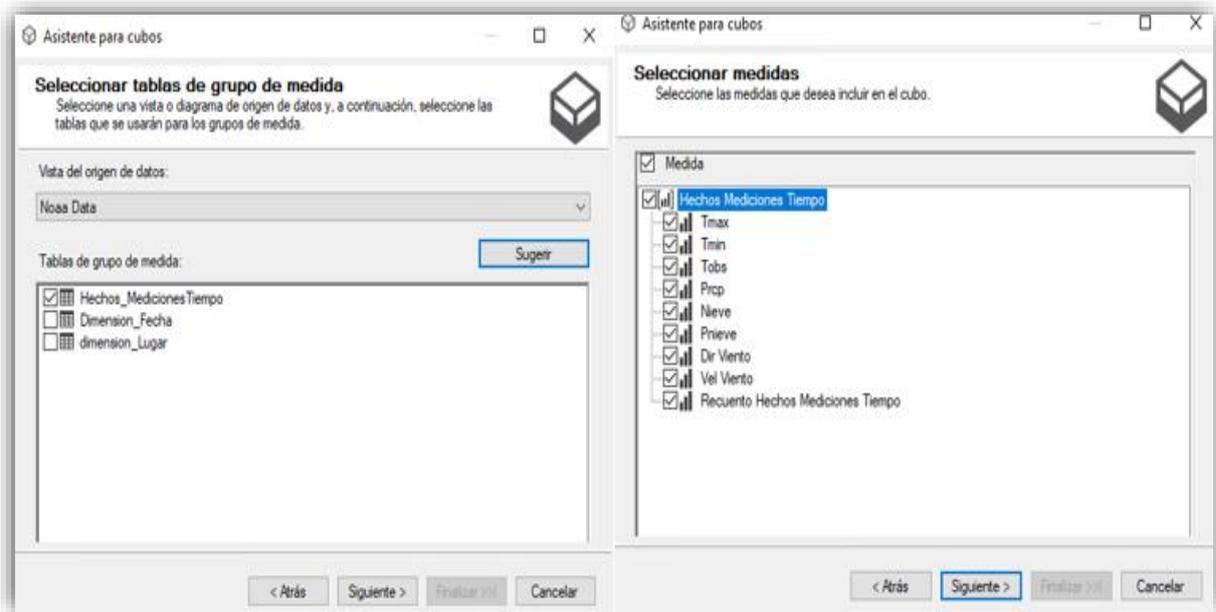


Figura 3.13: Asistente para elegir las medidas para la creación del cubo

Continuaremos con la **elección de dimensiones**, se seleccionan las dos creadas anteriormente. Finalmente, el asistente visualiza la vista previa del cubo (Figura 3.14) con los grupos de medida y dimensiones.

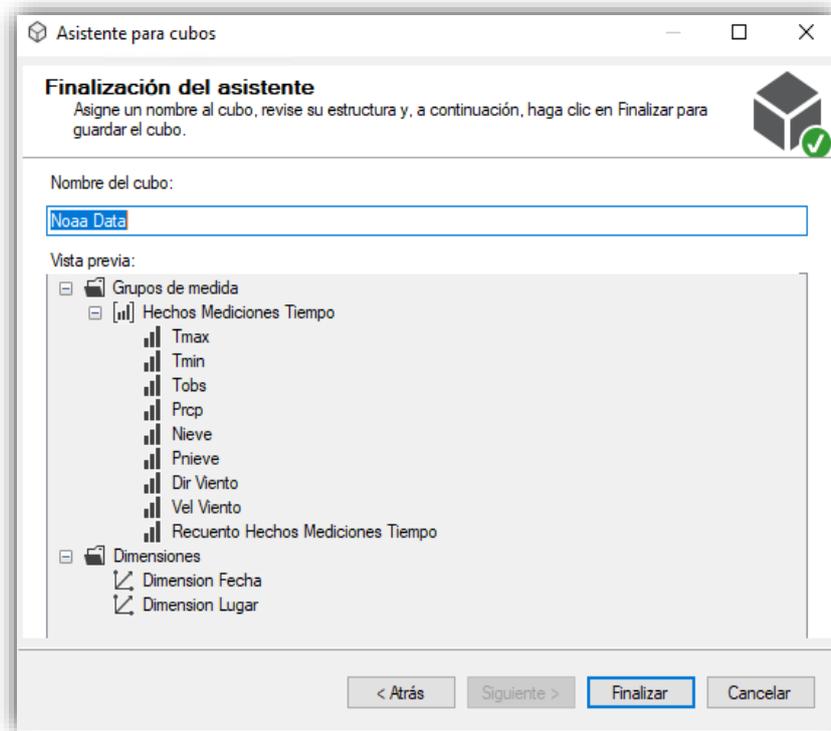


Figura 3.14: Finalización del asistente para crear el cubo

A continuación, se **implementa** el proyecto en el servidor (Figura 3.15).

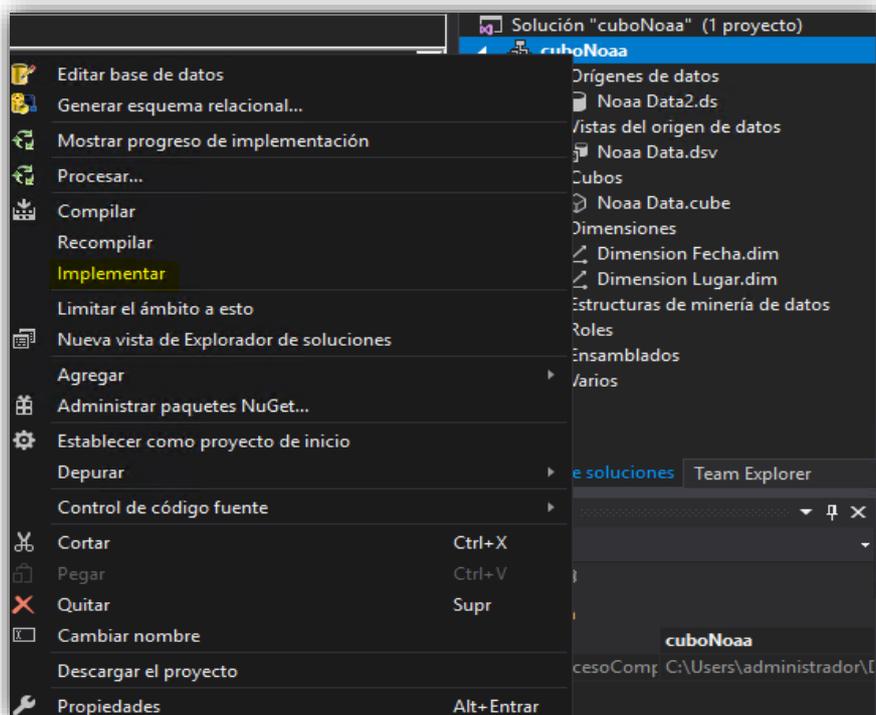


Figura 3.15: Implementación del proyecto

Tras la implementación se procesa el cubo, en la ventana emergente (Figura 3.16) podemos comprobar como el procesamiento del cubo se ha completado satisfactoriamente.

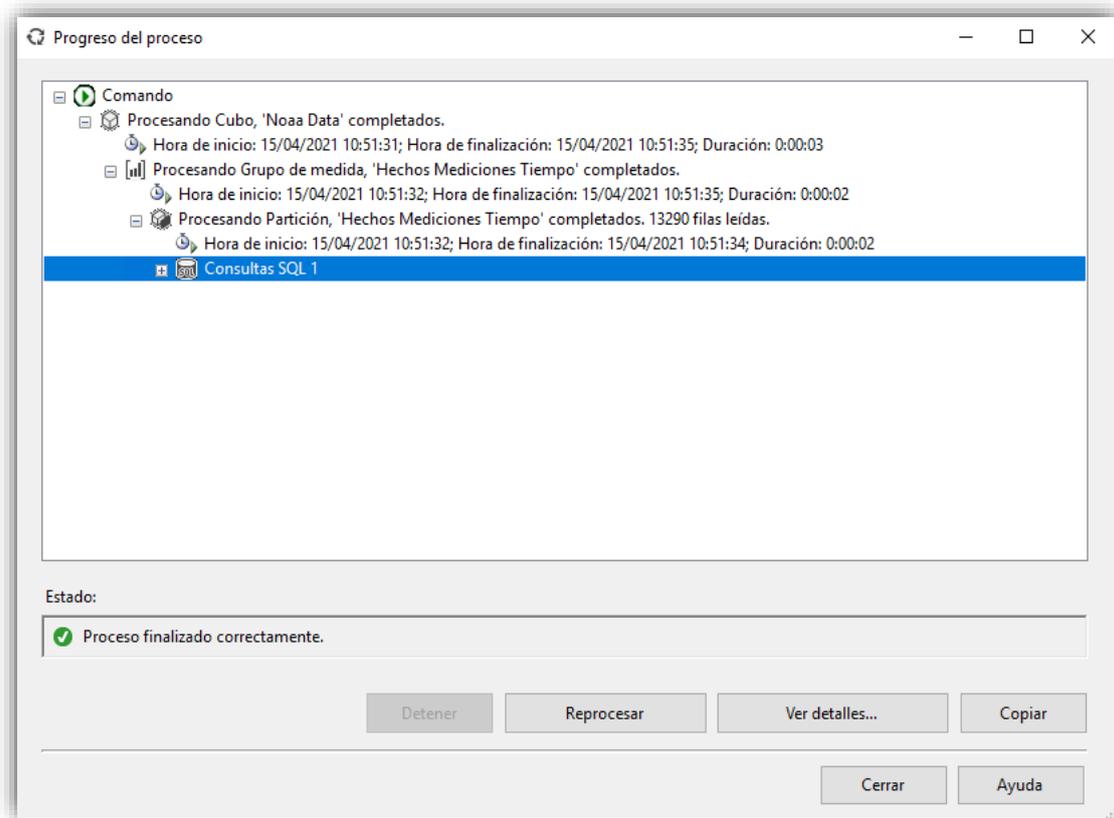


Figura 3.16: Procesamiento del cubo

La siguiente tarea es **crear medidas y columnas calculadas**. Para ello, seleccionamos en el explorador de soluciones el cubo implementado 'Noaa Data', en 'Estructura de cubo' se agrega una nueva medida, en este caso será la máxima precipitación (Figura 3.17).

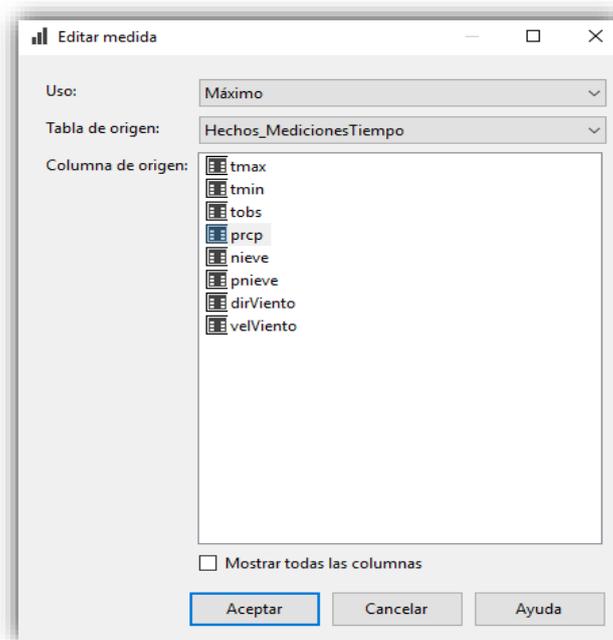


Figura 3.17: Creación de una medida (Máximo de Prcp)

Para crear las columnas calculadas en la pestaña cálculos, introducimos el nombre y la expresión deseada, la Figura 3.18, representa la conversión de la velocidad del viento de Millas por hora a Kilómetros hora, multiplicando mph por 1,609.

Mediante una consulta MDX, en el explorador podemos comprobar que la columna calculada funciona correctamente (Figura 3.19).

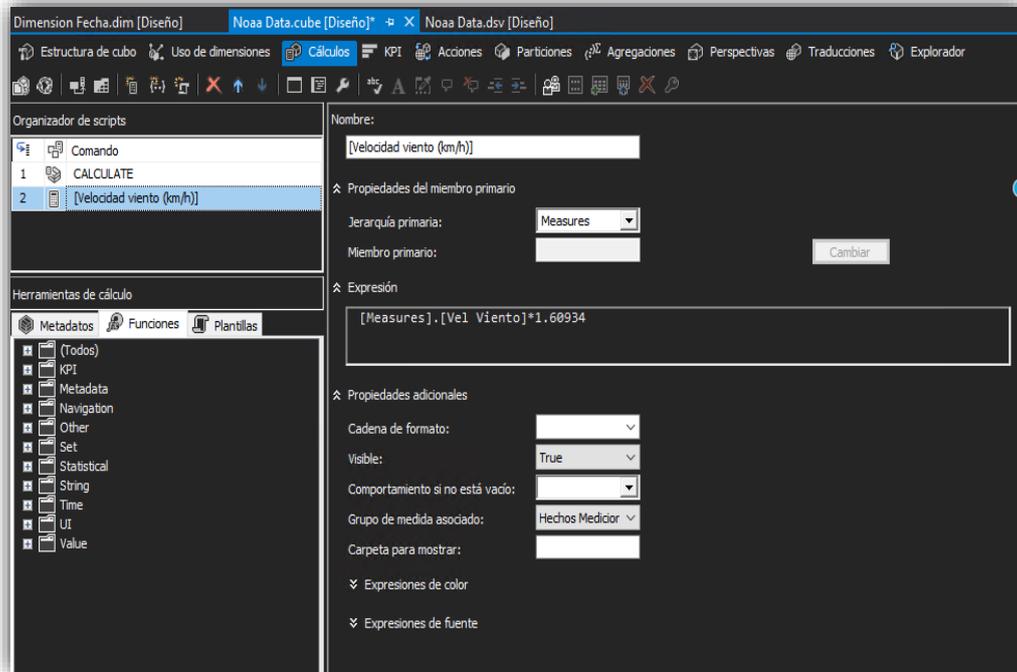


Figura 3.18: Creación columna calculada VelViento

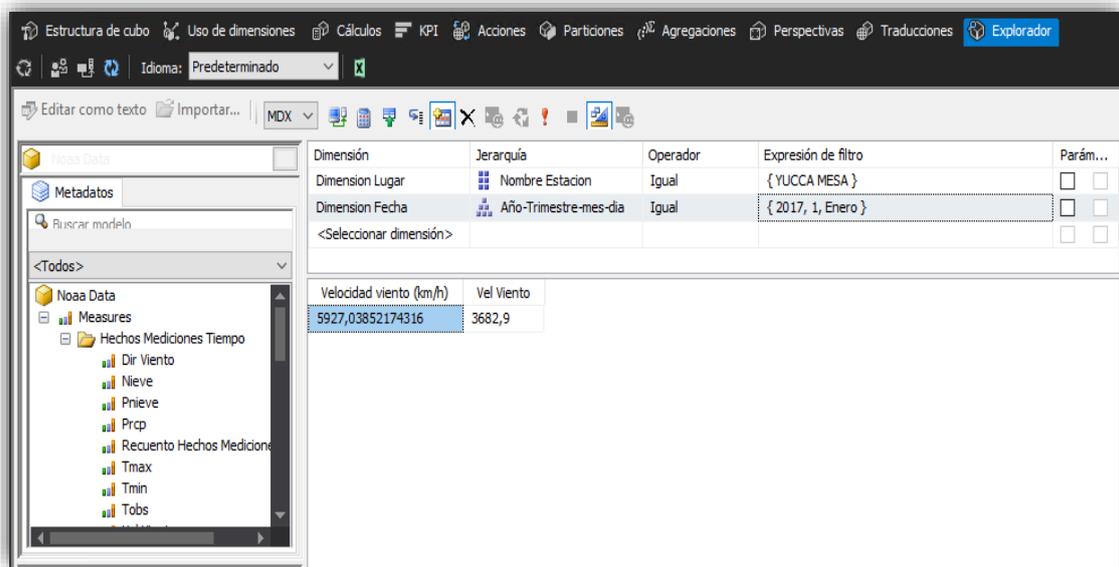


Figura 3.19: Consulta MDX para ver columna calculada

En la figura 3.20, podemos ver las nuevas medidas creadas tal y como se indicó en la especificación. Son conversiones de los atributos de las medidas usadas en EEUU a las usadas en Europa.

