

IBM SPSS Modeler Entity Analytics - erweiterte Konfiguration

Einführung

Dieses Handbuch richtet sich an Systemadministratoren, die IBM SPSS Modeler Entity Analytics (EA) für die Ausführung mit einem der folgenden Produkte konfigurieren:

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

Übersicht

In diesem Handbuch wird erläutert, wie Instanzen von IBM SPSS Collaboration and Deployment Services, Analytical Decision Management, Solution Publisher bzw. Modeler Server, die in einer COP-Clusterkonfiguration (Coordinator of Processes) ausgeführt werden, für den Zugriff auf ein EA-Repository konfiguriert werden können.

Das EA-Repository muss auf einem Computer eingerichtet werden, der Modeler Server als Host für das Repository verwendet. ADM, Solution Publisher und C&DS-Jobs sowie der Scoring-Service können dann den Knoten "Streaming von EA" in Modeler verwenden, der zum Durchsuchen dieses Repositorys konfiguriert ist.

Jedes EA-Repository wird von einem einzelnen Datenbankserver (in Release 17 von Modeler ist dies beispielsweise ein IBM DB2-Datenbankserver) und genau einer Instanz des EA-Service, die auf jedem Computer ausgeführt wird, bedient.

Hinweis: Wenn Sie nur Modeler Server verwenden, treffen die meisten Informationen in diesem Handbuch nicht auf Sie zu. Allerdings wollen Sie die EA-Services möglicherweise beenden können. In diesem Fall würden Sie das Script `manage_repository` verwenden. Weitere Informationen finden Sie in `Verwalten von DB2 und EA-Services` .

Das folgende Diagramm zeigt die Konfiguration für einen Cluster von Computern, auf denen der C&DS-Scoring-Service ausgeführt wird.

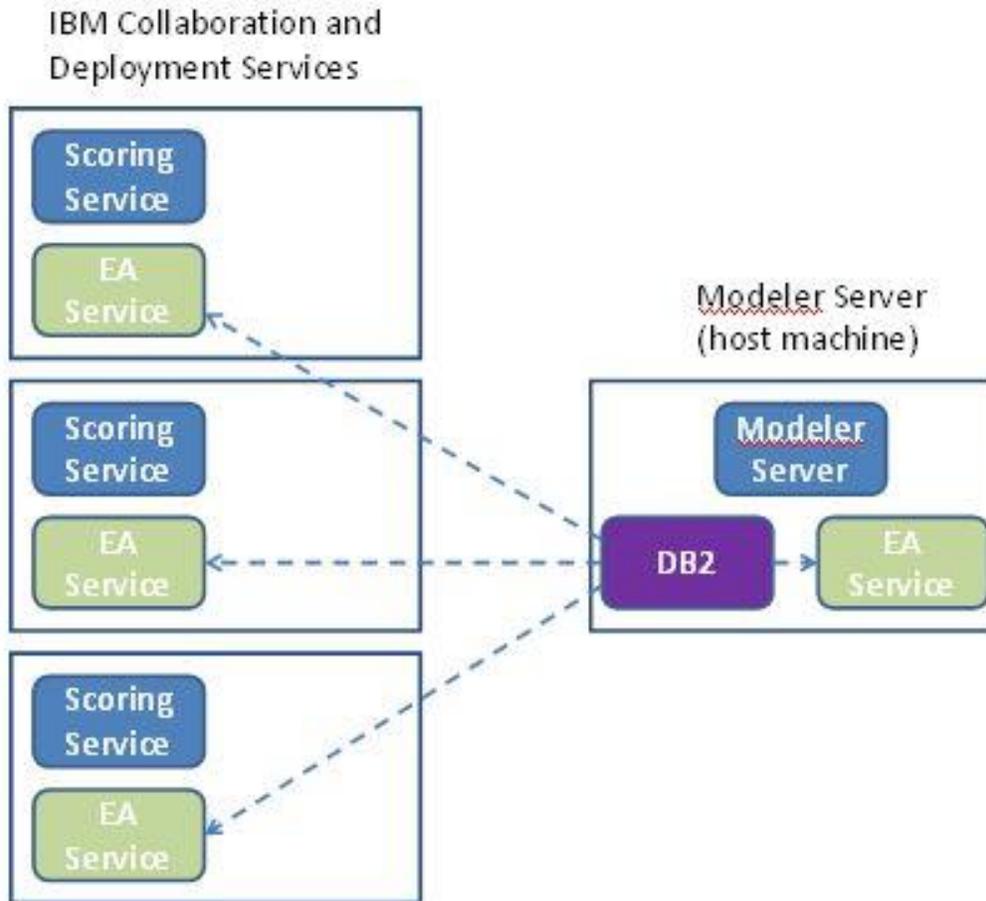


Abbildung 1 C&DS - Konfiguration von Cluster-Computern

Konfiguration

Notieren Sie sich zunächst den Speicherort des Ordners, in dem EA auf dem als Repository-Host verwendeten Modeler Server-Computer und den anderen Computern installiert ist. Sie benötigen diese Informationen während der Konfiguration.

