

Configuración avanzada de IBM SPSS Modeler Entity Analytics

Introducción

Los destinatarios de esta guía son los administradores de sistemas que están configurando IBM SPSS Modeler Entity Analytics (EA) para su ejecución de forma desplegada con uno de los programas siguientes:

- IBM SPSS Collaboration and Deployment Services (C&DS) versión 8
- IBM SPSS Modeler Server
- IBM Analytical Decision Management (ADM) versión 18
- IBM SPSS Modeler Solution Publisher

Conceptos básicos

En esta guía se muestra cómo configurar IBM SPSS Collaboration and Deployment Services, Analytical Decision Management, Solution Publisher o Modeler Server que se ejecuten en una configuración de clústeres de COP (Coordinator of Processes) para acceder a un repositorio de Entity Analytics.

El repositorio de EA debe configurarse en un equipo que utilice Modeler Server para alojar el repositorio. ADM, Solution Publisher y C&DS Jobs and Scoring Service podrán utilizar después el nodo de generación de análisis de entidad en el Modeler que se ha configurado para buscar en este repositorio.

Cada repositorio de EA es gestionado por un único servidor de bases de datos (por ejemplo, en el release 17 de Modeler, se trata de un servidor de bases de datos IBM DB2) y una instancia del servicio de EA que se ejecuta en cada máquina.

Nota: Si solo utiliza Modeler Server, la mayor parte de esta guía no se aplica en su caso; sin embargo, es posible que desee poder concluir los servicios de EA. En este caso, debe utilizar el script `manage_repository`. Para obtener más información, consulte la sección Administración de DB2 y servicios de EAe EA.

El diagrama siguiente muestra la configuración de un clúster de máquinas que ejecutan el servicio de puntuación C&DS.