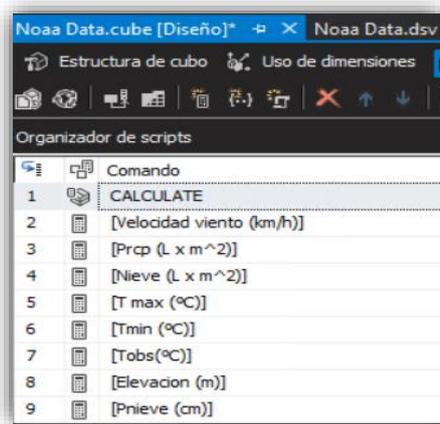


Figura 3.20: Resumen de columnas calculadas

Para diseñar las agregaciones, se abre el asistente en la pestaña ‘Agregaciones’, para optimizar el rendimiento teniendo en cuenta el almacenamiento, podemos ver como se diseñan todas las agregaciones dándole a la opción hacer click en detener (Figura 3.21). Como vemos en la gráfica alrededor de un 90% de rendimiento será un buen punto para diseñar las agregaciones porque conseguimos un alto rendimiento optimizando la memoria.

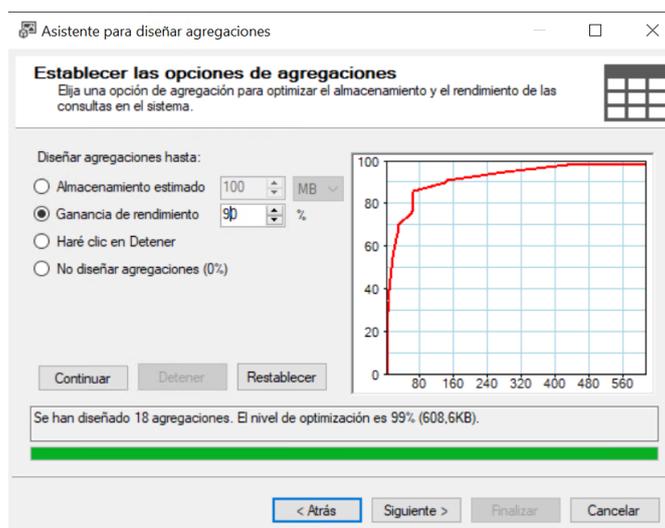


Figura 3.21: Diseño de agregaciones

3.4 Resumen del capítulo

Este capítulo, ilustra el proceso de cómo se crea el DataWarehouse en el PC ‘Puesto TI’ de la infraestructura IT. En primer lugar, se estudia la necesidad del DW en el proyecto y los conceptos necesarios para llegar a entender su implementación en el Sistema de Información de la empresa. A continuación, se explica el proceso requerido para la transformación de los datos originales y las distintas fases para la creación del cubo.

Una vez creado el cubo es necesario analizarlo para que los empleados del departamento TI ofrezcan informes interactivos e intuitivos al personal de los departamentos ‘Gerentes’ y ‘Analistas’. Para ello, se utilizará un software dedicado a Business Intelligence como es Power BI, esto se verá en detalle durante el próximo capítulo.

Palabras clave: *DataWarehouse, dimensión, ROLAP, origen de datos, MDX y diagrama en estrella.*

Capítulo 4

En este capítulo, se analizarán los datos del cubo creado en el apartado anterior usando la herramienta Power BI para crear los informes y posteriormente compartirlos con los distintos departamentos de la empresa bien, usando carpetas compartidas que se crearán en la Infraestructura IT o mediante Power BI Service, creando repositorios para la organización y publicándolos en estos o bien en una página web. Estos procesos se hacen usando como origen de datos el cubo y la Base de Datos relacional para poder crear informes más personalizados.

4.1 Business Intelligence

La inteligencia de negocios se encarga de analizar los datos transformándolos para optimizarlos y facilitar la toma de decisiones de un negocio mediante el uso de herramientas, aplicaciones, estrategias y tecnología. Su creación se remonta al año 1958 cuando el investigador *Peter Luhn* utilizó por primera vez este concepto. [3] [4] [25]

Las principales ventajas de usar BI son:

- Genera una superioridad competitiva.
- Permite crear informes personalizados, a partir de consultas de datos multidimensionales fácilmente comprensibles.
- Ahorra tiempo a la hora de resolver problemas.
- Acoplamiento entre almacenamiento y presentación de datos.
- Cuadros de mandos intuitivos, que permiten a un usuario que no posea grandes habilidades informáticas aplicar filtros para analizar los informes.
- La arquitectura DW permite a los integrantes de la infraestructura IT organizar, comprender y utilizar sus datos para la toma de decisiones trascendentales para la empresa.

Sus fases se definen como:

- **Observar:** Se observan todos los datos que tiene la empresa y pueden ser analizados.
- **Comprender:** Saber que indican los datos para hacer estudios relevantes, es decir, ¿Que ocurre y por qué?
- **Predecir:** A partir del aprendizaje de datos saber que ocurrirá, nos da la información de que le ocurriría a la empresa si se aplicaran cambios sobre ella en uno u otro sentido.
- **Colaborar:** los equipos deben colaborar entre sí para interpretar los datos de forma correcta.
- **Decidir:** Tras todo el análisis y simulaciones realizadas se decide el camino a seguir.

Gracias al uso de la inteligencia de negocio, se generan informes interactivos e intuitivos para facilitar la toma de decisiones a todas las personas de la empresa sin necesidad de conocimientos informáticos.

Existen muchos Software para la toma de decisiones, a continuación, se describen algunos de los más conocidos:

- **SAP Business Intelligence:** Ofrece informes y análisis, visualización en tiempo real de datos, soporte para empresas con gran cantidad de datos, posee aplicaciones para visualizar los datos incluyendo apps móviles. Está integrado con el paquete Office 365.
- **MicroStrategy:** Permite la creación de Dashboard muy poderosos y de alta velocidad, los análisis de datos ayudan a monitorear tendencias, las nuevas oportunidades y mejorar la productividad.
- **Oracle BI:** es una herramienta para la inteligencia de negocios proporcionada por Oracle. Esta tecnología proporciona la capacidad de crear Dashboard, paneles de control, consultas ad-hoc entre otras utilidades. Haciendo que sea una herramienta útil para empresas que necesiten analizar grandes volúmenes de datos.
- **Power BI:** Este software permite a los usuarios integrar sus apps y dar tanto informes como Dashboard en tiempo real. Será el elegido para el análisis de datos en este proyecto como se verá en el apartado [4.2 Power BI](#).

4.2 Power BI

Es un conjunto de aplicaciones de análisis de negocios que es utilizado para analizar datos y compartir la información. Power BI es una herramienta intuitiva que permite al usuario la creación de Dashboard, informes o gráficos de forma sencilla. Además, proporciona un gran número de conectores para obtener los datos, como veremos en este capítulo, donde se utilizarán los orígenes de datos 'Bases de datos SQL Server Analysis Services' y 'Base de datos SQL Server'.

La información estará disponible en los dispositivos de los usuarios o empleados y, los datos se encuentran protegidos de un extremo a otro. Mediante estas aplicaciones:

- **Power BI Desktop:** Aplicación de escritorio, que permite conectarse a los orígenes de datos que se desean modelar, transformar y visualizar. Siempre de forma local en el PC.
- **Power BI Service:** Permite visualizar y colaborar en informes en línea. También admite la creación de informes, pero con limitaciones, no cuenta con la vista de datos ni relaciones. Esta plataforma facilita el acceso a la información actualizada desde cualquier dispositivo.
- **Power BI Mobile:** Aplicación móvil con la que podemos acceder a los datos e informes actualizados en tiempo real desde cualquier lugar.

Además, hay dos tipos de cuentas:

- **Gratuita:** Permite trabajar en Power BI Desktop, publicar informes en el espacio personal de Service y visualizar estos en el móvil.
- **Pro:** Esta cuenta ofrece las mismas funciones en Power BI Desktop que la gratuita, en Power Bi Service admite crear varios workspaces, compartir y elaborar informes. En Mobile permite las Apps.

Power BI además facilita el trabajo en equipo mediante aplicaciones como Microsoft Teams, Excel u Office. Esto ayuda a la colaboración para interpretar los datos de forma correcta y poder tomar decisiones que beneficien a la empresa. [26]

A continuación, se definirán **conceptos** y elementos del software utilizado en este capítulo:

Power BI proporciona tres **vistas**:

- **Informe** : Para crear y publicar los informes y Dashboard usando los campos que se importan o especifican. Esta pestaña será la más utilizada en el proyecto. También, da la opción de usar Editor de Power Query que será utilizado para las tareas de limpieza y transformación de datos.
- **Datos** : Permite la visualización de los datos importados o creados, desde esta pestaña. También está permitido importar, crear y publicar datos, así como la agregación de nuevas columnas, tablas y medidas.
- **Modelo** : La funcionalidad de esta vista es la creación de un diagrama de las tablas del conjunto de datos con el que se trabaja, administrar las relaciones y crear jerarquías.

Query Editor es un motor que tiene Power BI que maneja lenguaje EMI. Con el objetivo de ayudar a conectarnos y transformar los datos.

Orígenes de datos: Son las fuentes desde donde se obtendrá la información con la que realizar las tareas de limpieza y transformación de datos, para el análisis y creación de informes.

Tareas de limpieza y transformación de datos: Su finalidad es conseguir datos íntegros y consistentes, la integración se realiza durante la carga de datos.

Las tareas de limpieza y transformación están encaminadas a eliminar los datos incorrectos o anómalos, valores duplicados, filas vacías o tratar las columnas donde se queden huecos sin datos.

Existen dos procesos para realizar estas tareas:

- **Transformación:** Crear atributos derivados o cambiar el tipo de un atributo.
- **Integración:** Busca separar un objeto en dos o más o, por el contrario, unificar varios objetos en uno.

Funciones DAX: Las expresiones de análisis de datos o DAX, son un lenguaje que recopila funciones y constantes que se pueden usar en una fórmula o expresión para calcular y devolver uno o varios valores. Se aplican en columnas calculadas, medidas y seguridad a nivel de fila. [5]

Dashboard o cuadro de mandos, visualiza los datos más importantes, permitiendo su análisis y detección de problemas, aumentando la efectividad a la hora de tomar decisiones

4.3 Especificación de la configuración

En esta sección se especificarán los detalles de transformaciones de datos realizadas en el informe con origen de datos la Base de Datos relacional, ya que los cubos no permiten transformaciones.

También se mostrará la estructura planificada para crear los dashboards y la interacción entre estos.



En primer lugar, veremos las transformaciones de datos necesarias:

Se realizarán pocas modificaciones debido a que los datos no necesitan grandes cambios.

- **Dimensión_Lugar**, se van a realizar las siguientes tareas de transformación de datos:
 - Reemplazar los símbolos '*' por '°' en los campos latitud y longitud.
 - Nueva columna que representa la elevación en metros, se crea a partir de la columna elevación cuyas medidas son en pies.
- **Dimension_Fecha**, se cambian los atributos que estaban establecidos como tipos nominales cuando eran valores de tipo numérico.
 - Se integran datos combinando filas, por ejemplo, las columnas día y mes en una única columna con el siguiente formato: (día, mes).
- **Hechos_MedicionesTiempo**
 - Los valores perdidos fecha que se reemplazaron 'null' por '31/12/2999', para establecer la relación de la tabla de hechos con la dimensión fecha, en el modelado de datos, se ha decidido eliminarlos, para evitar problemas de ruido en los informes. Debido a que sin la fecha no eran relevantes para los informes.
 - Se reemplazan los valores null de las columnas prcp, nieve, pnieve, dirViento y velViento por 0.
 - Se crean nuevas columnas con las temperaturas en grados, precipitaciones y nieve en litros por metro cuadrado, profundidad nieve en metros y velocidad viento en km/h.
 - Se encuentra un valor anómalo o perdido, es una nevada de más de 10 metros, se decide eliminar el registro de la tabla.

En cuanto a la interfaz se compone de un menú principal, en el que se permite la navegación a la página web de Noaa, ver el mapa de estaciones o dos submenús con páginas de informes, desde los que se puede interaccionar con el Dashboard deseado. Toda la navegación entre páginas se realiza por medio de iconos con funcionalidad de botones.

A continuación, se especificarán las plantillas (Tabla 4.1) e iconos (Tabla 4.2) utilizados para crear los informes.