In der folgenden Liste finden Sie die Standardspeicherorte für verschiedene Plattformen.

Windows Vista, Windows 7, Windows Server:

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

Für EA ist die Laufzeitbibliothek von Microsoft Visual C++ 2010 für Windows erforderlich. Installieren Sie erforderlichenfalls die Laufzeitbibliothek von Visual C++ 2010, bevor Sie EA-Funktionen ausführen. Sie kann über das Microsoft Download Center abgerufen werden.

UNIX:

<Modeler-Installationsverzeichnis>/ext/bin/pasw.entityanalytics/EA

Modeler Server - Konfiguration des Repository-Hosts

Es wird vorausgesetzt, dass Entity Analytics auf einem als Repository-Host verwendeten Modeler Server-Computer installiert ist, auf dem ein EA-Repository erstellt wurde. (Beachten Sie, dass dies ein lokaler oder ein ferner Computer sein kann.) Auf diesem Computer befindet sich das Repository, auf das andere Services zugreifen, beispielsweise:

- Collaboration and Deployment Services (C&DS)
- Decision Management
- Solution Publisher
- Andere Modeler Server-Instanzen, die in einem COP-Cluster (Coordinator of Processes) ausgeführt werden

Verwenden Sie das Script `manage_repository`, um Repository-Services auf dem als Repository-Host verwendeten Modeler Server-Computer zu starten und zu stoppen. Stellen Sie sicher, dass die Services eines bestimmten Repositories gestartet wurden, damit über C&DS, Analytical Decision Management oder Solution Publisher ausgeführte oder für das Scoring in C konfigurierte Streams erfolgreich auf das EA-Repository zugreifen können.

Weitere Informationen zum Verwenden des Scripts `manage_repository` finden Sie in `Verwalten von DB2 und EA-Services`.

C&DS-Konfiguration

Nach dem Konfigurieren des Repositories müssen Sie mit dem C&DS Deployment Manager-Client eine Entity Analytics-Serverdefinition hinzufügen. Dies ist erforderlich, um Echtzeitscoring für Entity Analytics zu aktivieren und um die Verwendung eines Modeler-Streams zu aktivieren, der einen EA-

Knoten in einem C&DS-Job enthält (um also EA-Ströme in C&DS auszuführen). Die Serverdefinition muss mit dem Repository-Namen im Stream übereinstimmen. Diese Definition teilt dem Stream mit, wo sich das Repository befindet, und liefert ihm die erforderlichen Verbindungsinformationen.

So erstellen Sie eine Entity Analytics-Serverdefinition:

- (1) Bearbeiten Sie das C&DS-Startscript `startServer.sh` und fügen Sie ihm wie im folgenden Beispiel gezeigt zwei Umgebungsvariablen hinzu.

Hinweis: In diesem Beispiel ist `/opt/ibm/V10.5` das DB2-Installationsverzeichnis und `/home/g2user` ist der Ausgangsordner der DB2-Instanz, das bzw. der während der Repository-Erstellung verwendet wurde.

```
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
```

Hinweis: Stellen Sie sicher, dass Sie Lese-/Schreib-/Bearbeitungsberechtigungen für den Ordner `ICC_PKCS11_ROOT` haben.