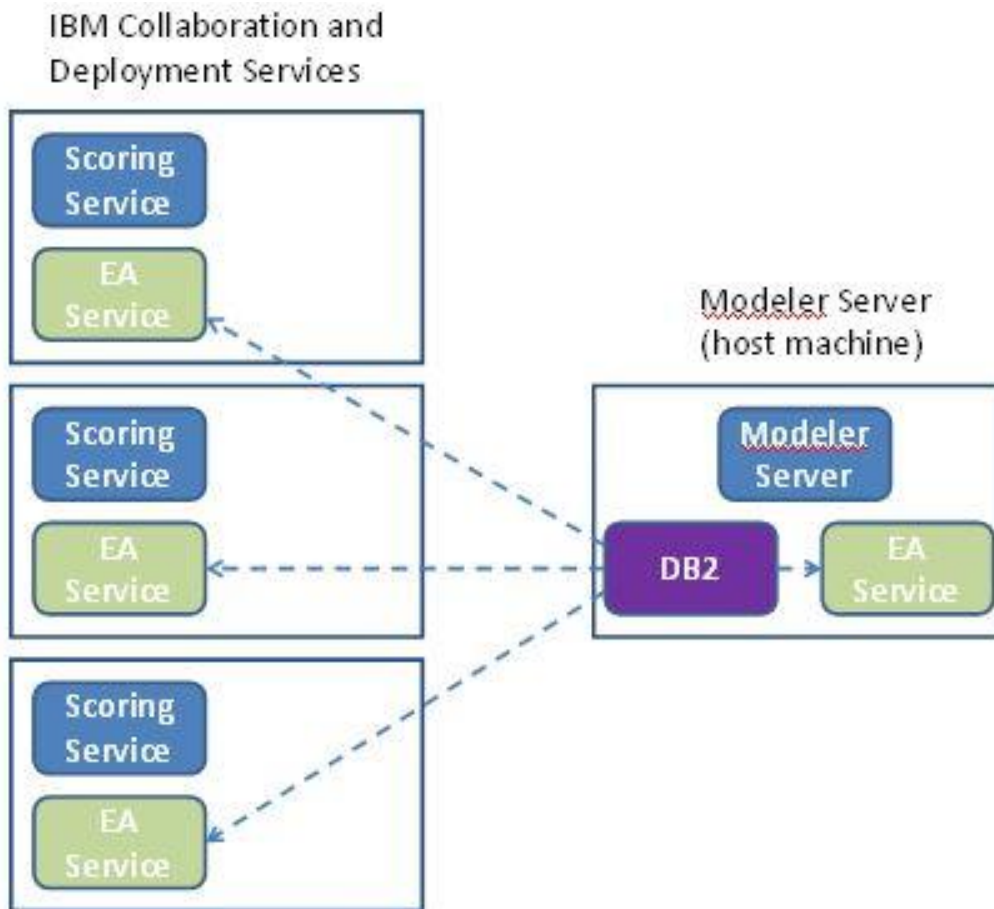


Figura 1 – C&DS – Configuración de máquinas de clúster

Configuración

El primer paso consiste en conocer la ubicación de la carpeta en la que se ha instalado EA en el Modeler Server host de repositorio y en cada una de las otras máquinas; necesitará esta información durante la configuración.

La lista siguiente muestra las ubicaciones predeterminadas para distintas plataformas.

Windows Vista, Windows 7, Windows Server:

C:\ProgramData\IBM\SPSS\Modeler\18\EA

Tenga en cuenta que EA requiere la biblioteca de tiempo de ejecución de Microsoft Visual C++ 2010 en Windows. Si es necesario, instale la biblioteca de tiempo de ejecución Visual C++ 2010 (que se encuentra disponible en el centro de descargas de Microsoft) antes de ejecutar las funciones de EA.

UNIX:

<directorio-instalación-modeler>/ext/bin/pasw.entityanalytics/EA

Modeler Server: Configuración de host de repositorio

Como requisito previo, Entity Analytics tiene estar instalado en un equipo con Modeler Server que aloje el repositorio y en el que deberá crearse un repositorio de Entity Analytics (tenga en cuenta que este equipo puede ser local o remoto). Esta máquina aloja el repositorio al que se accede desde otros servicios, como por ejemplo:

- Collaboration and Deployment Services (C&DS)
- Decision Management
- Solution Publisher
- Otros servidores Modeler Server que se ejecutan en un clúster Coordinator of Processes (COP).

Utilice el script `manage_repository` para iniciar y detener los servicios de repositorio en el equipo de Server Modeler que aloja el repositorio. Asegúrese de que se hayan iniciado los servicios de un repositorio concreto para que las rutas ejecutadas desde C&DS, Analytical Decision Management, o Solution Publisher, o que se haya configurado para la puntuación en C&DS, puedan acceder correctamente al repositorio de Entity Analytics.

Para obtener más información sobre cómo utilizar el script `manage_repository`, consulte Administración de DB2 y servicios de EA.

Configuración de C&DS

Después de configurar el repositorio, debe utilizar el cliente del Gestor de despliegue de C&DS para añadir una definición del servidor Entity Analytics. Esto es necesario para habilitar la puntuación en tiempo real para Entity Analytics y para habilitar el uso de una ruta de Modeler que contiene un nodo de Entity Analytics en un trabajo C&DS (en otras palabras, para ejecutar rutas de Entity Analytics en

C&DS). La definición del servidor debe coincidir con el nombre del repositorio en la ruta. Esta definición se utiliza para facilitarle a la ruta la ubicación del repositorio y proporcionarle la información de conexión que necesita.

Para crear una definición del servidor de Entity Analytics:

- (1) Edite el script de inicio C&DS `startServer.sh` añadiendo dos variables de entorno, como en el ejemplo siguiente.

Nota: En este ejemplo, `/opt/ibm/V10.5` es el directorio de instalación de DB2 y `/home/g2user` es la carpeta de inicio del usuario de la instancia de DB2 que se ha utilizado durante la creación del repositorio.

```
DB2_INSTALL_PATH=/opt/ibm/V10.5/
export DB2_INSTALL_PATH
DB2_INSTANCE_HOME=/home/g2user
export DB2_INSTANCE_HOME
ICC_PKCS11_ROOT=<modeler>/ext/bin/pasw.entityanalytics/g2
export ICC_PKCS11_ROOT
```

Nota: Asegúrese de que tiene permisos de lectura/escritura/modificación para la carpeta `ICC_PKCS11_ROOT`.