¿Dónde se utiliza la plantilla?	Plantilla
<p>Menú principal</p>	

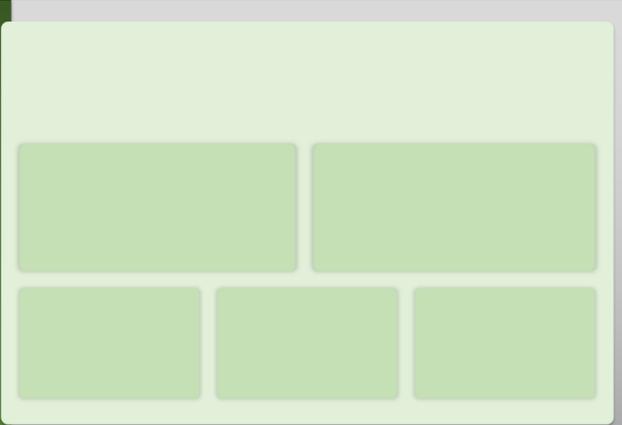
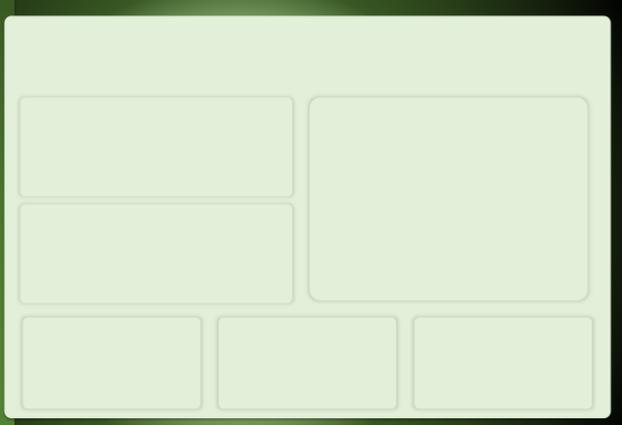
<p>Submenú informes básicos e informes</p>	
<p>Submenú informes avanzados</p>	

Tabla 4.1: Plantillas para informes de Power BI

¿Qué hace el icono?	Icono
<p>Permite volver al menú principal</p>	
<p>Navega a la página web de NOAA</p>	
<p>Muestra el mapa de estaciones de EEUU</p>	
<p>Muestra el submenú de dashboards básicos</p>	
<p>Muestra el submenú con dashboards más complejos</p>	

Retorna a la página anterior	
Muestra información de la página actual	
Muestra el informe de T ^a Máximas	
Muestra el informe de T ^a Mínimas	
Muestra el informe de precipitaciones	
Muestra el informe de nevadas	
Muestra el informe de vientos	
Muestra el informe de elementos más influyentes	
Muestra la página de preguntas y respuestas	
Muestra un esquema de jerarquías	
Muestra el informe del top N de precipitaciones por estaciones	
Muestra el informe del top N de nevadas por estaciones	
Muestra un informe de todas las Temperaturas	

Tabla 4.2: Iconos utilizados en los informes

Además, en el gráfico de dispersión del Dashboard de vientos, se realiza el siguiente formato condicional para colorear los datos según la velocidad del viento:

Velocidad del viento (KM/H)	Color burbuja
De 0 hasta 25	Verde
De 25 hasta 50	Amarillo
Más de 50	Rojo

Tabla 4.3: Color de datos en el gráfico de dispersión de vientos

Para compartir los datos se define un directorio de carpetas compartidas con la siguiente estructura:

Almacén empresa	Departamento	Departamento TI	00 - Compartida		
			01 - SMM - TI		
			02 - Jose-TI		
		Departamento Gerentes	00 - Compartida		
			03 - Antonio - Ger		
			04 - Laura - Ger		
		Departamento Analistas	00 - Compartida		
			05 - Luis -Ana		
			06 - Paco - Ana		
	Personal		01 - SMM - TI		
			02 - Jose-TI		
			03 - Antonio - Ger		
			04 - Laura - Ger		
			05 - Luis -Ana		
			06 - Paco - Ana		
			Público		01 - SMM - TI
					02 - Jose-TI
					03 - Antonio - Ger
04 - Laura - Ger					
05 - Luis -Ana					
06 - Paco - Ana					

Tabla 4.4: Estructura del directorio y subdirectorios compartido creado en DC

4.4 Detalles de la implementación

En primer lugar, se realiza la conexión al origen de datos (Figura 4.1).

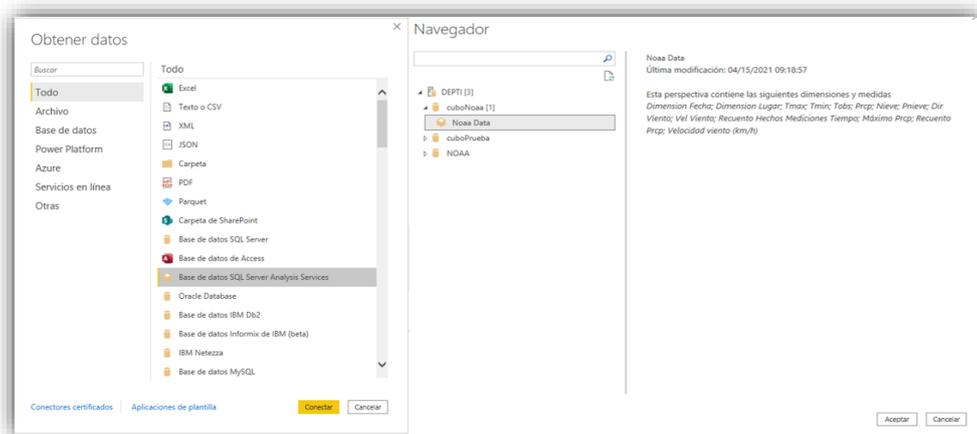


Figura 4.1: Conexión al origen de datos

Para realizar las transformaciones de datos, se usa el Editor de Power Query que se encuentra en la vista 'Informe'. En la sección Consultas → 'Transformar datos'.

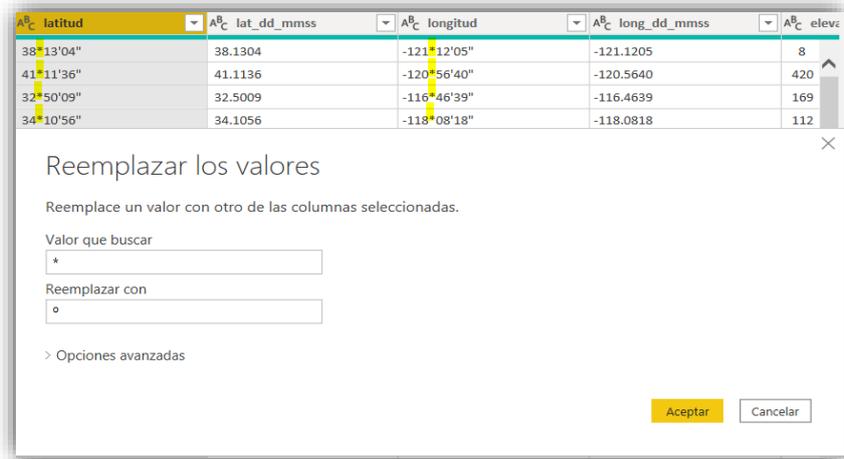


Figura 4.2: Reemplazar '*' por '°'

La figura 4.3 ilustra cómo se crea una nueva columna, a la derecha de la imagen se pueden ver los pasos que se aplican en Query Editor.

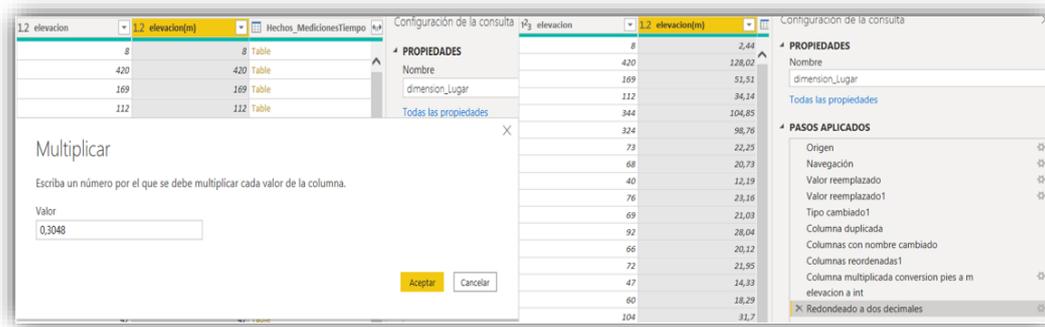


Figura 4.3: Crear columna elevación en metros

Tras las tareas de limpieza y transformación de datos obtenemos el siguiente modelo (Figura 4.3):

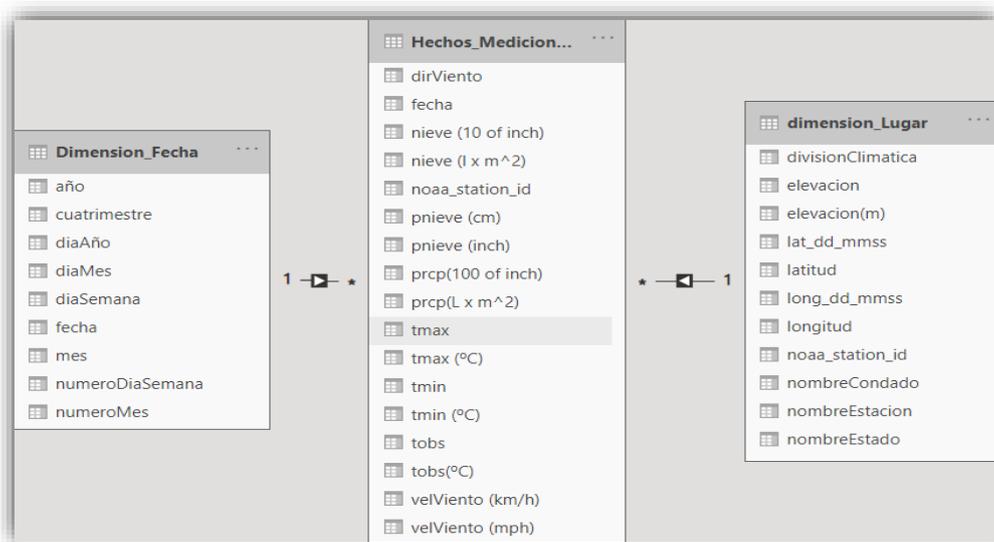


Figura 4.4: Modelo de datos del origen BD relacional tras las tareas de limpieza y transformación

Una vez importados los datos y transformados se procede a crear gráficos, en la vista de informes

tenemos todas los objetos visuales ofrecidos por Power BI para crear informes e interactuar con ellos de una forma sencilla. La primera representación del conjunto de datos será un mapa con la ubicación de las estaciones, la circunferencia alrededor de la ubicación de la estación viene determinada por el número de apariciones en la tabla de hechos (Figura 4.6). Se ha creado un nuevo campo en la dimensión Lugar, Ubicación que combina los campos nombreCondado y nombreEstado (Figura 4.5).

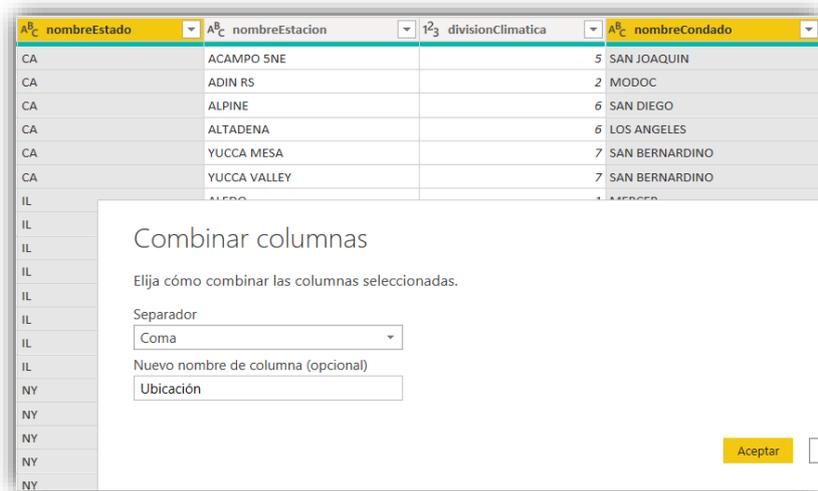


Figura 4.5: Combinar dos columnas

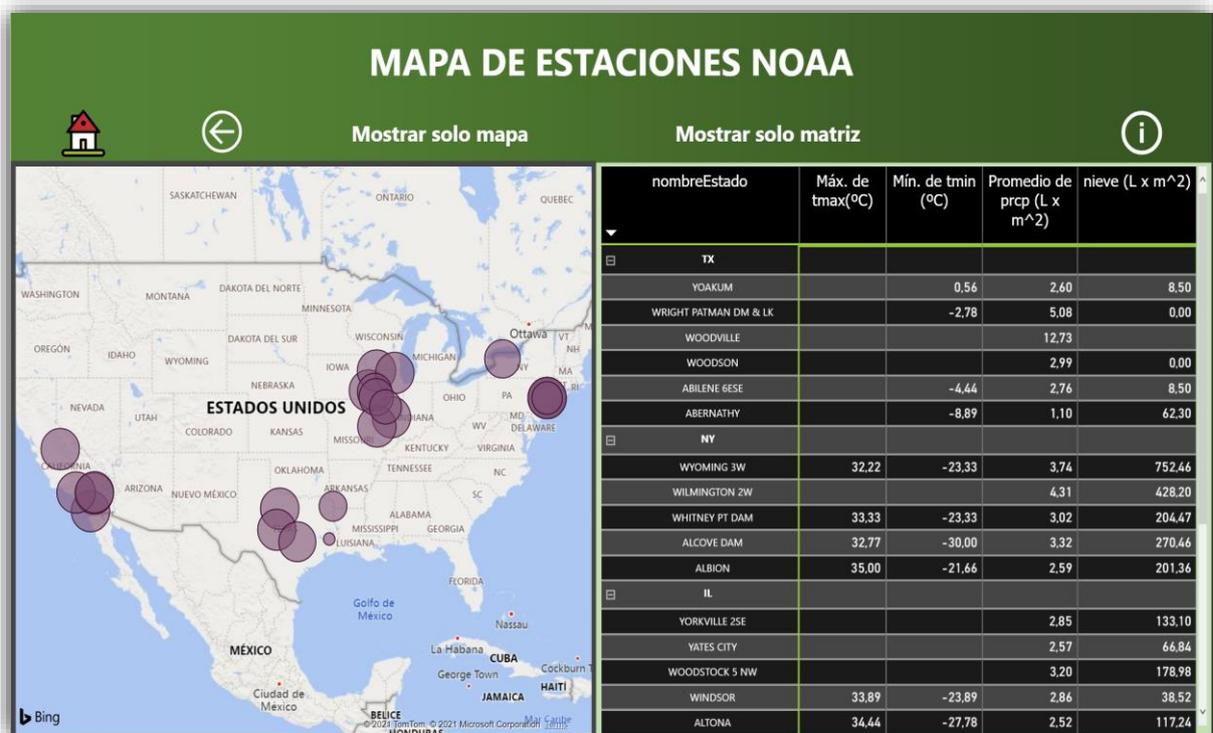


Figura 4.6: Mapa con matriz

También se crean informes interactivos usando segmentadores de datos y varios tipos de visualizaciones, como se observa en la siguiente figura:

