

Data Services 入门指南

第一章 Data Services产品套件概述

第一节 Data Services 和 Business Objects 产品套件

Business Objects产品套件通过专业的终端用户工具在一个单独的，可信的商务智能平台上提供了非常深刻的研究和探讨。整个平台由Data Services支持。在Data Services的顶层，Business Objects设计了最可靠的，扩展的和可管理的BI平台，该平台支持业界集成度最好的报表，查询和分析，绩效管理仪表盘，记分板 and 应用程序。真正的数据集成组合了批量抽取，转换及加载(ETL)技术和跨越多个扩展的企业应用之间的实时双向的数据流。通过创建一个关系数据仓库并智能地组合使用对来自企业ERP系统和其他数据源的数据的直接实时访问和批量数据访问方法，BusinessObjects开发了一个功能强大的，高性能数据集成的产品,它允许充分平衡你的ERP和企业应用的基础设施以满足多种业务使用。

Business Objects提供了一个批量和实时数据集成系统来驱动目前新一代的分析和供应链管理应用。使用Business Objects高可扩展性的数据集成解决方案，企业可以维护与客户，供应商，员工和合作方的一个实时的在线的对话，给他们提供交易和业务分析需要的关键信息。

第二节 Data Services产品优点

使用Data Services来开发批量和实时使用的企业数据集成。通过Data Services:

- ◆ 可以创建一个单独的用于批量和实时数据移动的基础架构从而获得一个更快的更低成本的实现。
- ◆ 企业可以在独立于任何单独的系统的的前提下将数据作为企业资产来管理。集成多个系统的数据并将这些数据重新用于许多其他目的。
- ◆ 可以选择使用预打包的数据解决方案已获得快速的部署和快的投资回报。这些解决从业务操作系统抽取历史的和日常的数据并在一个开放的关系数据库中缓存这些数据。

Data Services定制和管理数据访问并综合业界领先的，专利的技术来给业务分析，供应链管理，客户关系管理和Web应用提供数据

平台统一

Data Services提供了几点平台统一：

- ◆ 获取端到端的数据家谱和影响分析
- ◆ 创建语义层(Universe)和在ETL设计环境中管理变化

Business Objects 将整个ETL过程和商务智能平台进行深度集成，这样做可以获得下列好处：

- ◆ 简单的元数据管理
- ◆ 简化的统一的系统管理
- ◆ 生命周期管理
- ◆ 可以信赖的信息

易用性和高的生产效率

Data Services 为了给来自任何信息源用于任何信息使用目的信息管理提供一个单一的数据集成平台，Data Services综合了批量和实时的数据的移动和管理。

使用Data Services可以：

- ◆ 在操作数据库，数据仓库或者数据集市存储数据
- ◆ 以批量或者实时的方式更新存储的数据
- ◆ 为开发，测试和部署整个数据集成平台提供一个单一的图形开发环境
- ◆ 为捕捉不同抽取和访问方法之间的关系管理一个单独的元数据资源库并提供集成的宗谱和影响分析

高可用性和性能

Data Services的高性能引擎和经过验证的数据移动和管理能力包括：

- ◆ 可扩展的，多实例的快速执行的数据移动
- ◆ 负载均衡
- ◆ 变化数据的捕捉
- ◆ 并行处理

第三节 Data Services相关的产品

选择几个Business Objects的产品的可选产品来进一步支持和改进你的Data Services产品的处理能力。

BusinessObjects Composer

即使在开始一个数据仓库的实现之前，Business Objects 已经知道你需要收集和来自信息用户的关键信息以便“构成”一个蓝本，从这个蓝本可以开始创建你的Data Services作业。BusinessObjects Composer是一个独立的基于Web的用于设计，转换和加载项目的应用程序。给ETL作业开发高级的设计可以使实现处理进程更加高效并改进由此开发的作业的质量。Composer 是对其他ETL工具的补充，它提供了一个在开始创建数据仓库之前创建你的设计的一个平台。使用Composer来识别源数据，实现转换和编写设计文档

BusinessObjects元数据管理

BusinessObjects元数据管理提供了跨越部分和全部Business Objects产品系列的完整的商务智能项目的元数据和它的多个关系的一个集成的视图。使用元数据管理可以：