- (2) Reinicie el servidor C&DS.
- (3) En el cliente del Gestor de despliegue, añada una nueva definición de credenciales. El nombre de usuario y la contraseña deben coincidir con la credencial utilizada en la máquina de Modeler Server.
- (4) En el cliente del Gestor de despliegue, añada una nueva definición de servidor:
 - a. Escriba el nombre del servidor. Debe coincidir con el nombre del repositorio utilizado en la máquina de Modeler Server.
 - b. Seleccione **Entity Analytics Server** como tipo y pulse **Siguiente**.
 - c. Escriba la ruta del directorio o el URL del servidor de Entity Analytics con el que se realizará la conexión. Este es el valor “connect” que se muestra en el archivo `g2.ini`.
 - d. Entre el nombre y la suma de comprobación de Salt relacionados con el repositorio que se utiliza para los datos cifrados; esto se muestra en el archivo `g2.ini`.
 - e. Seleccione las credenciales que se utilizarán. Son las credenciales que ha añadido en el Paso 1.

- f. Exporte el valor de Salt de Modeler Server e impórtelo a C&DS. Tenga en cuenta que antes de exportar o importar el valor de Salt utilizando la línea de comandos, debe establecer/exportar la variable de entorno: ICC_PKCS11_ROOT.
- Para exportar un valor de Salt especial:
manage_repository -exportsalt contraseña nombresalt nombrearchivoexport
 - Para exportar todos los valores de Salt:
manage_repository -exportsalt -all contraseña carpetaexport
 - Para importar un valor especial de Salt:
manage_repository -importsalt contraseña nombresalt nombrearchivosalt
 - Para importar todos los valores de Salt:
manage_repository -importsalt -all contraseña archivoíndicesalt

Notas:

- Puede encontrar el nombre de Salt en el archivo g2.ini del repositorio.
- El nombre de Salt que importe debe ser el mismo que el nombre de Salt establecido en el archivo g2.ini del repositorio.

Tenga en cuenta que también debe añadir información acerca del servidor de Entity Analytics en la pestaña Entity Analytics del trabajo.

Encontrará más detalles relacionados con la configuración de Entity Analytics Server en la Guía de despliegue de IBM SPSS Modeler (*ModelerDeployment.pdf*).

Configuración de Analytical Decision Management para ejecutar rutas que contienen nodos de Generación de análisis de entidad

Después de haber utilizado el cliente del Gestor de despliegue de C&DS para añadir una definición de servidor Entity Analytics, tal como se ha explicado en Configuración de C&DS, no es necesario que haga nada más para habilitar Analytical Decision Management para ejecutar rutas.

Configuración de Solution Publisher

Cuando utilice Solution Publisher para ejecutar imágenes que contengan cualquier nodo EA en un repositorio, configure y administre Entity Analytics en cada máquina de Solution Publisher del modo siguiente:

- (1) Edite el script modelerrun, como en el ejemplo siguiente.

Nota: En este ejemplo, */opt/ibm/V10.5* es el directorio de instalación de DB2 y */home/g2user* es la carpeta de inicio del usuario de la instancia de DB2 que se ha utilizado durante la creación del repositorio.

```
DB2_INSTALL_PATH=/opt/ibm/V10.5/
export DB2_INSTALL_PATH
DB2_INSTANCE_HOME=/home/g2user
export DB2_INSTANCE_HOME
ICC_PKCS11_ROOT=<modeler>/ext/bin/pasw.entityanalytics/g2
export ICC_PKCS11_ROOT
```

Nota: Asegúrese de que tiene permisos de lectura/escritura/modificación para la carpeta ICC_PKCS11_ROOT.

- (2) Configure la instalación de Solution Publisher según las instrucciones para las máquinas satélite que se ofrecen en el apartado Configuración para un clúster de COP

Tenga en cuenta que en el ejemplo de las instrucciones se muestran rutas de Modeler Server, que deberá cambiar por rutas equivalentes para Solution Publisher.

- (3) Edite el archivo *<carpeta de instalación de solution publisher>/ext/bin/pasw.entityanalytics/ea.cfg*

Añada la línea siguiente:

```
service_management, 0
```

Administración de DB2 y servicios de EA

En la máquina de Modeler Server que aloja el repositorio, los procesos de DB2 y el servicio de EA se gestionan de forma independiente de Modeler Server. Se inician bajo petición, si no se están ya en funcionamiento cuando hay que ejecutar una ruta de Modeler, y permanecen en ejecución independientemente de si Modeler Server está en funcionamiento o se ha detenido. También un administrador puede iniciarlos, comprobarlos y detenerlos utilizando el script `manage_repository`.