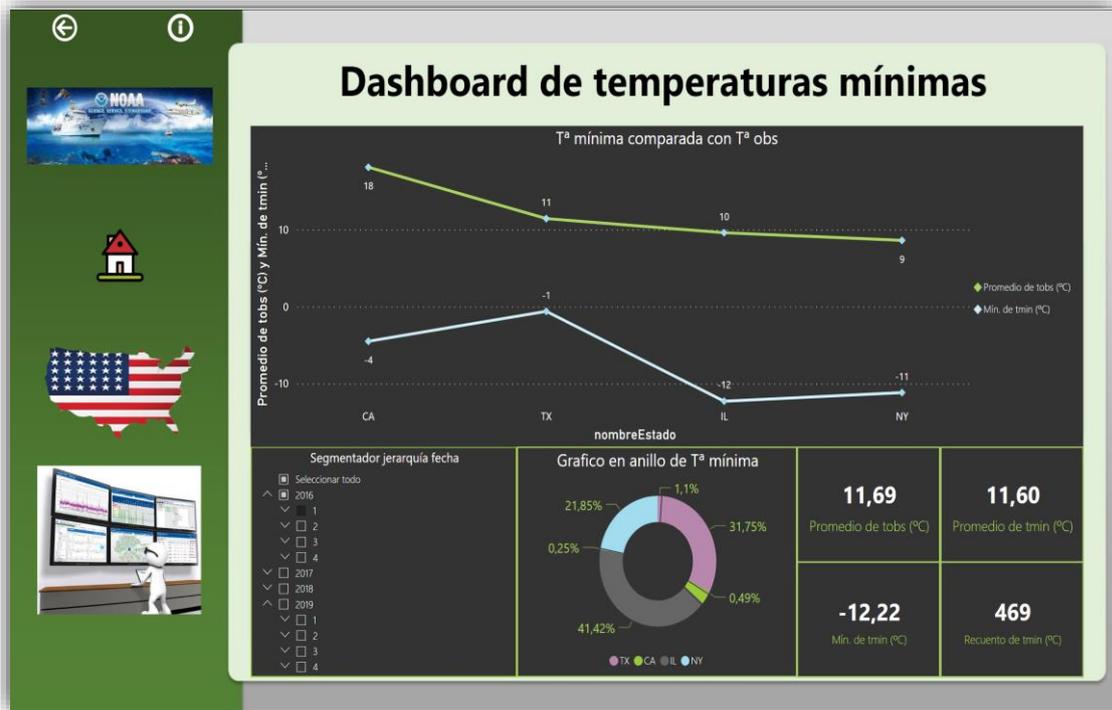


Figura 4.7: Dashboards de Tª mínima

Como se mencionó en el apartado anterior, para el informe de vientos se ha creado un formato condicional para colorear los elementos del gráfico de dispersión (Figura 4.8), como resultado final se obtiene figura 4.9.

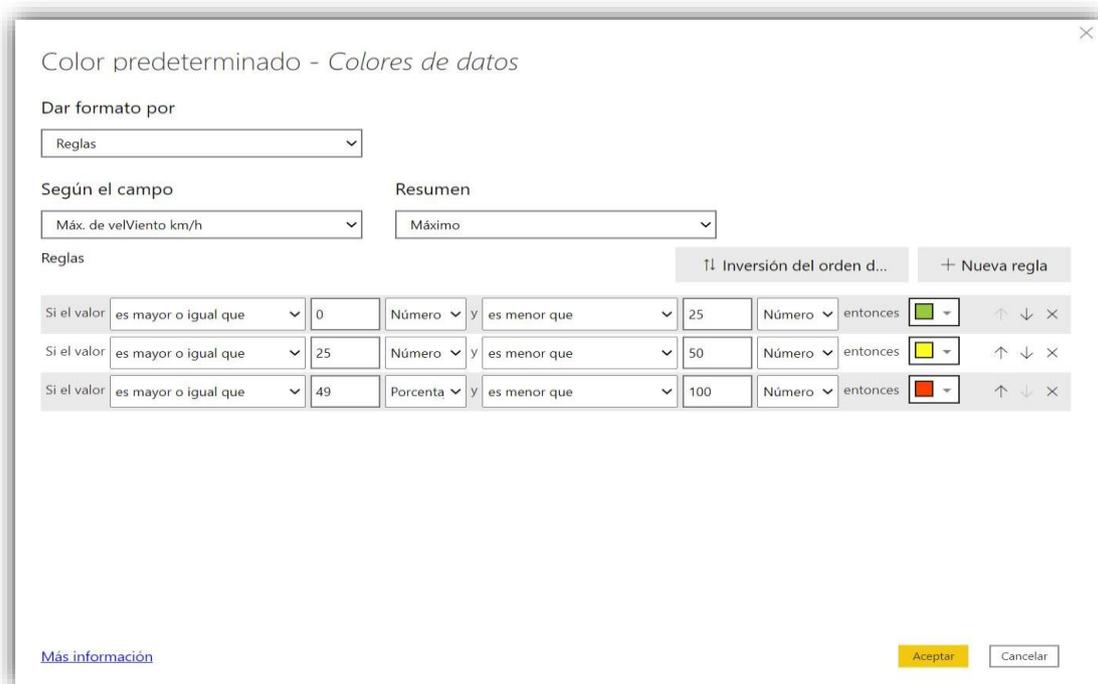


Figura 4.8: Coloreado de datos

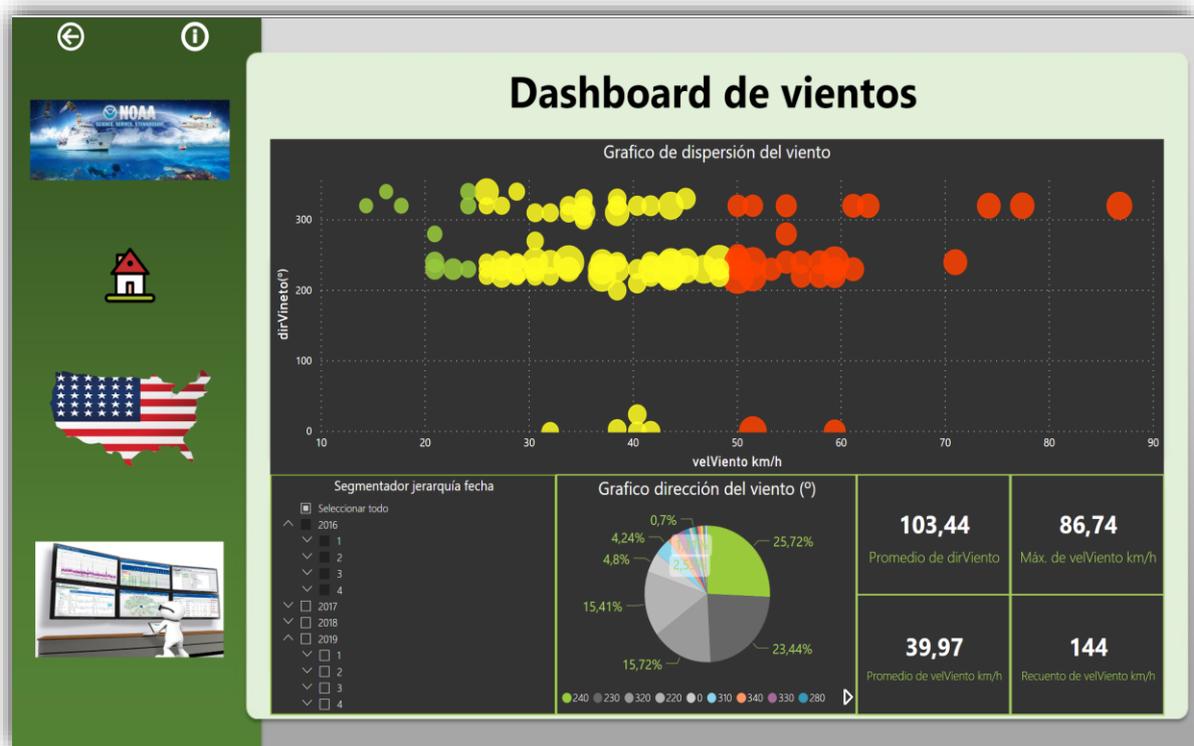


Figura 4.9: Dashboard de vientos

En la figura 4.10 se muestra un Dashboard que permite interactuar con el Top 5,10,15 o 20 de las estaciones donde más ha llovido en el periodo de tiempo que indiquemos por medio del segmentador. Para ello, se necesita crear una tabla que tenga la Posición (Top 5, Top 10...) y el valor (5,10...). Además, se agregan dos nuevas medidas la primera:

```
1 SeleccionTop = if(HASONEVALUE('Tabla_Posición'[Posición]), Values('Tabla_Posición'[Valor]), 25)
```

Esta medida se encarga de seleccionar el valor deseado de la tabla creada, el valor por defecto es 25 el total de todas las estaciones.

La segunda medida es 'Total lluvia por estación' en Var estación se guardarán los nombres de las estaciones, en Var posición se guarda la posición seleccionada en el segmentador de datos donde se encuentra 'SeleccionTop' (Top 5, Top 10...) para devolver el total de lluvias para esos N campos ordenados de mayor a menor.

```
1 Total Lluvia por estación = Var estación = Values(dimension_Lugar[nombreEstacion])
2 Var posicion = 'Tabla_Posición'[SeleccionTop]
3 Return CALCULATE([Total Lluvia], TOPN (posicion,ALL(dimension Lugar[nombreEstacion]), [Total Lluvia], desc), estación)
```

El resultado de aplicar estas funciones DAX en visualizaciones se muestra en la figura 4.10. Compuesta por un segmentador que permite filtrar por fecha, una matriz que muestra el top N de estaciones ordenadas de mayor a menor cantidad de precipitaciones en el periodo seleccionado, el segmentador para elegir la cantidad de estaciones mostradas y un gráfico de barras con la cantidad de lluvia por estación.



Figura 4.10: Top N estaciones por lluvia

El Anexo [Informes Power BI](#), ilustra un ejemplo de navegación por la interfaz donde se visualiza el resto de Dashboard creados para el proyecto.

4.4.1 Acceso a los recursos de la infraestructura

Para finalizar el apartado, se necesita compartir los informes, para ello se utiliza Power BI Service y el directorio de carpetas compartidas en la infraestructura IT creada.

- Directorio de carpetas compartidas, es necesaria la creación del directorio ‘AlmacenEmpresa’ alojado en el Controlador de Dominio, este recurso se comparte con los usuarios de la empresa, como se indica en la figura 4.11:

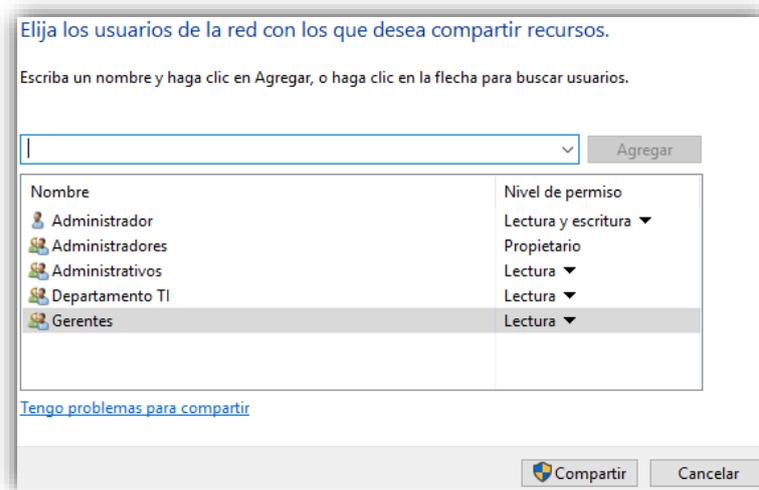


Figura 4.11: Compartir carpetas compartidas con los grupos de usuarios

Dentro de esta carpeta se crean otras tres carpetas de recursos, una personal donde cada empleado podrá guardar los archivos que no quiere compartir con ningún otro compañero.

Otra pública para que cualquier persona pueda subir archivos a esa carpeta y los miembros de la empresa tengan acceso a su contenido.

Y por último, una carpeta departamento, que estará compuesta por un conjunto de subcarpetas para cada departamento, es decir, tres carpetas una para Departamento TI, Gerentes y Analistas, dentro de cada carpeta de departamento contendrá una carpeta para cada usuario del departamento (los demás usuarios del departamento solo tendrán permisos de lectura) y una carpeta compartida a todos los usuarios del departamento en la que tendrán permisos de lectura y escritura los componentes de ese departamento. La estructura de carpetas se especificó en la tabla 4.4 del apartado [4.3 Especificación de la configuración](#).

Se hace de la siguiente manera:

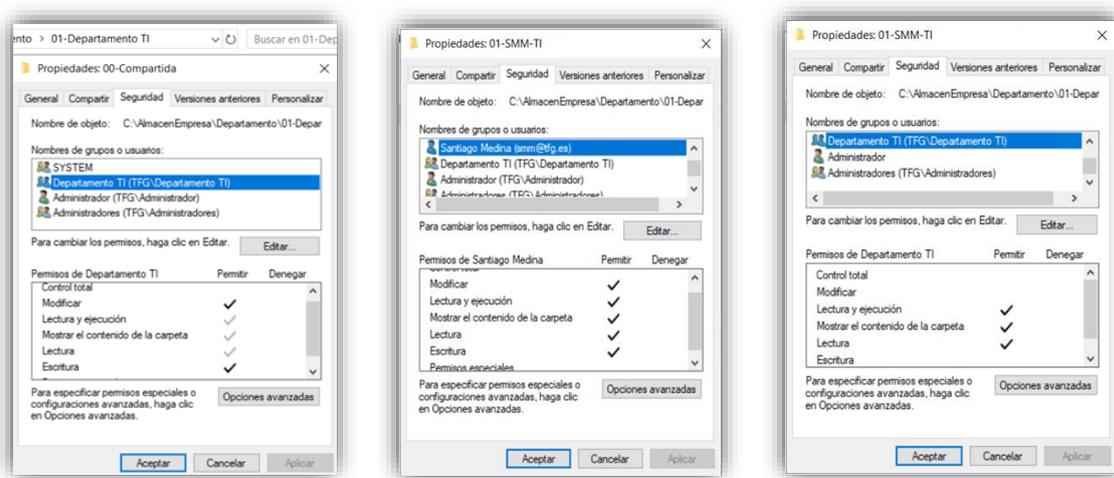


Figura 4.12: Asignación de permisos de carpetas por departamento y personales

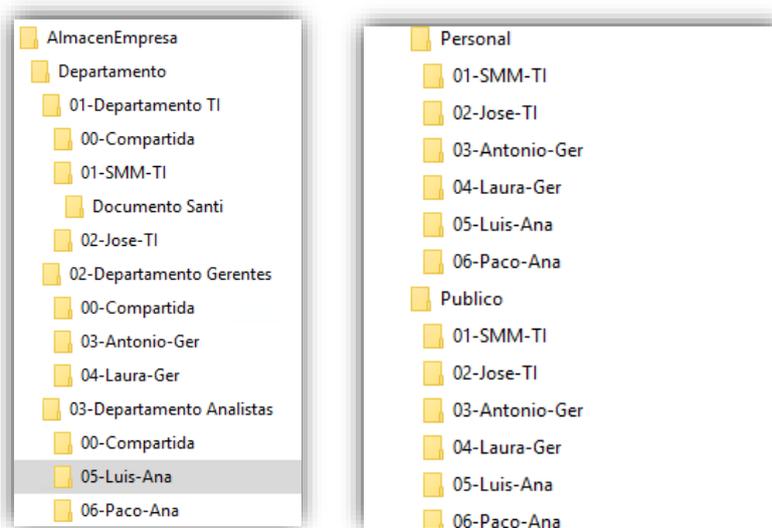


Figura 4.13: Directorio de carpetas compartidas creado

Una vez permitido el acceso a los usuarios de la empresa a sus correspondientes carpetas compartidas se crea en el escritorio de cada PC un acceso directo a la carpeta del recurso compartido que estará en el controlador de dominio y será visible a través de la red interna de la infraestructura IT.