- ◆ 从一个单独的资源库来查看关于Business Objects报表,文档和数据源
- ◆ 分析宗谱来确定Business Objects文档和报表的数据源
- ◆ 分析原表，数据列，元素或者字段的修改会对现存的Business Objects文档和报表产生的影响
- ◆ 跟踪随着时间的推移每个对象的不同修改版本
- ◆ 通过一个DateTime来将操作元数据(如处理的数据行数和CPU利用率)作为历史数据来查看
- ◆ 用不同的语言来查看元数据

第四节 Data Services接口

Data Services 提供了许多类型的接口组件。你的Data Services版本可能提供这些接口的全部或者只是其中的一部分。

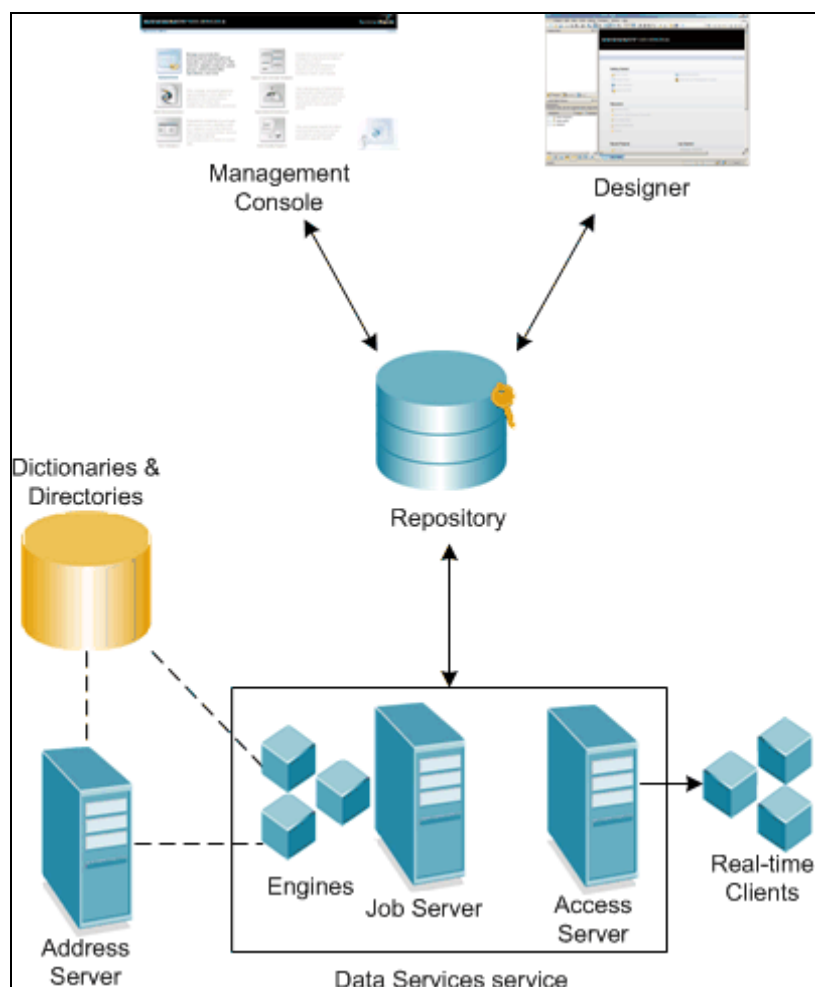
可以使用Data Services接口开发工具包来开发从其他应用程序读取或者向其他应用程序写入数据的适配器。除了上面列出的接口，Data Services嵌入的关系数据模型允许你完全利用SQL转换的处理能力来操作，处理和丰富分层次的业务文档。关于支持的环境和硬件需求的详细信息参见Business Objects支持站点<https://service.sap.com/bosap-support>中的“支持的平台”文档。该文档包括数据库应用程序，Web应用程序，Web浏览器和操作系统的特定的版本和补丁级别要求。

