

IBM SPSS Modeler Entity Analytics

高级配置

简介

本指南的目标读者为系统管理员，他们配置 IBM SPSS Modeler Entity Analytics (EA) 以部署在下列某个系统中：

- IBM SPSS Collaboration and Deployment Services (C&DS) V6
- IBM SPSS Modeler Server
- IBM Analytical Decision Management (ADM) V8
- IBM SPSS Modeler Solution Publisher

概述

本指南介绍如何配置使用 Coordinator of Processes (COP) 聚类配置运行的 IBM SPSS Collaboration and Deployment Services、Analytical Decision Management、Solution Publisher 或 Modeler Server 以访问 Entity Analytics 存储库。

必须在使用 Modeler Server 托管存储库的机器上设置 EA 存储库。然后，ADM、Solution Publisher 及 C&DS Jobs and Scoring Service 可以使用在 Modeler 中配置为搜索此存储库的流 EA 节点。

每个 EA 存储库由单一的数据库服务器（例如，在 Modeler 16 发行版中为 IBM DB2 数据库服务器）以及每台机器上运行的某个 EA 服务实例提供服务。

下图显示运行 C&DS 评分服务的机器聚类的配置。

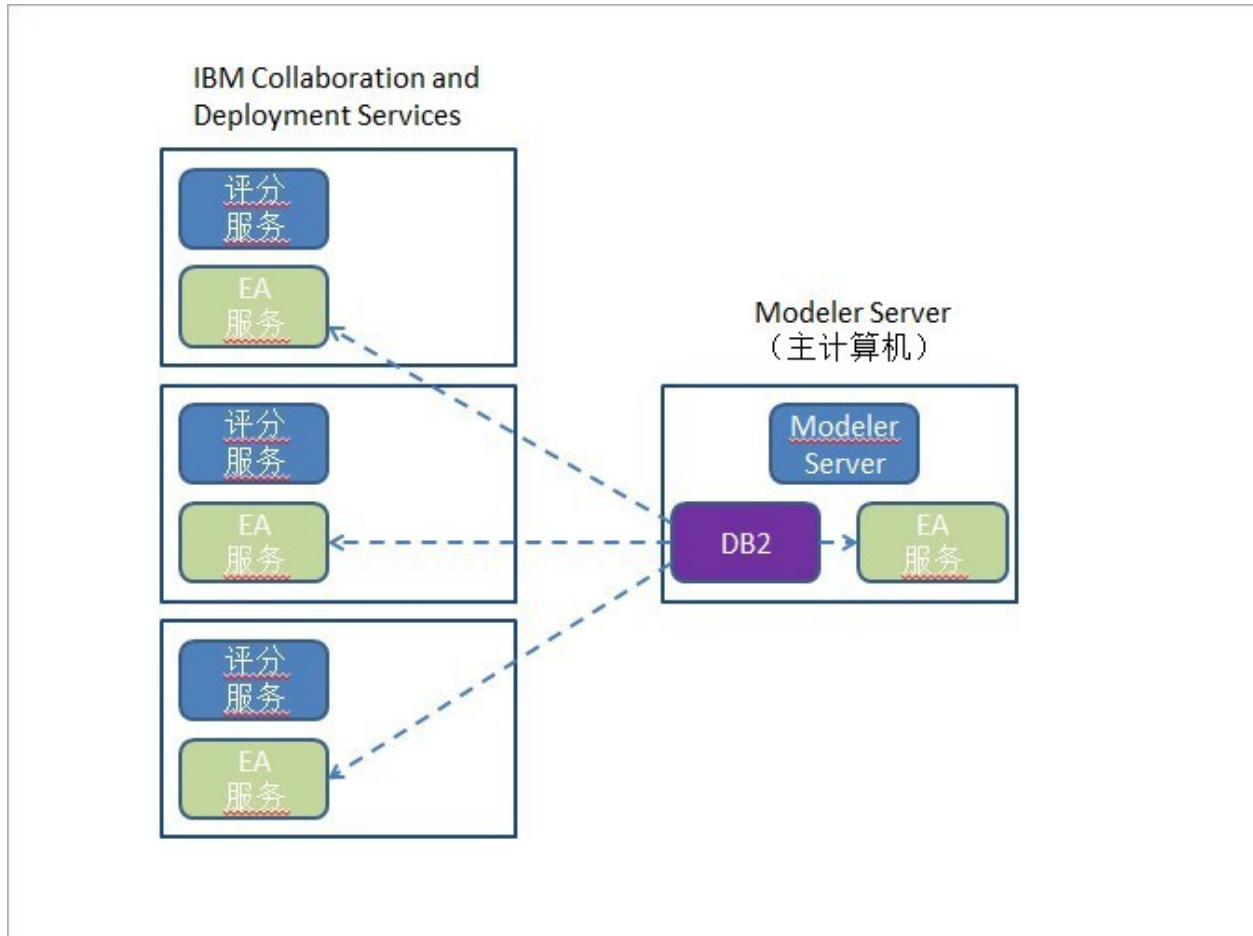


图 1 - C&DS - 聚类机器的配置

配置

第一步，记下存储库主机 Modeler Server 和每台其他机器上的 EA 安装文件夹的位置；您在进行配置时需要此信息。

以下列表显示了不同平台的缺省位置。

Windows Vista、Windows 7、Windows Server:

C:\ProgramData\IBM\SPSS\Modeler\16\EA

UNIX:

<modeler-install-directory>/ext/bin/pasw.entityanalytics/EA

Modeler Server - 存储库主机配置

作为先决条件，Entity Analytics 必须安装到存储库主机 Modeler Server 和在此机器上创建的 Entity Analytics 存储库上。此机器托管从其他服务访问的存储库，例如：

- Collaboration and Deployment Services (C&DS)
- Decision Management
- Solution Publisher
- 在 Coordinator of Processes (COP) 聚类中运行的其他 Modeler Server。

使用 `manage_repository` 脚本在存储库主机 Modeler Server 上启动和停止存储库服务。请确保已启动特定存储库的服务，这样从 C&DS、Analytical Decision Management 或 Solution Publisher 执行的流或为 C&DS 中的评分进行配置的流才能成功访问 Entity Analytics 存储库。

有关如何使用 `manage_repository` 脚本的更多信息，请参阅 管理 DB2 和 EA 服务。

C&DS 配置

在配置存储库后，您必须使用 C&DS Deployment Manager 添加 Entity Analytics 服务器。为此：

- (1) 在 Deployment Manager 中，添加新凭证；用户名和密码必须与 Modeler Server 机器上使用的凭证匹配。
- (2) 在 Deployment Manager 中，添加 Entity Analytics 服务器详细信息：
 - a. 输入服务器名称；必须与 Modeler Server 机器上使用的存储库的名称匹配。
 - b. 选择类型；您可以选择使用主机和端口或使用连接字符串。

- c. 如果选择使用主机和端口，请输入服务器所在的主机名和要用于连接到该主机的端口号。端口必须与 Modeler Server 机器上的存储库的端口匹配。
- d. 如果选择使用连接字符串，请输入与其建立连接的 Entity Analytics 服务器的目录路径或 URL。必须与 Modeler Server 机器上的存储库的连接字符串匹配。
- e. 选择要使用的凭证；这是您在第 1 步中添加的凭证。

IBM SPSS Modeler Deployment Guide (*modeler_deployment_book.pdf*) 中提供了与 Entity Analytics Server 设置相关的更多详细信息。

Analytical Decision Management 用于运行包含流 EA 节点的流的配置

使用 C&DS Deployment Manager 添加 Entity Analytics 服务器后（如 C&DS 配置中所述），无需执行其他步骤即可启用 Analytical Decision Management 来运行流。

Solution Publisher 用于运行包含流 EA 节点的流的配置

使用 Solution Publisher 对存储库运行包含流 EA 节点的映像时，请配置和管理每台 Solution Publisher 机器上的 Entity Analytics，如下所示：

- (1) 根据 配置 COP 聚类 部分中针对附属机器的指示信息配置 Solution Publisher 安装。

请注意，指示信息中的示例显示 Modeler Server 路径 - 更改这些路径以使用 Solution Publisher 的等效路径。

- (2) 编辑文件 `<solution_publisher_install_folder>/ext/bin/pasw.entityanalytics/ea.cfg`

添加以下行：

```
service_management, 0
```

- (3) 使用 `manage_repository` 脚本在 Solution Publisher 机器上启动和停止存储库服务。请确保已启动特定存储库的服务，这样从 Solution Publisher 执行的映像才能成功访问 Entity Analytics 存储库。

有关如何使用 `manage_repository` 脚本的更多信息，请参阅 管理 DB2 和 EA 服务。

管理 DB2 和 EA 服务

在存储库主机 Modeler Server 上，DB2 与 EA 服务进程是从 Modeler Server 单独管理的。在要求执行 Modeler 流时，如果它们未处于运行状态，那么将根据需要启动，并且无论 Modeler Server 是否运行，它们都将保持独立运行。管理员也可以使用 `manage_repository` 脚本来启动、检查和停止这些进程。

manage_repository 脚本必须用于其他（附属）机器上以启动和停止 EA 服务（在这些机器上，DB2 实例不在本地运行，因而无法启动或停止）。EA 服务不会自动启动。

要启动 EA 服务和 DB2 服务器（如果适用）：

```
cd <modeler-install-directory>\ext\bin\pasw.entityanalytics  
manage_repository.bat|.sh -start <repository-name> <username> <password>
```

示例：

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

要检查 EA 服务器与 DB2 是否处于运行状态：

请使用 -check 选项调用 manage_repository 脚本：

```
cd <modeler-install-directory>\ext\bin\pasw.entityanalytics  
manage_repository.bat|.sh -check <repository-name> <username> <password>
```