Para conectar, siempre que se use la VPN o un puesto de la organización, se escribe en el explorador de archivos: ‘\\192.168.1.1\AlmacenEmpresa’.

En el siguiente conjunto de imágenes (figura 4.14) se verá un ejemplo, donde se verá como visualiza las carpetas el usuario Luis del departamento Analistas, así como el intento de editar un archivo de texto del que solo tiene permisos de lectura.

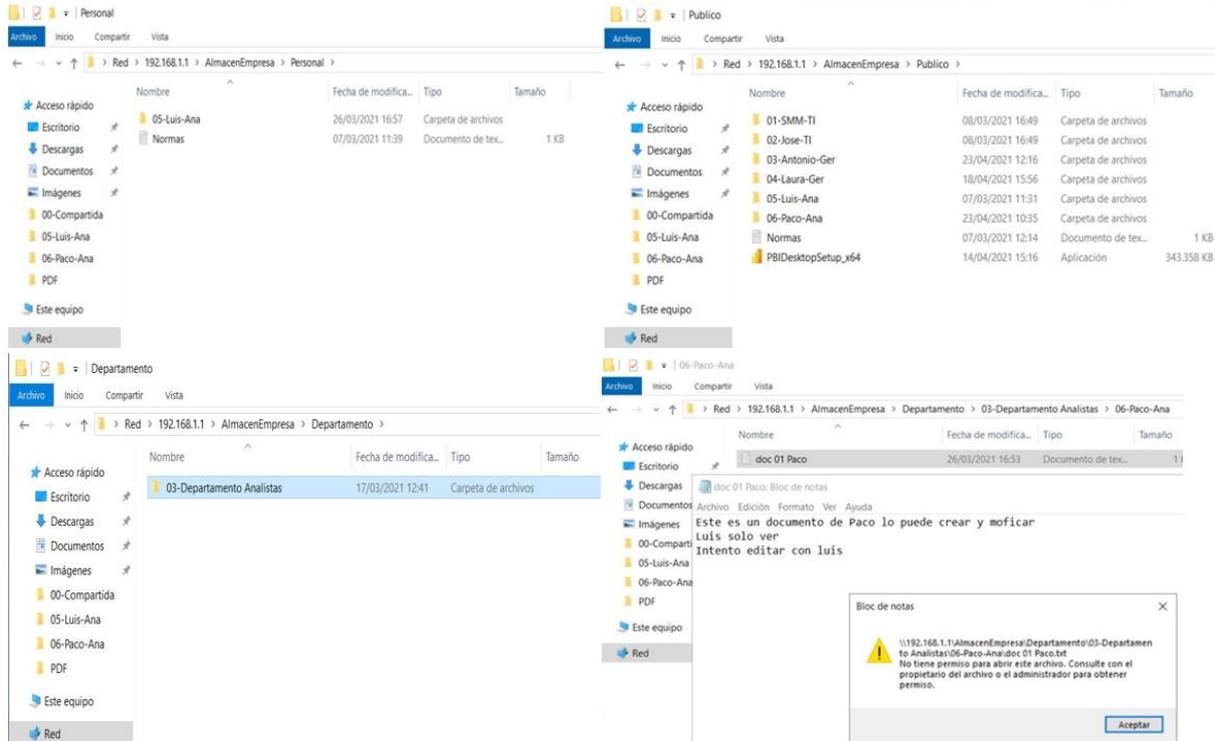


Figura 4.14: Vista del directorio compartido a través del usuario Luis

Ahora, en la figura 4.15, se comparte un archivo al que solo Luis puede acceder, deshabilitando los permisos a todos los usuarios de la carpeta de informes Power BI y solo concediéndole a él permisos de lectura y ejecución.

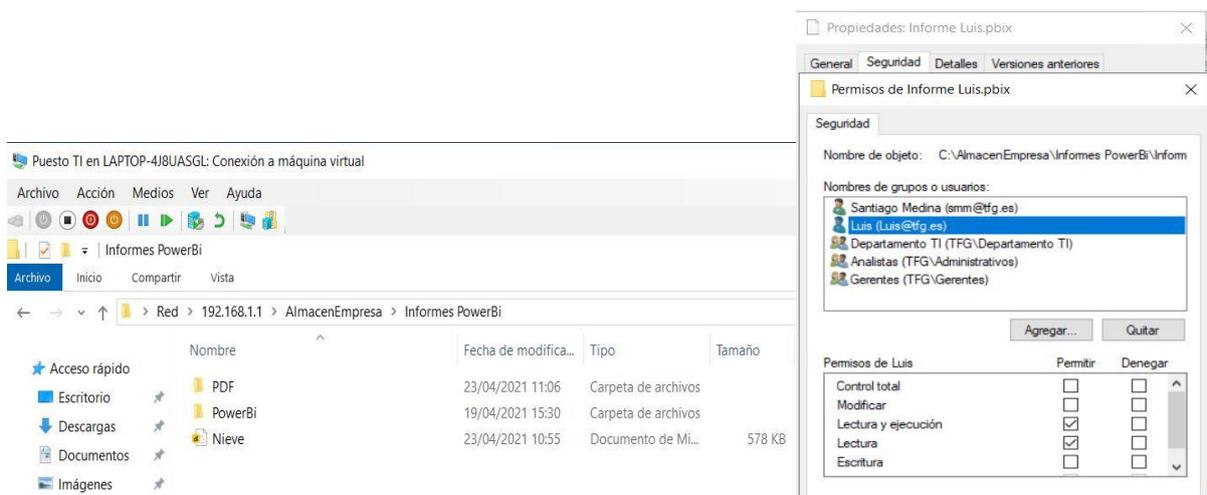


Figura 4.15: Directorio antes de compartir el informe y permisos que se aplican

En la figura 4.16, se muestra el resultado obtenido, teniendo el informe solo visible para el Analista Luis.

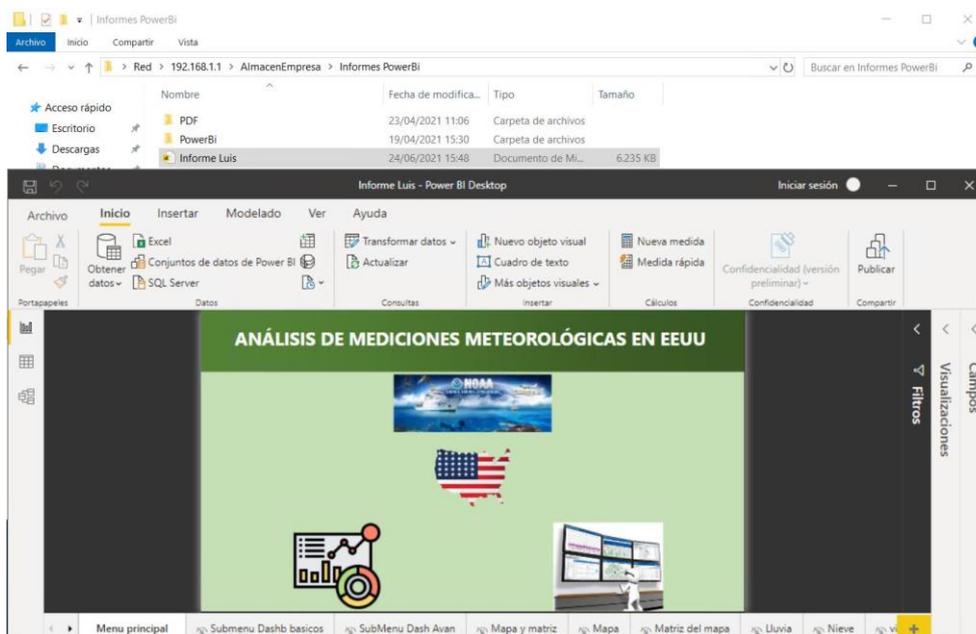


Figura 4.16: Informe compartido correctamente

- Publicación de informes en Power Bi Services, permite compartir informes en un repositorio de PowerBi como se muestra en la figura 4.17, donde podemos visualizar los informes realizados permitiendo crear áreas de trabajo y añadir a usuarios en esta, con diferentes permisos cómo: administradores, miembros, visores y colaboradores.

Esta parte del proyecto no se pudo completar porque se necesitan direcciones de correo '@inlumine.ual.es'. No obstante, la figura 4.18 indica donde se agregarían los colaboradores. Utilizando la prueba de 60 días de la cuenta Power BI Pro.

Además, se puede visualizar un informe publicado desde PowerBi Mobile (Figura 4.19).

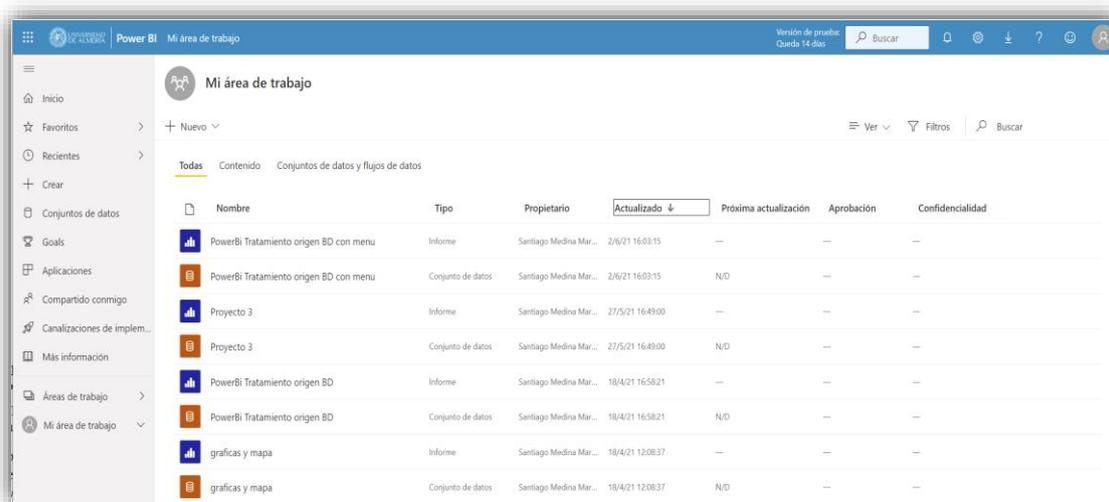


Figura 4.17: Compartir un proyecto en Power BI Service

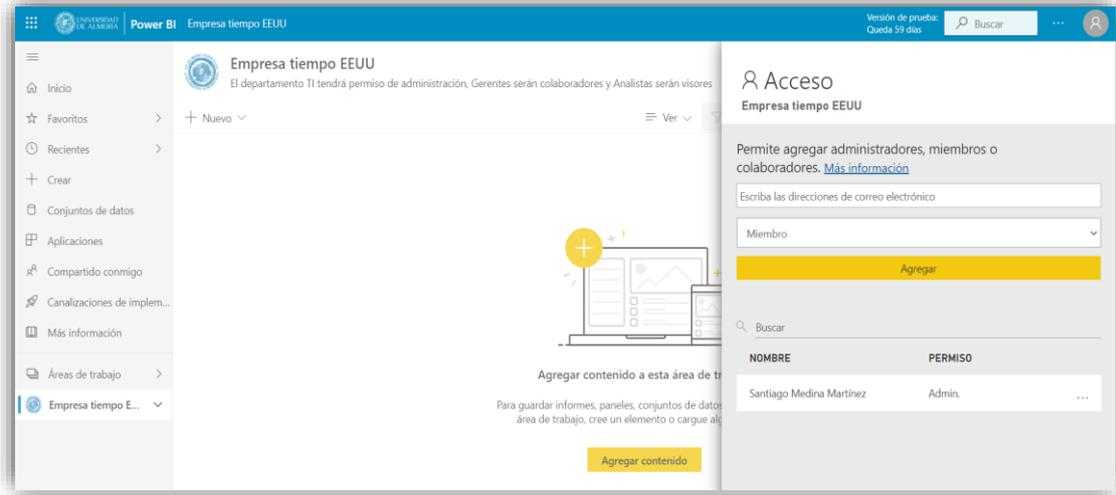


Figura 4.18: Crear un área de trabajo

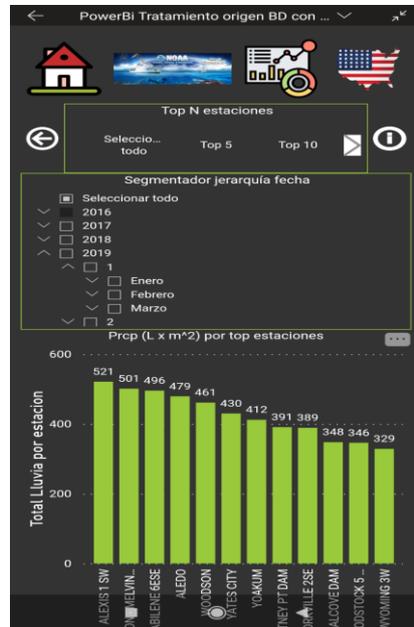


Figura 4.19: Vista de informe desde Power BI Mobile

Para publicar los informes en una página web se utilizará Blogger, en la figura 4.20 se muestra cómo se obtiene el código HTML, para insertarlo en la página web. En la figura 4.21 se visualiza la creación de la entrada y la vista final del informe.