第二章 Data Services架构

Data Services的架构是分层设计的以允许支持各种不同的开放的业界标准的API做得数据集成，从而获得对数据和元数据的更加优化的管理。

第一节 标准的Data Services组件

下图概述了Data Services的组件之间的关系：



Data Services设计器

Designer是一个具有易于使用的图形用户界面的开发工具。它允许开发人员定义包括数据映射，转换和控制逻辑的数据管理应用程序。使用Designer来创建包含工作流(作业执行定义)和数据流（数据转换定义）的应用程序。使用Designer来创建对象，然后通过选择流图，表布局 and 嵌入的工作空间页中选择图标来拖放并配置它们。Designer中的对象代表了元数据。Designer接口允许管理存储在Data Services资源库中的元数据。从Designer中可以触发Data Services作业服务器来执行你的作业以便进行应用程序的初始测试。

Data Services资源库

Data Services资源库是一套存储用户创建的和预定义的系统对象，源和目标元数据，和转换规则的表。在一个开放的客户端、服务器平台上设置资源库来适应与其他企业工具之间共享元数据。将每个资源库存储在一个现存的RDBMS上。

每个资源库与一个或者多个运行你创建的作业的数据服务作业服务器相关联。有两种类型的资源库：

- ◆ 应用程序设计器使用的本地资源库用来存储Data Services对象(如项目，作业，工作流，和数据流)的定义和源和目标的元数据。
- ◆ 中央资源库是可以用来支持多用户开发的一个可选的组件。中央资源库提供了允许签入和迁出本地资源库的对象的共享的对象库。

Data Services作业服务器

Data Services 作业服务器启动数据移动的从多个不同种类的源集成数据的引擎，执行复杂的数据转换，并管理从ERP系统和其他源的抽取和事务。Data Services作业服务器可以以批量和实时模式移动数据并使用分布式的查询优化，多线程，内存缓存，内存数据转换和并行处理来提供高的数据吞吐量和可扩展性。

在设计一个作业时可以从设计器中运行这个作业，设计器通知作业服务器来运行这个作业。作业服务器从相关的资源库调出这个作业然后启动Data Services引擎来处理作业。在生产环境中，作业服务器通过一个调度程序或者通过一个由Data Services访问服务器管理的实时的服务触发一个作业的执行。在生产环境中，可以通过创建一个按照整体系统负载来执行作业的作业服务器租来

平衡作业负载。

Data Services引擎

在Data Services作业执行时，作业服务器启动Data Services引擎进程来执行数据抽取，转换和移动。Data Services引擎进程使用并行的处理和内存数据转换来获得高的数据吞吐量和可扩展性。

Data Services访问服务器

访问服务器是一个实时的请求应答的消息代理，它收集消息请求，将它们路由给一个实时的服务并在一个用户特定的时间帧内提交一个消息应答。访问服务器将消息进行排队并在任何数量的计算资源之间把它们发送下一个可用的实时服务。这种方法提供了自动的可扩展性，因为如果给定的实时服务的吞吐量比较高时，访问服务器可以启动位于另外的计算资源上的实时服务。你可以配置多个访问服务器。

Data Services地址服务器

Data Services地址服务器提供了对Global Address Cleanse EMEA引擎和Global Suggestions List的地址验证和修正。地址服务器必须在处理包含通过EMEA引擎激活的Global Suggestion List转换或者Global Address Cleanse转换的数据流之前启动。

Data Services管理员

管理员(Administrator)提供对Data Services资源的基于浏览器的管理。这些Data Services资源管理包括：

- ◆ 调度，监视和执行批量作业
- ◆ 配置，启动和停止实时服务
- ◆ 配置作业服务器，访问服务器和资源库的使用
- ◆ 配置和管理适配器
- ◆ 管理用户
- ◆ 通过Web服务公布批量作业和实时服务

Data Services元数据报表应用程序

元数据报表应用程序提供对与下列内容相关联的元数据的基于浏览器的分析和报表功能：

- ◆ 你的Data Services作业
- ◆ 其他与Data Services相关联的其他Business Objects应用

元数据报表为管理你的元数据提供了四个应用程序：

1) 影响和宗谱分析报告

影响和宗谱分析报告包括：

- ◆ 数据仓库分析 — 针对每个数据连接，查看概述，表，函数，和层次报表。Data Services用户可以确定下列事项：
 - ✓ 有什么数据源组装成它们的表