- (2) Starten Sie den C&DS-Server erneut.
- (3) Fügen Sie im Deployment Manager-Client eine neue Berechtigungsnachweisdefinition hinzu. Der Benutzername und das Kennwort müssen mit dem Berechtigungsnachweis übereinstimmen, der auf dem Modeler Server-Computer verwendet wird.
- (4) Fügen Sie im Deployment Manager-Client eine neue Serverdefinition hinzu:
 - a. Geben Sie den Servernamen ein. Dieser muss mit dem Namen des Repositorys übereinstimmen, das auf dem Modeler Server-Computer verwendet wird.
 - b. Wählen Sie **Entity Analytics Server** als Typ aus und klicken Sie auf **Next**.
 - c. Geben Sie den Verzeichnispfad oder die URL des Entity Analytics-Servers ein, zu dem die Verbindung hergestellt wird. Dies ist der in der Datei `g2.ini` angezeigte "connect"-Wert.
 - d. Geben Sie den Saltnamen für das Repository an, das für verschlüsselte Daten verwendet wird; dieser Wert wird in der Datei `g2.ini` angezeigt.
 - e. Wählen Sie den zu verwendenden Berechtigungsnachweis aus. Hierbei handelt es sich um den in Schritt 1 hinzugefügten Berechtigungsnachweis.

- f. Exportieren Sie den Saltwert aus Modeler Server und importieren Sie ihn in C&DS. Beachten Sie, dass Sie die Umgebungsvariable ICC_PKCS11_ROOT vor dem Export oder Import des Saltwerts über die Befehlszeile setzen/exportieren müssen.
- So exportieren Sie einen bestimmten Saltwert:
manage_repository -exportsalt Kennwort Salname Exportdateiname
 - So exportieren Sie alle Saltwerte:
manage_repository -exportsalt -all Kennwort Exportorder
 - So importieren Sie einen bestimmten Saltwert:
manage_repository -importsalt Kennwort Salname Saltdateiname
 - So importieren Sie alle Saltwerte:
manage_repository -importsalt -all Kennwort Saltindexdatei

Hinweise:

- Sie finden den Saltnamen in der Datei g2.ini des Repositorys.
- Der Saltnamen, den Sie importieren, muss mit dem Saltnamen übereinstimmen, der in der Datei g2.ini des Repositorys festgelegt ist.

Sie müssen die Entity Analytics-Serverinformationen auch der Entity Analytics-Registerkarte des Jobs hinzufügen.

Weitere Details zum Einrichten des Entity Analytics-Servers finden Sie im Bereitstellungshandbuch zu IBM SPSS Modeler (*ModelerDeployment.pdf*).

Konfiguration von Analytical Decision Management für die Ausführung von Streams, die Knoten des Typs "Streaming von EA" enthalten

Nachdem Sie mit dem C&DS Deployment Manager-Client eine Entity Analytics-Serverdefinition wie in *C&DS-Konfiguration* beschrieben hinzugefügt haben, sind keine weiteren Schritte erforderlich, um Analytical Decision Management für die Ausführung von Streams zu aktivieren.

Konfiguration für Solution Publisher

Wird Solution Publisher verwendet, um Images, die einen beliebigen EA-Knoten enthalten, für ein Repository auszuführen, müssen Sie Entity Analytics auf jedem Solution Publisher-Computer wie folgt konfigurieren und verwalten:

- (1) Bearbeiten Sie das Script modelerrun wie im folgenden Beispiel.

Hinweis: In diesem Beispiel ist */opt/ibm/V10.5* das DB2-Installationsverzeichnis und */home/g2user* ist der Ausgangsordner der DB2-Instanz, das bzw. der während der Repository-Erstellung verwendet wurde.

```
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
```

Hinweis: Stellen Sie sicher, dass Sie Lese-/Schreib-/Bearbeitungsberechtigungen für den Ordner ICC_PKCS11_ROOT haben.

- (2) Konfigurieren Sie die Installation von Solution Publisher wie in den Anweisungen für die Satellitencomputer im Abschnitt *Konfiguration für einen COP-Cluster* beschrieben.

Hinweis: Das Beispiel in den Anweisungen zeigt Modeler Server-Pfade. Ändern Sie diese in die entsprechenden Pfade für Solution Publisher.

- (3) Bearbeiten Sie die Datei *<Solution Publisher-Installationsordner>/ext/bin/pasw.entityanalytics/ea.cfg*.

Fügen Sie die folgende Zeile hinzu:

```
service_management, 0
```

Verwalten von DB2 und EA-Services

Auf dem als Repository-Host verwendeten Modeler Server-Computer werden die Prozesse von DB2 und der EA-Services getrennt von Modeler Server verwaltet. Sofern sie nicht bereits ausgeführt werden, werden sie auf Anforderung gestartet, wenn sie zur Ausführung eines Modeler-Streams benötigt werden, und sie laufen unabhängig davon weiter, ob Modeler Server ausgeführt wird oder angehalten wurde. Außerdem können sie von einem Administrator mithilfe des Scripts `manage_repository` gestartet, überprüft und gestoppt werden.