Figura 4.20: Código HTML del informe

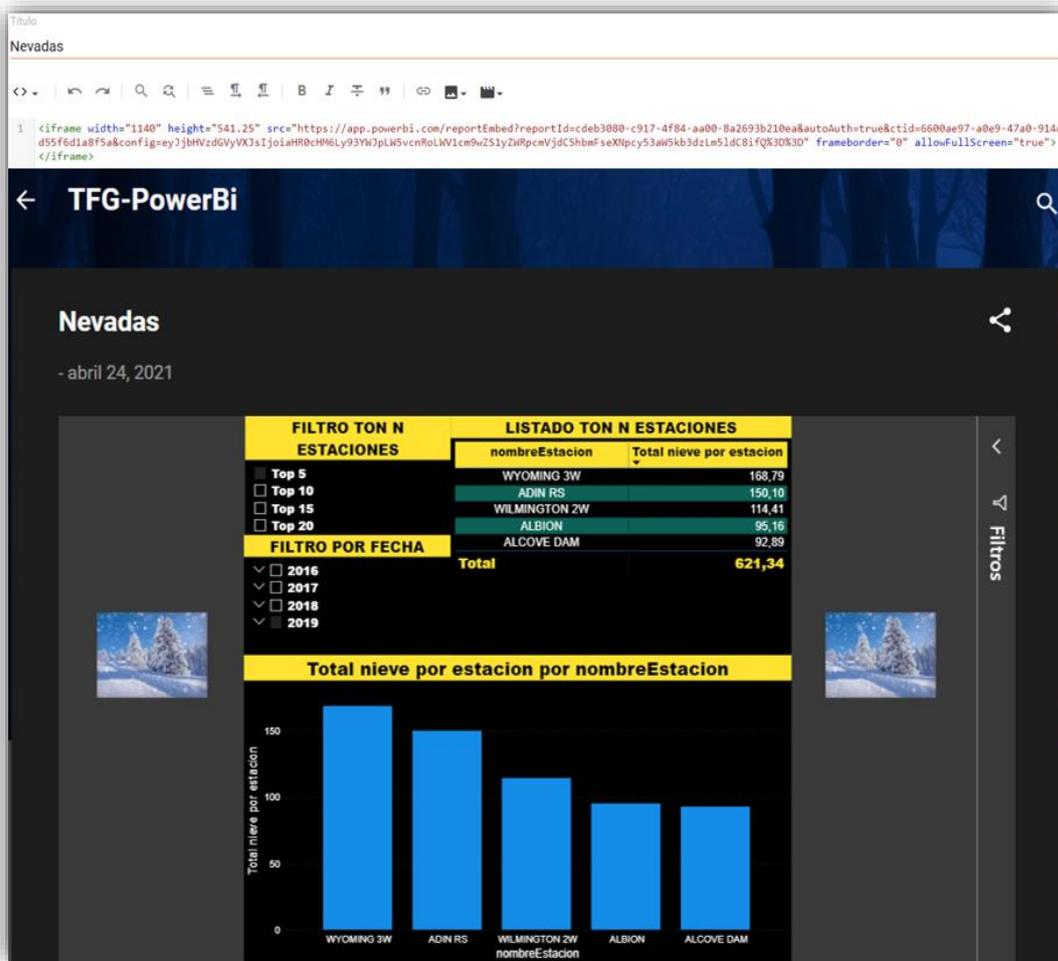


Figura 4.21: Informe publicado en Blogger

4.5 Resumen del capítulo

En este capítulo, se han elaborado y estudiado los informes de datos tratados en el capítulo anterior. Hemos visto cómo se fabrican plantillas para los informes, la navegación entre páginas, las tareas de limpieza y transformación de datos con Power Query y multitud de Dashboard incluyendo el uso de funciones DAX.

Para finalizar, se han creado los recursos compartidos necesarios en el servidor DC, asignando permisos para acceder a ellos, a usuarios y grupos de usuarios. Estas carpetas permiten el intercambio de información entre los usuarios de la infraestructura. También se ha visto como compartirlos con Power BI Service y publicarlos en la web.

En el siguiente capítulo, se verán los resultados obtenidos y, se realizará un estudio de las posibles mejoras que se puedan aplicar a este proyecto y conclusiones.

Palabras clave: Power BI, Power BI Service, Power BI Mobile, Power Query, DAX, Dashboard.



5. Conclusiones y trabajo futuro

Como último apartado se mostrarán las conclusiones a los resultados obtenidos durante el proyecto y posibilidades de mejoras del mismo.

5.1 Conclusiones

Gracias a este proyecto cuyo objetivo era la simulación de una PYME con toda la infraestructura IT, que admitiera el teletrabajo usando una VPN, he tenido la oportunidad de conocer en detalle el sistema operativo Microsoft Windows Server 2019. He necesitado profundizar en el estudio de Active Directory en especial en aspectos como la creación, administración de infraestructuras de redes y GPOs. También he tenido que adquirir conocimientos del rol “Servicio y Acceso Remoto” para poder implementar la VPN’.

Una vez desplegada la infraestructura IT se crea el DataWarehouse para tener los datos meteorológicos almacenados de forma organizada y visualizar la información con Dashboard, al objeto de permitir su análisis y apoyar la toma de decisiones. Esta parte del proyecto me ha hecho investigar y profundizar en el estudio de asignaturas estudiadas en el grado como Bases de Datos, Almacenes de Datos e Inteligencia de negocio. Incluyendo la herramienta Power BI para la visualización y análisis de datos.

Durante el desarrollo del proyecto se han cumplido los objetivos y plazos establecidos obteniendo como resultado la infraestructura con el Sistema de Información descrito.

El uso de estas tecnologías, han hecho que este proyecto sea multidisciplinar permitiéndome ampliar mis conocimientos en varios ámbitos de la informática.

5.2 Trabajo futuro

En esta sección, se hablará de las posibles modificaciones o mejoras que pueden realizarse en este trabajo.

- Un aspecto interesante sería utilizar **Cloud Computing**, donde se implementaría toda la infraestructura IT creada en la nube utilizando AWS, Google Cloud o Azure, compatibles con AD. Esto ayudaría al estudio de esta parte de la informática que se encuentra en un continuo crecimiento en la actualidad. Pero elevando el coste de esta.
- Crear una página web para que los usuarios puedan publicar información a través de esta.
- El proyecto se podría realizar con una empresa real que quisiera implementar el teletrabajo y con una Base de Datos con la que poder desplegar Business Intelligence.

Por último, sería interesante la creación de un área de trabajo en PowerBi, donde cada usuario tenga roles y permisos de los datos dependiendo de su puesto en el organigrama.



6. Bibliografía

- [1] H. L. Bernal, *GUIA PARA LA CONSTRUCCION DE UN DATA WAREHOUSE*, 2002.
- [2] PowerData, «Data Warehouse: todo lo que necesitas saber sobre almacenamiento de datos,» [En línea]. Available: <https://www.powerdata.es/data-warehouse>. [Último acceso: 05 04 2021].
- [3] Media, «¿Qué es Business Intelligence (BI) y qué herramientas existen?,» 05 10 2017. [En línea]. Available: <https://blog.signaturit.com/es/que-es-business-intelligence-bi-y-que-herramientas-existen#:~:text=1958%3A%20el%20investigador%20de%20IBM,del%20t%C3%A9rmino%20inteligencia%20de%20negocio>. [Último acceso: 20 04 2021].
- [4] S. Ramos, *Business Intelligence (BI) & Analytics: El arte de convertir los datos en conocimiento*, SolidQ, 2016.
- [5] M. R. y. A. Ferrari, *The Definitive Guide to DAX: Business intelligence with Microsoft Excel, SQL Server Analysis Services, and Power BI*, Microsoft Press, 2017.
- [6] J. R. A. G. L.-N. R. A. Brian Desmond, *Active Directory: Designing, Deploying and Running Active Directory*, O'REILLY, 2003.
- [7] C. Simmons, *Active Directory Bible*, IDG Books, 2013.
- [8] D. Heywood, *Redes con Microsoft TCP/IP*, PRENTICE HALL, 1999.
- [9] R. M. Hicks, *Implementing DirectAccess with Windows Server 2016*, Apress, 2016.
- [10] ostec, «VPN SSL, entienda las características y los beneficios,» [En línea]. Available: <https://ostec.blog/es/generico/vpn-ssl-beneficios/>. [Último acceso: 28 03 2021].
- [11] A. y. R. Siddaway, *Pro Microsoft Hyper-V 2019*, Apress, 2018.
- [12] J. A. Castillo, «Active Directory Que es y para qué sirve,» 15 Diciembre 2018. [En línea]. Available: https://www.profesionalreview.com/2018/12/15/active-directory/#Que_es_Active_Directory.
- [13] Paessler, «IT Explained: Active Directory,» 20 febrero 2021. [En línea]. Available: <https://www.es.paessler.com/it-explained/active-directory>.
- [14] J. Moskowitz, *Group Policy: Fundamentals, Security, and the Managed Desktop*, SYBEX, 2010.
- [15] G. Delprato, «Cómo Funciona DNS,» 24 06 2011. [En línea]. Available: <https://windowserver.wordpress.com/2011/06/24/cmo-funciona-dns-parte-3-integracin-con-active-directory/>.

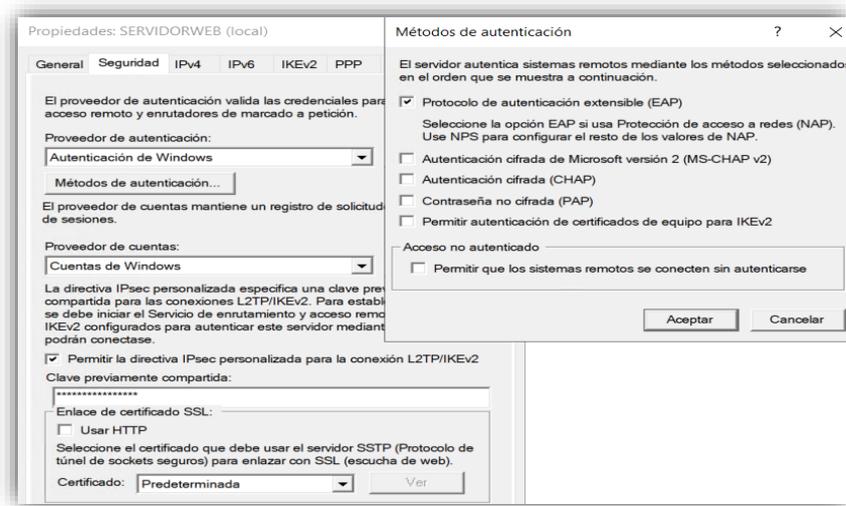
- [16] D. Simpson, «Opciones de autenticación de VPN,» 27 julio 2017. [En línea]. Available: <https://docs.microsoft.com/es-es/windows/security/identity-protection/vpn/vpn-authentication>. [Último acceso: 25 marzo 2021].
- [17] PureVPN, «VPN IKEv2,» [En línea]. Available: <https://www.purevpn.com/what-is-vpn/protocols/ikev2>. [Último acceso: 27 03 2021].
- [18] R. Cotaquispe, «¿Qué es VMware? Aprende a crear computadoras virtuales para ser experto en cloud computing,» 9 04 2021. [En línea]. Available: <https://www.crehana.com/es/blog/web/que-es-vmware/>. [Último acceso: 23 06 2021].
- [19] J. A. Castillo, «Crear nuestra VPN con OpenVPN en Windows: todo el proceso,» 18 04 2021. [En línea]. Available: <https://www.profesionalreview.com/2020/04/18/crear-vpn-openvpn-windows/>. [Último acceso: 23 06 2021].
- [20] S. D. Luz, «WireGuard VPN: Instalación y configuración de la mejor VPN del momento,» 4 03 2021. [En línea]. Available: <https://www.redeszone.net/tutoriales/vpn/wireguard-vpn-configuracion/>. [Último acceso: 24 06 2021].
- [21] J. Jiménez, «¿Tengo que abrir puertos para que la VPN funcione mejor?,» 4 diciembre 2020. [En línea]. Available: <https://www.redeszone.net/noticias/redes/abrir-puertos-usar-vpn/>. [Último acceso: 24 marzo 2021].
- [22] R. W. y. C. K. Robert Wrembel, *Data Warehouses and OLAP Concepts, Architectures and Solutions*, IGI Global, 2006.
- [23] «Historical Observing Metadata Repository (HOMR),» NOAA, [En línea]. Available: <https://www.ncdc.noaa.gov/homr/reports>. [Último acceso: 20 02 2021].
- [24] R. Dobson, «Almacén de datos meteorológicos en SQL Server: crear y completar,» 16 09 2020. [En línea]. Available: <https://www.mssqltips.com/sqlservertip/6547/weather-data-warehouse-in-sql-server--create-and-populate/>. [Último acceso: 09 04 2021].
- [25] Tableau, «¿Qué es la inteligencia de negocios? Guía sobre la inteligencia de negocios y por qué es importante,» 28 02 2021. [En línea]. Available: <https://www.tableau.com/es-es/learn/articles/business-intelligence>.
- [26] «¿Qué es Power BI Desktop?,» Microsoft, 10 03 2021. [En línea]. Available: <https://docs.microsoft.com/es-es/power-bi/fundamentals/desktop-what-is-desktop#:~:text=Power%20BI%20Desktop%20es%20una,los%20datos%2C%20transformarlos%20y%20visualizarlos.&text=Con%20Power%20BI%20Desktop%2C%20puede,en%20un%20modelo%20de%20datos>. [Último acceso: 22 04 2021].

Anexo

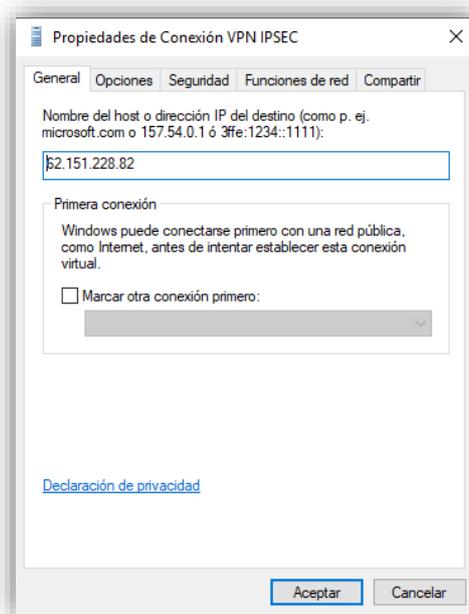
Conexión VPN

En este anexo, se verán las configuraciones de VPN tanto en el Servidor de la infraestructura como en ordenadores externos a esta.

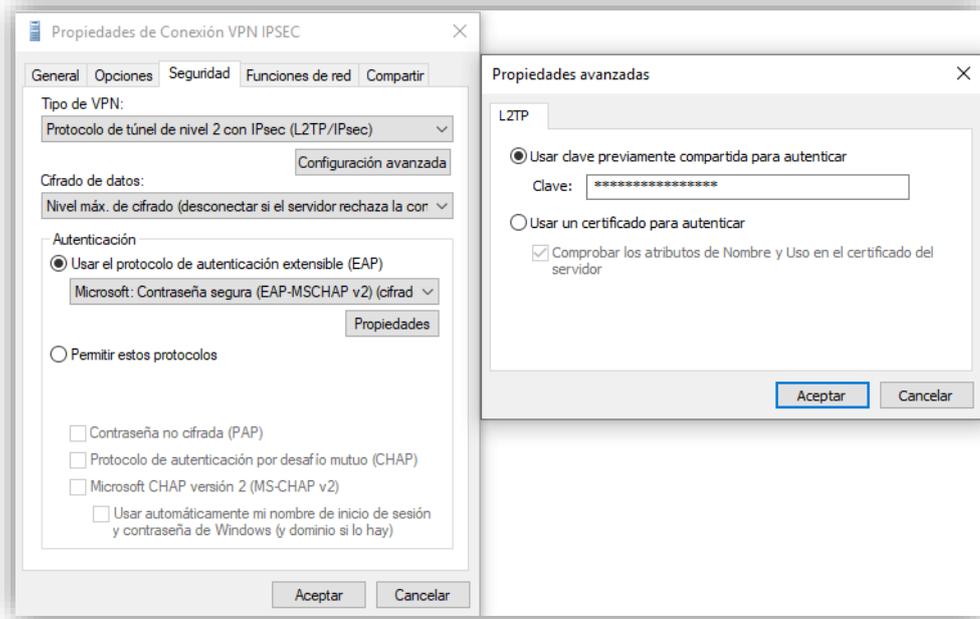
La siguiente imagen muestra los parámetros que se deben seleccionar en las propiedades del servidor en la consola de Enrutamiento y Acceso remoto para conseguir una VPN L2TP/IPSEC.