- ✓ 他们的表组装成了什么样的目标表
- ✓ 是一个还是多个下列Business Objects报告使用了来自他们的表的数据：
 - ◆ 商务视图
 - ◆ 水晶报表
 - ◆ Universes
 - ◆ Web Intelligence文档
 - ◆ Desktop Intelligence文档
- ◆ Universe分析 — 查看Universe，类和对象宗谱。Universe用户可以确定有什么数据源组装成他们的Universe，以及哪些报表使用了他们的Universe
- ◆ 商务视图分析 — 查看中央管理服务器中的商务视图。可以查看每个商务视图的商务元素和商务字段的宗谱报告。水晶报表商务视图用户可以确定由什么数据源组装成他们的商务视图及什么报告使用了他们的商务视图
- ◆ 报表分析 — 查看中央管理服务器中的报表的数据源。可以查看CMS管理的每个水晶报表和Web Intelligence文档的数据库表和数据列的宗谱报表。报表编写人员可以确定有什么数据源组装成他们的报表
- ◆ 依赖性分析— 在资源库搜索特定的对象并了解这些对象如何影响其他的Data Services或Business Objects Universe对象和报表，或者如何被其他的Data Services或Business Objects Universe对象

和报表影响。元数据搜索结果提供追溯到关联报表的链接

为了查看Business Objects应用程序的影响和宗谱分析，必须配置元数据集成器。

2) 操作仪表盘报表

操作仪表盘报表提供Data Services作业执行的图形化的解释。该反馈信息可以一眼就看出一个给定的时间段内一个或者多个资源库的作业执行的状态和性能。有了这些信息，就可以使你的作业调度和管理流水流水作业并监视作业的调度和管理以便获得整体的效率和性能。

3) 自动编制文档报告

自动编制文档报表提供一个创建你在Data Services中创建的所有对象的打印文档的一个方便综合的方法。自动编制文档报表捕捉理解你的Data Services作业的关键信息，这样就可一眼看出整个ETL的处理过程。

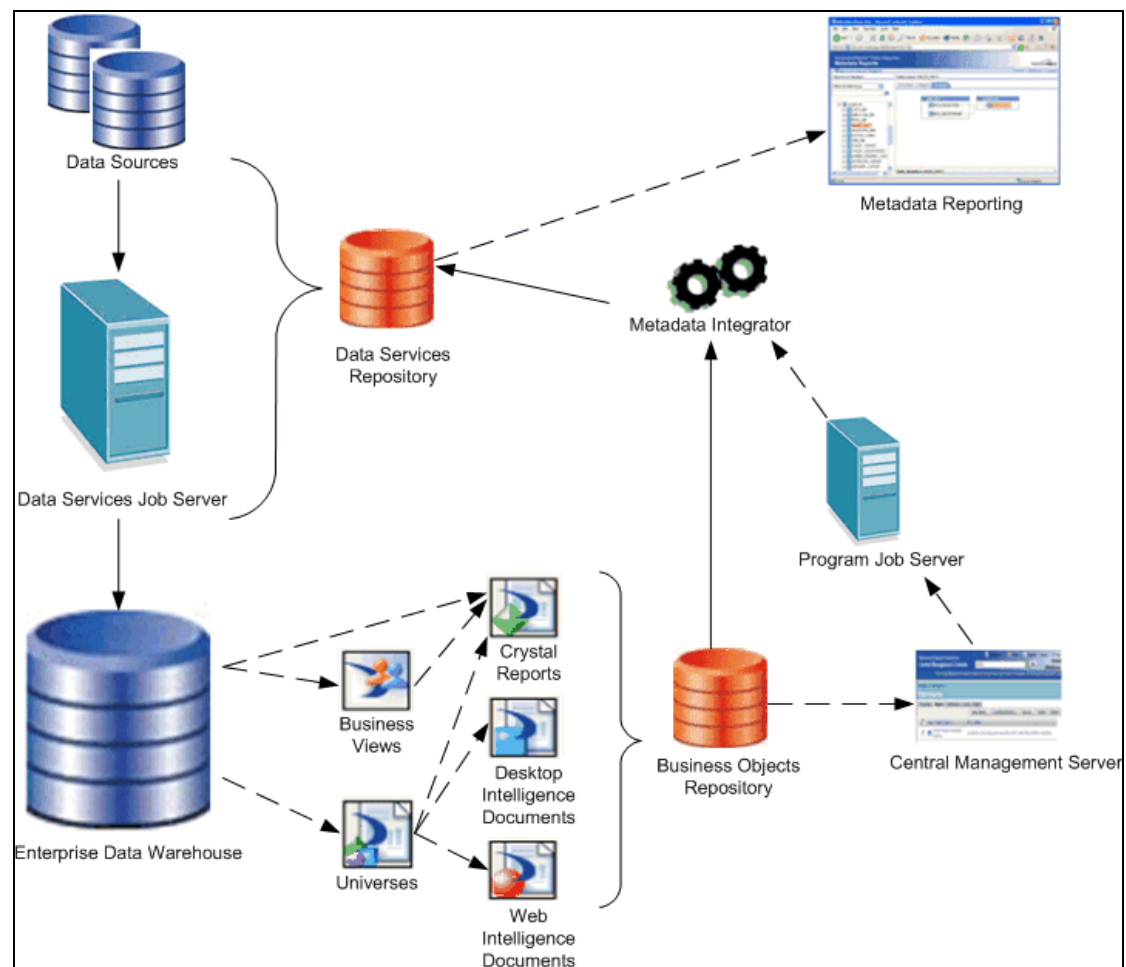
创建一个项目后可以利用自动编制文档报表快速创建一个PDF或Microsoft Word文件，该文件捕捉了包含图形展示和关键映射细节的选择的作业， workflow，和/或数据流信息。