El script `manage_repository` se puede utilizar en otras máquinas (satélites) para iniciar y detener el servicio de EA (en estas máquinas la instancia de DB2 no se está ejecutando de forma local y no se inicia ni se detiene). El servicio de EA no se inicia de forma automática.

Para iniciar el servicio de EA (y el servidor de DB2 si es necesario):

```
cd <directorio-instalador-modeler>\ext\bin\pasw.entityanalytics
manage_repository.bat|.sh -start <nombre-repositorio> <nombre usuario> <contraseña>
```

Ejemplo:

```
pasw.entityanalytics >./manage_repository.sh -start AAA myuser mypass
```

Iniciando servidor DB2...

Servidor DB2 iniciado

Iniciando servicio EA...

Para comprobar si el servicio de EA y DB2 se están ejecutando:

Invoque el script `manage_repository` con la opción de selección:

```
cd <directorio-instalador-modeler>\ext\bin\pasw.entityanalytics  
manage_repository.bat|.sh -check<nombre-repositorio> <nombre de usuario> <contraseña>
```

Ejemplo:

```
pasw.entityanalytics > ./manage_repository.sh -check AAA myuser mypass  
el servidor DB2 se está ejecutando en el host localhost, puerto 1320  
el servicio EA se está ejecutando en el host localhost, puerto 1321
```

Para detener el servicio de EA (y DB2 si es aplicable):

Invoque el script `manage_repository` con la opción de detención:

```
cd <directorio-instalación-modeler>\ext\bin\pasw.entityanalytics  
manage_repository.bat|.sh -stop <nombre-repositorio> <nombre usuario> <contraseña>
```

Ejemplo:

```
pasw.entityanalytics> ./manage_repository.sh -stop AAA myuser mypass  
Deteniendo servicio EA local  
Comando de conclusión enviado al servicio EA  
Deteniendo servidor DB2 local  
Comando de conclusión enviado a servidor DB2
```

Múltiples instancias de DB2

Si tiene varias instancias de DB2, para especificar qué instancia debe utilizarse en el repositorio debe establecer la variable de entorno `DB2INSTANCE` de `modelersrv.sh` en el nombre de la instancia necesaria.

Configuración para un clúster de COP

Elija una máquina del Modeler Server para alojar el repositorio de EA y vuelva a configurarlo de acuerdo con las instrucciones del apartado Modeler Server: Configuración de host de repositorio. Se hace referencia a esta máquina como la máquina host de Modeler Server.

Son necesarios los cambios de configuración siguientes para las máquinas de Modeler Server no host restantes, referidas como máquinas satélite:

(1) Cree una carpeta: `<carpeta EA>/repositories/<nombre repositorio>`.

(Donde `<nombre-repositorio>` es el nombre del repositorio de EA en la máquina Modeler Server del host.)

(2) Edite el archivo `<carpeta EA>/g2_config.xml` y añada una nueva instancia para cada repositorio.

Los componentes de EA utilizan este archivo para ubicar el servicio Web de g2 para un repositorio con nombre. Por ejemplo:

```
<instance external="false" g2host="localhost" g2port="1321" host="9.30.214.79"
name="AAA"
path="C:\ProgramData\IBM\SPSS\Modeler\18\EA\repositories\AAA\g2.ini"
port="1320"/>
```

La información más importante es:

- El atributo **name** (el nombre del repositorio) debe coincidir con el nombre del repositorio utilizado en la máquina Modeler Server del host.
- El atributo **path** (la ubicación del repositorio):
 - Windows:
C:\ProgramData\IBM\SPSS\Modeler\18\EA\repositories\<<nombre-repositorio>\g2.ini
 - UNIX:
<directorio-instalación-modeler>/ext/bin/pasw.entityanalytics/EA/repositories/<nombre-repositorio>/g2.ini
- Establezca el atributo **g2port** (que identifica el puerto en el cual escuchará el servicio g2) en un puerto que no se utilice actualmente en la máquina satélite.
- Establezca el atributo **g2host** en "localhost".
- Apunte el atributo **host** al nombre de la máquina host de Modeler Server.
- Establezca el atributo **port** en el número de puerto utilizado por el repositorio de EA en la máquina host de Modeler Server.