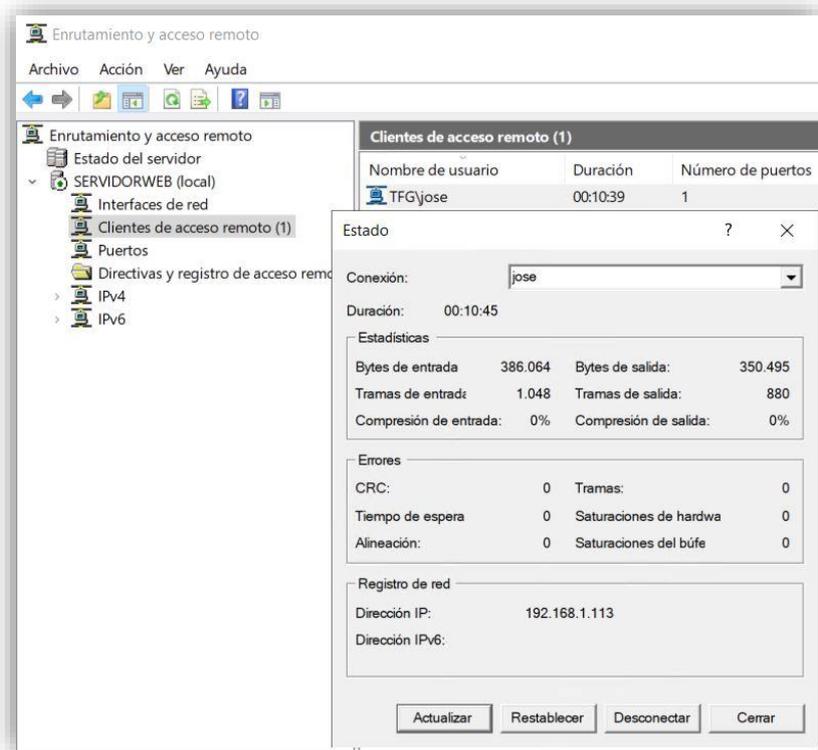
Una vez configurado el servidor y reiniciado, se procede a realizar la configuración en un PC externo. En primer lugar, se apunta como dirección de destino a la IP pública del servidor donde se ha configurado la VPN.



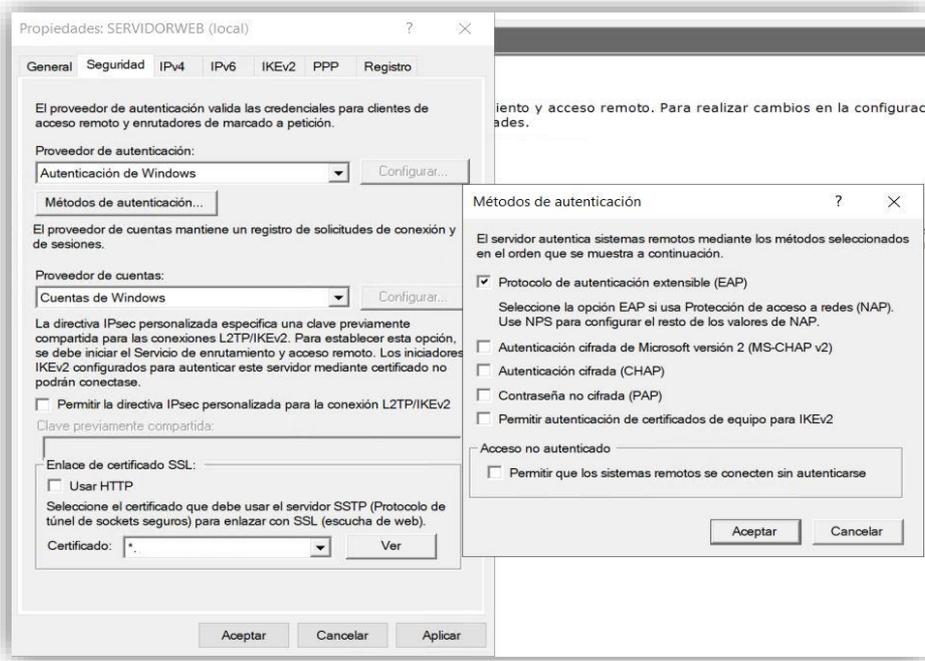
Después, se configuran las propiedades con los siguientes parámetros y la clave compartida:



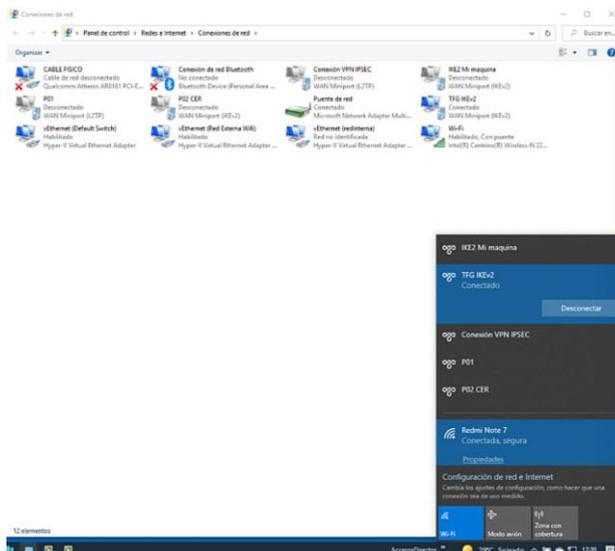
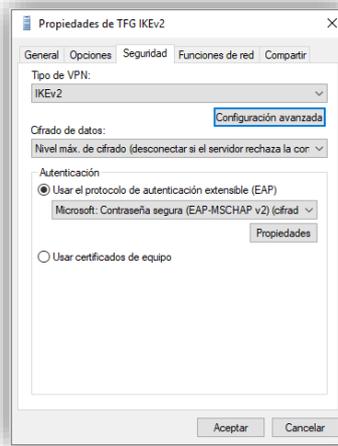
Una vez configurada la conexión se puede conectar correctamente:



Para configurar la VPN IKEv2, hay que instalar los certificados en los equipos cliente y servidor. Después, en el servidor se establecen los siguientes parámetros y se establece el certificado utilizado:



Para los PC externos se establece la siguiente configuración:



También habilitar las siguientes reglas en el router para permitir estas conexiones vpn:

Personalizar reglas						
estado	aplicación / servicio	puerto interno	puerto externo	protocolo	IPv4 del dispositivo	
	FTP Server	21	21	TCP		añadir
✓	UDP500	500	500	UDP	10.10.10.101	delete
✓	UDP1500	1500	1500	UDP	10.10.10.101	delete
✓	UDP4500	4500	4500	UDP	10.10.10.101	delete
✓	TCP1701	1701	1701	both	10.10.10.101	delete

Para finalizar el anexo, en el siguiente video se muestra cómo se crea una conexión L2TP desde 0:



Scripts para modelado de datos

Para la inserción de los archivos planos de datos en la Base de Datos se usa el siguiente Script:

```

-- Paso 1
use noaa_data
create table #CopiarArchivoCALIFORNIA
(
    noaa_station_id nvarchar(15)
    , elevacion real
    , latitud_grad_dec real
    , longitud_grad_dec real
    --da error tipo date lo cambiare más adelante
    , fecha nvarchar(15)
    , mdpr real
    , dapr real
)

--Paso 2
-- Importación fichero
BULK INSERT ImportedFileTable
FROM 'C:\Users\santi\Desktop\TFG\BD_WEATHER\EstacionesMeteorologicas\CALIFORNIA.txt'
insert into #CopiarArchivoCalifornia
(
    noaa_station_id
    , elevacion
    , latitud_grad_dec
    , longitud_grad_dec
    , fecha
)
    
```

<pre> ,prcp real ,nieve real ,pNieve real ,tmax int ,tmin int ,tobs int ,dirViento int ,velViento real) Go if EXISTS(SELECT * FROM sys.objects WHERE name = 'ImportedFileTable' and TYPE = 'u') DROP TABLE ImportedFileTable CREATE TABLE ImportedFileTable(textvalue varchar(max)) GO </pre>	<pre> ,mdpr ,dapr ,prcp ,nieve ,pNieve ,tmax ,tmin ,tobs ,dirViento ,velViento) </pre>
<pre> --Paso 3.1 select substring(textvalue,12,6) [STATON ID] ,substring(textvalue,70,10) elevation ,substring(textvalue,81,10) lat_dec_de- gree ,substring(textvalue,92,10) long_dec_de- gree ,substring(textvalue,107,2) + '/' + subs- tring(textvalue,109,2) + '/' + subs- tring(textvalue,103,4) [date], case when substring(textvalue,112,8) = '- 9999' then null else substring(textvalue,112,8) end M DPR, case when substring(textvalue,121,8) = '- 9999' then null else substring(textvalue,121,8) end D APR, case when substring(textvalue,130,8) = '- 9999' then null else substring(textvalue,130,8) end P RCP, case when substring(textvalue,139,8) = '- 9999' then null else substring(textvalue,139,8) end S NWD, case when substring(textvalue,148,8) = '- 9999' then null else substring(textvalue,148,8) end S NOW, </pre>	<pre> --Paso 3.2 case when substring(textvalue,157,8) = '- 9999' then null else substring(textvalue,157,8) end T MAX , case when substring(textvalue,166,8) = '- 9999' then null else substring(textvalue,166,8) end T MIN , case when substring(textvalue,175,8) = '- 9999' then null else substring(textvalue,175,8) end T OBS , case when substring(textvalue,184,8) = '- 9999' then null else substring(textvalue,184,8) end W DFG , case when substring(textvalue,193,8) = '- 9999' then null else substring(textvalue,193,8) end W SFG from ImportedFileTable where isnumeric(substring(textva- lue,12,6))=1 order by [STATON ID],date </pre>
<pre> -- Paso 4: Copia a la tabla select * into CALIFORNIA from #CopiarArchivoCalifornia order by noaa_station_id, fecha </pre>	

En el caso de Texas, Nueva York, Illinois y Florida se reutilizará el mismo Script mostrado anteriormente con pequeñas modificaciones como el archivo escogido, nombre de atributo y nombres de tabla. Tras importar todos los archivos tenemos las siguientes tablas en la base de datos:

NUEVA YORK	ILLINOIS	FLORIDA
noaa_station_id	noaa_station_id	noaa_station_id
latitud_grad_dec	latitud_grad_dec	elevacion
longitud_grad_dec	longitud_grad_dec	latitud_grad_dec
elevacion	elevacion	longitud_grad_dec
fecha	fecha	fecha
dapr	dapr	mdpr
mdpr	mdpr	dapr
prcp	prcp	prcp
profNieve	nieve	nieve
nieve	profNieve	pNieve
tMedia	tmax	tmax
tmax	tmin	tmin
tmin	tobs	tobs
tobs	wesd	dirViento
wesd		velViento
wesf		
vMediaViento		
dirRachaMasRapida2min		
dirRachaMasRapida5seg		
velRachaMasRapida2min		
velRachaMasRapida5seg		

TEXAS	CALIFORNIA
noaa_station_id	noaa_station_id
elevacion	elevacion
latitud_grad_dec	latitud_grad_dec
longitud_grad_dec	longitud_grad_dec
fecha	fecha
mdpr	mdpr
dapr	dapr
prcp	prcp
nieve	nieve
pNieve	pNieve
tmax	tmax
tmin	tmin
tobs	tobs
dirViento	dirViento
velViento	velViento

Para construir el DataWarehouse se realiza un modelo en estrella, con una tabla de hechos y sus dimensiones, con los datos presentados anteriormente en las tablas, se construye el diagrama:

- **Hechos_MedicionesTiempo:** Se define una tabla de hechos con los siguientes atributos usando el script (Tabla 4.3), esta tabla tendrá las medidas meteorológicas y está conectada a dimensión_Tiempo y dimensión_Lugar, a través de las Foreign Key.
 - **FK:** noaa_station_id y fecha.
 - **Medidas:** Temperaturas máxima, mínima y observada, lluvia, nieve, profundidad de la nieve, dirección y velocidad del viento.

```

select *
into Hechos_MedicionesTiempo
from
(select noaa_station_id, fecha, tmax, tmin
, tobs, prcp, nieve, pnieve, dirViento, velViento
from [dbo].[CALIFORNIA]
where noaa_station_id in
(select distinct noaa_station_id from
[dbo].[CALIFORNIA])
union
select
noaa_station_id, fecha, tmax, tmin
, tobs, prcp, nieve, pnieve, dirViento
, velViento
from [dbo].[FLORIDA]
where noaa_station_id in
(select distinct noaa_station_id from
[dbo].[FLORIDA])

```

```

union
select
  noaa_station_id
, fecha, tmax, tmin, tobs, prcp, nieve
, profNieve, null dirViento, null velViento
from [dbo].[ILLINOIS]
where noaa_station_id in
(select distinct noaa_station_id from
[dbo].[ILLINOIS])

union
select
  noaa_station_id, fecha, tmax, tmin, tobs
, prcp, nieve, profNieve, dirRachaMasRa-
pida2min, vMediaViento
from [dbo].[NUEVA_YORK]
where noaa_station_id in
(select distinct noaa_station_id from
[dbo].[NUEVA_YORK])

union
select
  noaa_station_id, fecha, tmax, tmin, tobs, prcp, nieve, pNieve, dirViento, velViento
from [dbo].[TEXAS]
where noaa_station_id in
(select distinct noaa_station_id from [dbo].[TEXAS])) Hechos_MedicionesTiempo
order by noaa_station_id, fecha
-- Mostrar tabla
select * from Hechos_MedicionesTiempo
order by noaa_station_id, fecha

```

➤ **Dimension_Fecha:** Almacenará los datos de las fechas que comprenden los hechos, es decir, desde 01-01-2016 hasta 31-12-2019. El Script para generar esta dimensión se muestra en (Tabla 4.4). La dimensión se compone:

- **Primary Key:** fecha.
- **Campos:** numeroDiaSemana, diaSemana, diaMes, diaAño, mes día semana, numero mes y año.

```

--Se crean datos desde 01-01-2016 hasta
31-12-2019
declare @año int = 4, @fechaInicio
date = '2016-01-01'; --aaaa/mm/dd
declare @fechaFin date =
  DATEADD(DAY, -1, DATEADD(YEAR, @año,
@fechaInicio));
;WITH seq(n) AS
(SELECT 0 UNION ALL SELECT n + 1 FROM
seq
WHERE n < DATEDIFF(DAY, @fechaInicio,
@fechaFin)),
d(d) AS
(SELECT DATEADD(DAY, n, @fechaInicio)
FROM seq)
select
d fecha ,
datepart(weekday,d) numeroDiaSemana,
datename(weekday,d) diaSemana,
datename(month, d) mes,
month(d) numeroMes,
case
  when month(d) >=1 and month(d) <=3
then 'C1'
  when month(d) >=4 and month(d) <=6
then 'C2'
  when month(d) >=7 and month(d) <=9
then 'C3'
  when month(d) >=10 and month(d) <=12
then 'C4'
  else null
end cuatrimestre
,year(d) año
into Dimension_Fecha
from d
ORDER BY d
OPTION (MAXRECURSION 0);
--Muestra dimension fecha

```

<pre>datetime(DAY,d) diaMes, datetime(DAYOFYEAR, d) diaAño,</pre>	<pre>select * from Dimension_Fecha</pre>
---	--

- **Dimensión Lugar:** indicará los datos demográficos de las estaciones de la tabla de hechos, se crean previamente dos tablas y se realiza una intersección de los datos comunes.