示例：

```
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
```

要停止 EA 服务和 DB2（如果适用）：

请使用 -stop 选项调用 manage_repository 脚本：

```
cd <modeler-install-directory>\ext\bin\pasw.entityanalytics  
manage_repository.bat|.sh -stop <repository-name> <username> <password>
```

示例：

```
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
```

配置 COP 聚类

选择某个 Modeler Server 机器以托管 EA 存储库并根据 Modeler Server - 存储库主机配置中的指示信息对其进行重新配置。该机器被称为主机 Modeler Server。

对于被称为附属机器的其余非主机 Modeler Server 机器，需要以下配置更改：

(1) 创建文件夹: `<EA folder>/repositories/<repository-name>`。

(其中 `<repository-name>` 为主机 Modeler Server 上的 EA 存储库的名称。)

(2) 编辑文件 `<EA folder>/g2_config.xml` 并为每个存储库添加新的实例。

EA 组件使用此文件来查找指定存储库的 g2 Web 服务。例如:

```
<instance external="false" g2host="localhost" g2port="1321" host="9.30.214.79"
name="AAA"
path="C:\ProgramData\IBM\SPSS\Modeler\16\EA\repositories\AAA\g2.ini"
port="1320"/>
```

其中的重要信息为:

- **name** 属性 (存储库名称) 必须与主机 Modeler Server 上使用的存储库的名称匹配。
- 将 **g2port** 属性 (用于标识 g2 服务将侦听的端口) 设置为附属机器上当前未使用的端口。
- 将 **g2host** 属性设置为 “localhost”。
- 将 **host** 属性指向主机 Modeler Server 的名称。
- 将 **port** 属性设置为 Modeler Server 主机上 DB2 使用的端口号。

(3) 将 `<modeler-install-directory>/ext/bin/pasw.entityanalytics/templates/g2.ini` 文件复制到在第 (1) 步中创建的新文件夹, 然后编辑 `g2.ini` 的新副本:

- 将两处 `{DataFolder}` 更改为该文件夹的完整路径: `<modeler-install-directory>/ext/bin/pasw.entityanalytics/data`。
- 将 `{Database}` 更改为 Modeler Server 主机上使用的存储库的名称。
- 将 `{Salt}` 更改为在 Modeler Server 主机上使用的存储库的 `g2.ini` 文件中找到的 Salt 值

(4) 使用以下命令对 Modeler Server 主机上使用的存储库进行编目:

```
catalog tcpip node <Your Node Name> remote <host machine IP> server 50000
db2 catalog database <Repository on Host Modeler Server Machine> at node <Your Node Name>
```

(5) 重新启动附属机器上的 Modeler Server。

(6) 启动 Modeler Server 附属机器上的存储库。

(7) 从 Modeler Client 连接到 Modeler Server 附属机器并连接到已配置的存储库。

下图显示运行 Modeler Server 的机器的 COP 聚类配置。

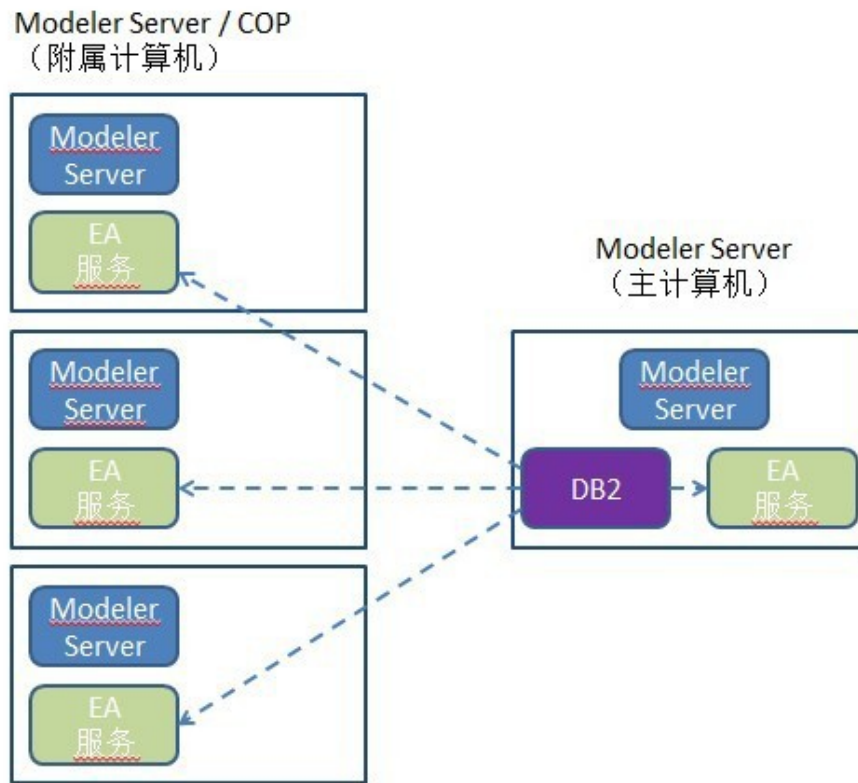


图 2 - Modeler Server - COP 聚类机器的配置