(3) Copie el archivo <directorio-instalación-modeler>/ext/bin/pasw.entityanalytics/templates/g2.ini en la nueva carpeta creada en el paso 1 y edite la nueva copia de g2.ini:

- Cambie las dos apariciones de {DataFolder} por la ruta de acceso completa a la carpeta:
<directorio-instalación-modeler>/ext/bin/pasw.entityanalytics/data.
- Cambie el valor de {Database} por el nombre del repositorio utilizado en la máquina host de Modeler Server.
- Cambie el valor de {Salt} por el valor de sal que se encuentra en el archivo g2.ini del repositorio utilizado en la máquina host de Modeler Server.

(4) Cambie el usuario por el que ha creado el repositorio (por ejemplo: g2user) y, después, utilice los comandos siguientes para catalogar el repositorio utilizado en la máquina host de Modeler Server.:

```
su - g2user
db2 catalog tcpip node <nombre del nodo> remote <IP máquina host> server 50000
```


db2 catalog database <Repositorio en máquina host de Modeler Server> at node <nombre del nodo>

- (5) Reinicie Modeler Server en la máquina satélite.
- (6) Conéctese a la máquina satélite de Modeler Server desde Modeler Client y conéctese al repositorio configurado.

El siguiente diagrama muestra la configuración de un clúster de COP de equipos que ejecutan Modeler Server.

Modeler Server / COP
(satellite machines)

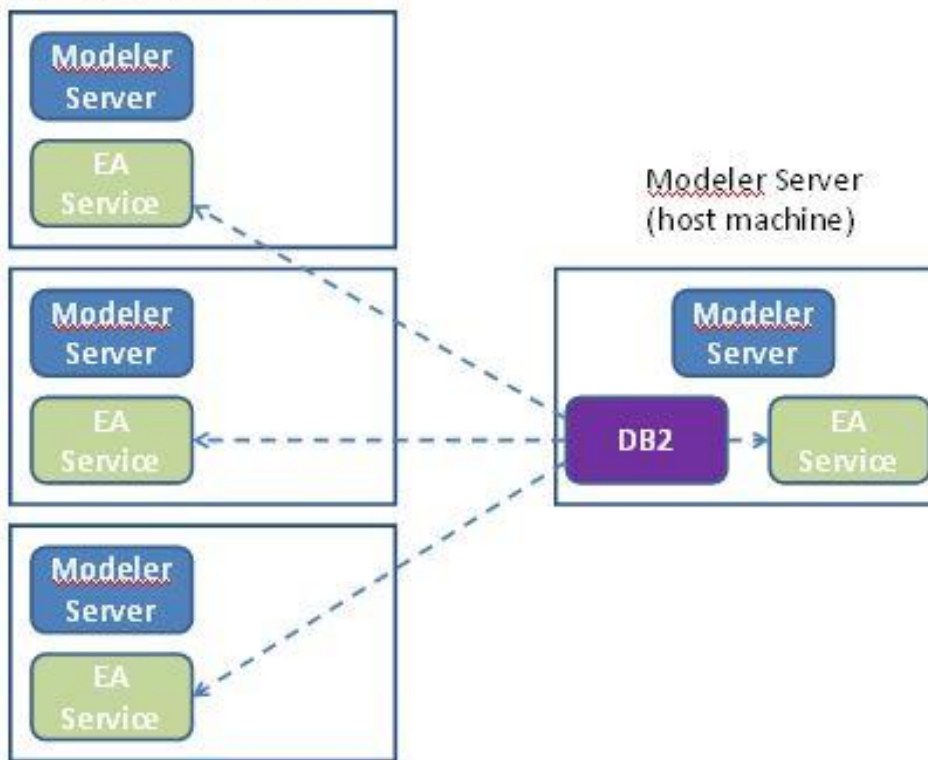


Figura 2 – Modeler Server – Configuración de máquinas en clúster COP