La primera “unionEstaciones” contendrá todas las estaciones de los estados California, Texas, Illinois, Florida y Nueva York a partir de la tabla “ccsl_still_in_use” que contiene todas las estaciones de EEUU y su ubicación (Tabla 4.5).

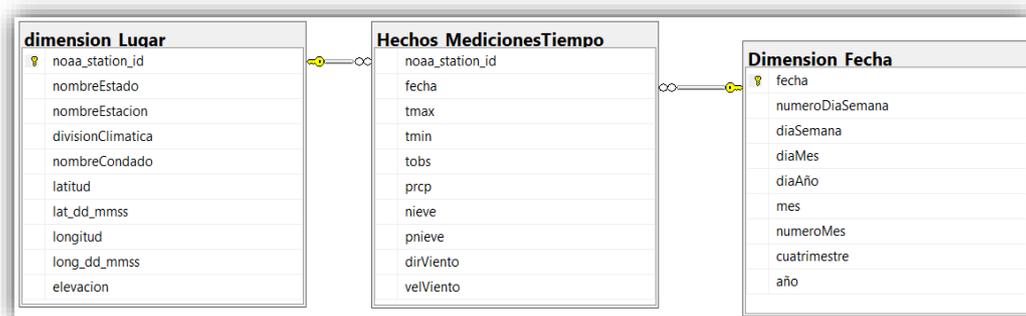
La segunda “estacionesDistintasEstados” creará la tabla a partir de una consulta (tabla 4.6) de las estaciones que tengan un id, latitud y longitud distintos de las tablas CALIFORNIA, FLORIDA, ILLINOIS, NUEVA_YORK y TEXAS.

<pre>--Crear tabla con estaciones de CA TX IL FL NY use noaa_data go select * into unionEstaciones from (select * from [dbo].[ccsl_still_in_use] where STATEPROV = 'ca' union</pre>	<pre>select * from [dbo].[ccsl_still_in_use] where STATEPROV = 'tx' union select * from [dbo].[ccsl_still_in_use] where STATEPROV = 'ny' union select * from [dbo].[ccsl_still_in_use] where STATEPROV = 'il') unido;</pre>
---	---

```
--Tabla estacionesDistintasEstados
select *
into estacionesDistintasEstados
from
(select distinct noaa_station_id, [latitud_grad_dec], [longitude_grad_dec],
'CA' state_abbr
from [dbo].[CALIFORNIA]
union
select distinct noaa_station_id, [latitud_grad_dec], [longitude_grad_dec],
'FL' state_abbr
from [dbo].[FLORIDA]
union
select distinct noaa_station_id, [latitud_grad_dec], [longitude_grad_dec],
'IL' state_abbr
from [dbo].[ILLINOIS]
union
select distinct noaa_station_id, [latitud_grad_dec], [longitude_grad_dec],
'NY' state_abbr
from [dbo].[NUEVA_YORK]
union
select distinct noaa_station_id, [latitud_grad_dec], [longitude_grad_dec],
'TX' state_abbr
from [dbo].[TEXAS]) estacionesDistintasEstados
order by state_abbr
```

```
--Tabla dimension_Lugar
select
STATEPROV ,[STATION NAME] as nombreEstacion ,[COOP STATION ID] noaa_station_id
,[ccsl_still_in_use].[CLIMATE DIVISION] as divisionClimatica
,[ccsl_still_in_use].[COUNTY NAME] as condado
,[ccsl_still_in_use].[NCDC STATION ID]
,[ccsl_still_in_use].[LATITUDE] as latitud
,[ccsl_still_in_use].[lat_dd.mmss] ,[ccsl_still_in_use].[LONGITUDE] as longitud
,[ccsl_still_in_use].[long_dd.mmss] ,[ccsl_still_in_use].[ELEVATION] as elevacion
into dimension_Lugar
from [dbo].[ccsl_still_in_use]
where [COOP STATION ID] in
(select noaa_station_id
from [dbo].[estacionesDistintasEstados]
intersect
select [COOP STATION ID] from [dbo].[unionEstaciones])
```

El diagrama en estrella creado es el siguiente:



Informes Power BI

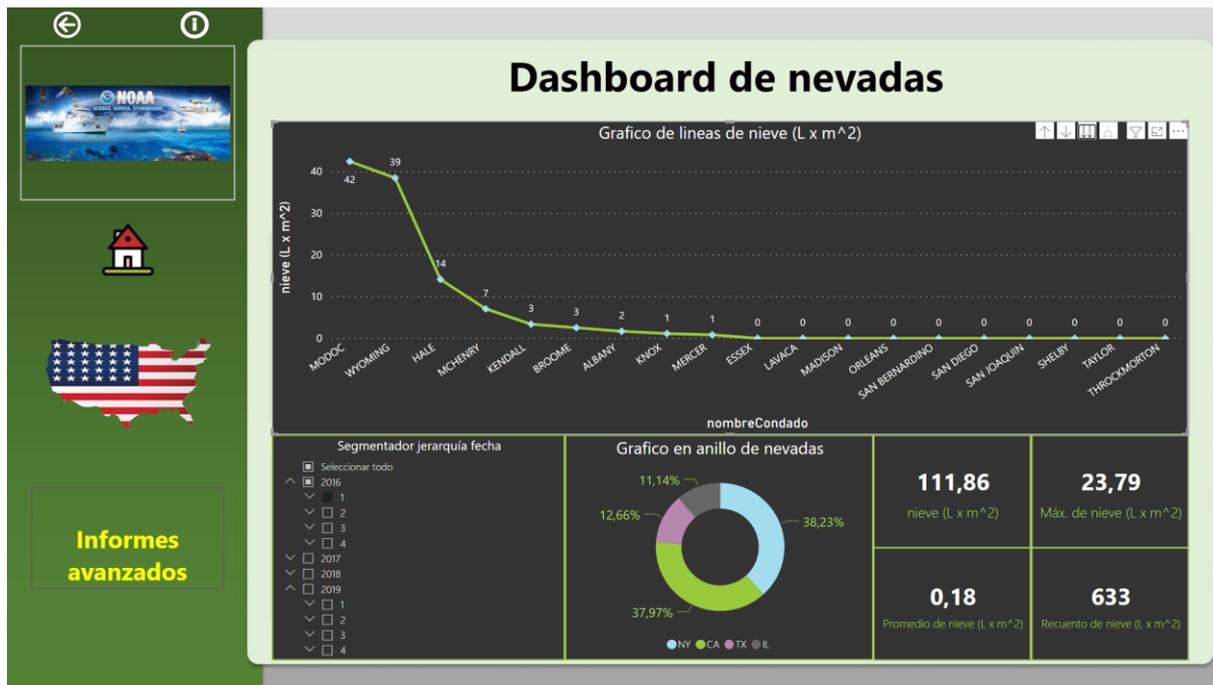
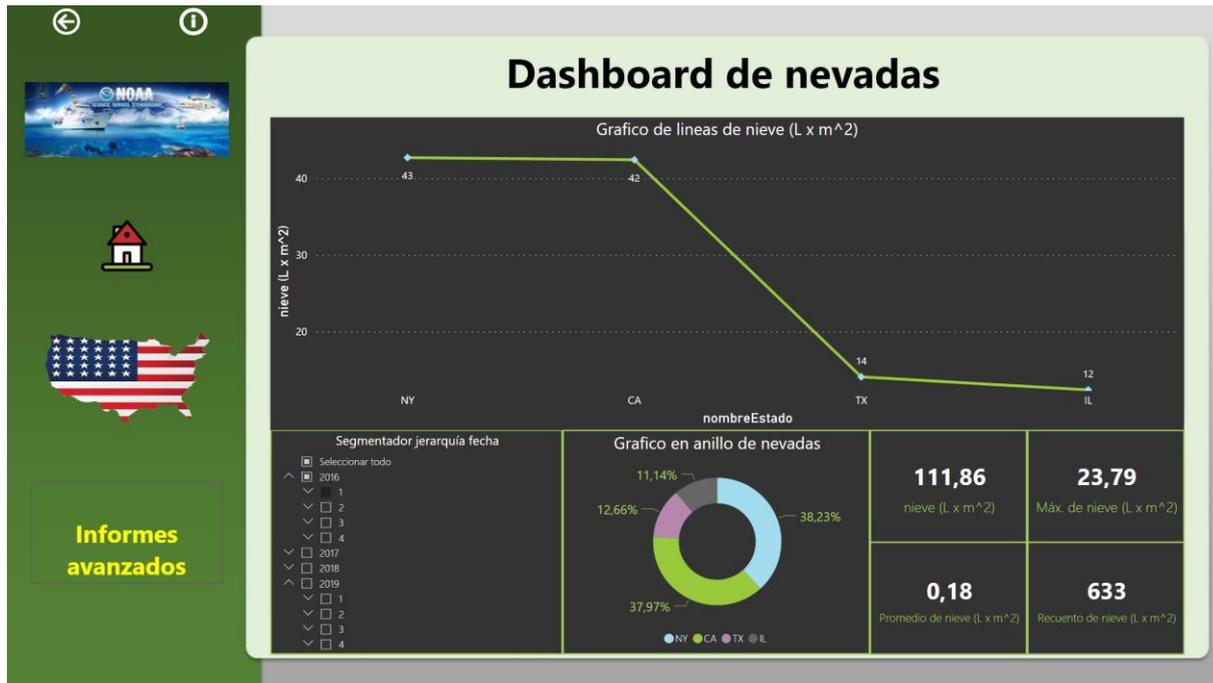
En la siguiente imagen se muestra el menú principal donde se permite la navegación, a través de botones, a la página web de Noaa, al mapa de estaciones, al submenú de informes básicos o al submenú de informes avanzados.



A continuación, se visualiza el submenú de informes básicos, donde se permite el acceso a 5 informes. Además, en la barra izquierda se permite navegar a la página web de Noaa, volver al inicio, ir mapa de estaciones o al submenú de informes avanzados. Facilitando el uso de la interfaz al usuario.



Todos los informes siguen la misma tipología, el mostrado en la siguiente imagen, permite la navegación por la jerarquía de lugar (País-Estado-Condado-Estación).



Por último, se mostrará el submenú de informes avanzados y un ejemplo de informe de este, el submenú se compone de 6 iconos para navegar a diferentes tipos de estudios como: Elementos más influyentes, preguntas y respuestas, Top N lluvias, Top N nevadas y un esquema jerárquico. En la barra izquierda se permite navegar a la página web de Noaa, volver al inicio, ir al submenú de informes básicos o ir al mapa de estaciones.

Estudios meteorológicos avanzados en EEUU

En las siguientes imágenes puede seleccionar el tipo de informe que desea visualizar.

TOP

TOP

Dashboard temperaturas

Segmentador jerarquía fecha

Seleccionar todo

- 2016
- 2017
- 2018
 - 1
 - 2
 - 3
 - 4
- 2019

°C

Promedio de tmax(°C), Promedio de tobs (°C) y Promedio de tmin (°C) por nombreEstado

°F

	CA	IL	NY
Promedio de tmax(°C)	25,9	26	15
Promedio de tobs (°C)	19	17	5
Promedio de tmin (°C)	12	9	7,23

°F

	CA	NY	IL
Min. de tmin (°C)	-3	-23	-28
Promedio de tobs (°C)	19	9	8
Máx. de tmax(°C)	43	35	34

Segmentador jerarquía fecha

Seleccionar todo

- Estados Unidos
 - CA
 - LOS ANGELES
 - MODOC
 - ADIN RS
 - SAN BERNARDINO
 - SAN DIEGO
 - SAN JOAQUIN
 - IL
 - NY
 - TX



Este proyecto es un compendio de tecnologías para desplegar una PYME ficticia, dedicada al estudio de datos meteorológicos de EEUU, estos datos son introducidos en un DataWarehouse y estudiados con Power BI, para aplicar Business Intelligence. La infraestructura IT implementa una VPN para permitir el teletrabajo de los empleados.

El proyecto trata conceptos tan diversos como:

- Active Directory: creación de un dominio y usuarios de este, directivas de grupo, Unidades Organizativas, carpetas compartidas y permisos.
- Configuración de adaptadores de red y de un Servidor Web, creación e implementación de una VPN en la infraestructura IT.
- Obtención de datos, creación de Base de datos SQL Server y creación DataWarehouse.
- Transformación de datos, funciones DAX, análisis de datos, creación de informes y publicación de datos mediante Power BI en los directorios compartidos de la empresa y en la web.

Todos estos conceptos son necesarios para construir las distintas partes que conforman la infraestructura que se crea en el proyecto.

This project uses some technologies to deploy a small fictitious company, dedicated to the study of meteorological data from USA, these data are entered in a DataWarehouse and studied with Power BI, to apply Business Intelligence. The IT infrastructure implements a VPN to allow employees to teleworking.

The project works with concepts as diverse as:

- Active Directory: creation of a domain and its users, group policies, Organizational Units, shared folders and permissions.
- Configuration of network adapters and a Web Server, creation and implementation of a VPN in the IT infrastructure.
- Obtaining data, creating SQL Server Database and creating DataWarehouse.
- Data transformation, DAX functions, data analysis, reporting, and publishing of data using Power BI to shared company directories and a web blog.

All these concepts are necessary to build the different parts that make up the company that is created in the project.