Das Script `manage_repository` kann auf den anderen Computern (sogenannten Satellitencomputern) verwendet werden, um den EA-Service zu starten und zu stoppen (auf diesen Computern wird die DB2-Instanz nicht lokal ausgeführt und wird nicht gestartet oder gestoppt). Der EA-Service wird nicht automatisch gestartet.

So starten Sie den EA-Service (und gegebenenfalls den DB2-Server):

```
cd <Modeler-Installationsverzeichnis>\ext\bin\pasw.entityanalytics
manage_repository.bat|.sh -start <Repository-Name> <Benutzername> <Kennwort>
```

Beispiel:

```
pasw.entityanalytics >./manage_repository.sh -start AAA myuser mypass
Starting DB2 server...
DB2 server started
Starting EA service...
```

So überprüfen Sie, ob der EA-Service und DB2 ausgeführt werden:

Rufen Sie das Script `manage_repository` mit der Option `-check` auf:

```
cd <Modeler-Installationsverzeichnis>\ext\bin\pasw.entityanalytics  
manage_repository.bat|.sh -check<Repository-Name> <Benutzername> <Kennwort>
```

Beispiel:

```
pasw.entityanalytics > ./manage_repository.sh -check AAA myuser mypass  
DB2 server is running on host localhost, port 1320  
EA service is running on host localhost, port 1321
```

So stoppen Sie den EA-Service (und gegebenenfalls DB2):

Rufen Sie das Script `manage_repository` mit der Option `-stop` auf:

```
cd <Modeler-Installationsverzeichnis>\ext\bin\pasw.entityanalytics  
manage_repository.bat|.sh -stop <Repository-Name> <Benutzername> <Kennwort>
```

Beispiel:

```
pasw.entityanalytics> ./manage_repository.sh -stop AAA myuser mypass  
Stopping local EA service  
Shutdown command sent to EA service  
Stopping local DB2 server  
Shutdown command sent to DB2 server
```

Mehrere DB2-Instanzen

Wenn Sie mehrere DB2-Instanzen haben und die Instanz angeben wollen, die für das Repository verwendet werden soll, setzen Sie die Umgebungsvariable `DB2INSTANCE` in `modelersrv.sh` auf den Namen der erforderlichen Instanz.

Konfiguration für einen COP-Cluster

Wählen Sie einen Modeler Server-Computer als Host für das EA-Repository aus und konfigurieren Sie ihn entsprechend den Anweisungen in *Modeler Server - Konfiguration des Repository-Hosts*. Dieser Computer wird als Modeler Server-Host-Computer bezeichnet.

Die folgenden Konfigurationsänderungen sind für die übrigen Modeler Server-Computer erforderlich, die nicht als Host dienen und als Satellitencomputer bezeichnet werden:

(1) Erstellen Sie einen Ordner: `<EA-Ordner>/repositories/<Repository-Name>`.

(Dabei ist `<Repository-Name>` der Name des EA-Repositorys auf dem Modeler Server-Host-Computer.)

(2) Bearbeiten Sie die Datei `<EA-Ordner>/g2_config.xml` und fügen Sie für jedes Repository eine neue Instanz hinzu.

Diese Datei wird von den EA-Komponenten verwendet, um den g2-Web-Service für ein angegebenes Repository ausfindig zu machen. Beispiel:

```
<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"/>
```

Wichtige Informationen:

- Das Attribut **name** (Repository-Name) muss mit dem Namen des Repositories übereinstimmen, das auf dem Modeler Server-Computer verwendet wird.
- Das Attribut **path** (die Repository-Position):
 - Windows:
C:\ProgramData\IBM\SPSS\Modeler\18\EA\repositories\<<Repository-Name>\g2.ini
 - UNIX:
<Modeler-Installationsverzeichnis>/ext/bin/pasw.entityanalytics/EA/repositories/<Repository-Name>/g2.ini
- Setzen Sie das Attribut **g2port** (gibt den Port an, den der g2-Service als Port überwacht) auf einen Port, der derzeit nicht auf dem Satellitencomputer verwendet wird.
- Setzen Sie das Attribut **g2host** auf "localhost".
- Setzen Sie das Attribut **host** auf den Namen des Modeler Server-Host-Computers.
- Setzen Sie das Attribut **port** auf die Portnummer, die das EA-Repository auf dem Modeler Server-Host-Computer verwendet.

(3) Kopieren Sie die Datei <Modeler-

Installationsverzeichnis>/ext/bin/pasw.entityanalytics/templates/g2.ini in den neuen Ordner, der in Schritt (1) erstellt wurde, und bearbeiten Sie die neue Kopie von g2.ini:

- Ändern Sie die beiden Vorkommen von {DataFolder} in den vollständigen Pfad des Ordners: <Modeler-Installationsverzeichnis>/ext/bin/pasw.entityanalytics/data.
- Ändern Sie {Database} in den Namen des Repositories, das auf dem Modeler Server-Host-Computer verwendet wird.
- Ändern Sie {Salt} in den Saltwert aus der Datei g2.ini des Repositories, das auf dem Modeler Server-Host-Computer verwendet wird.

- (4) Ändern Sie den Benutzer in den Benutzer, der das Repository erstellt hat (z. B. g2user), und katalogisieren Sie dann das Repository, das auf dem Modeler Server-Host-Computer verwendet wird, mit den folgenden Befehlen:

```
su - g2user
```

```
db2 catalog tcpip node <Name Ihres Knotens> remote <IP des Host-Computers> server 50000
```

```
db2 catalog database <Repository auf Modeler Server-Host-Computer> at node <Name Ihres Knotens>
```

- (5) Starten Sie Modeler Server auf dem Satellitencomputer erneut.
- (6) Stellen Sie von Modeler Client eine Verbindung zum Modeler Server-Satellitencomputer her und stellen Sie dann eine Verbindung zum konfigurierten Repository her.

Das folgende Diagramm zeigt die Konfiguration für einen COP-Cluster von Computern, auf denen Modeler Server ausgeführt wird.

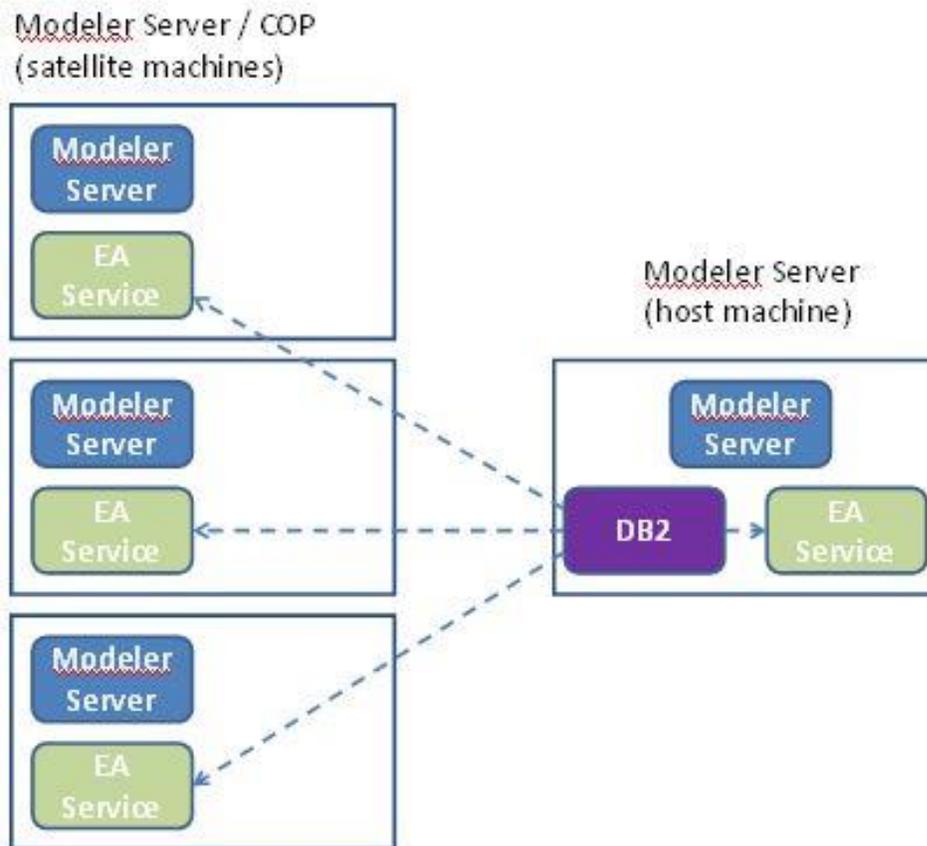


Abbildung 2 Modeler Server - Konfiguration von COP-Cluster-Computern