4) 数据验证仪表盘

数据验证仪表盘报表提供根据你在你的Data Services批量作业上创建的验证规则来评估目标数据的可靠性的图形解释。该反馈信息允许商业用户快速地检查，评估，和识别元数据中的潜在的不一致性和错误。

Data Services元数据集成器

元数据集成器允许Data Services与Business Objects智能产品之间的无缝共享。运行元数据集成器为水晶报表，Desktop Intelligence文档和Web Intelligence文档使用的商务视图和Universe将原数据收集到Data Services的资源库。



Data Services服务

Data Services服务在Data Services作业和访问服务器安装时安装。Data Services在重新启动系统时启动作业服务器和访问服务器。Windows 服务名是Data Services Service。UNIX对等的是一个叫做AL_JobService的后台程序。

Data Services SNMP代理

Data Services错误事件可以使用简单网络管理协议(SNMP)支持的应用程序来通信以便更好地监视错误。在运行作业服务器的任何计算机上安装一个Data Services SNMP 代理。Data Services SNMP 代理监视和记录运行在安装了代理的计算机上的作业服务器和作业的有关信息。可以配置网络管理软件(NMS)应用程序与Data Service的SNMP代理进行通信.这样可以使用你的NMS应用程序来监视Data Services作业的状态。

Data Services Adapter SDK

Data Services 适配器SDK为快速开发与其他应用程序和中间件产品如EAI系统的适配器提供了一个Java平台。Data Services适配器使用业界标准的XML和Java技术简化学习曲线。适配器提供与包括下列内容进行交互需要的方式：

- ◆ 从Data Services到其他系统的读取，写入和请求应答
- ◆ 从其他系统到Data Services的请求应答

详细信息参见《Data Services Adapter SDK User's Guide》。

第二节 可选的Data Services组件

Data Services多用户

Data Services多用户是一个可选组件，它允许你的开发团队在应用程序开发的所有阶段的相互独立的部分共同协作。尽管每个用户在一个唯一的本地资源库中开发应用，团队使用了一个中央资源库来存储整个项目的原版拷贝。中央资

源库保存了一个应用程序对象的所有版本，这样如果需要的话可以回复到前一个版本。

多用户开发包括其他的高级特性，如加标签和过滤，给你提供了对应用程序对象的更多的灵活性和控制管理。

详细信息参见 《Services Management Console: Administrator Guide》 和 《Data Services Advanced Development Guide》。

第三节 Data Services管理工具

Data Services有几个管理工具帮助你管理Data Services组件。

许可证管理器

许可证管理器显示你目前拥有许可证的Data Services组件。

资源库管理器

资源库管理器允许你创建，升级和检查本地资源库和中央资源库的版本。

服务器管理器

服务器管理器允许你添加，删除，或者编辑作业服务器和访问服务器的属性。它自动安装在安装了作业服务器和访问服务器的每台计算机上。使用服务器管理器来定义作业服务器和资源库之间的连接。可以将不同机器上的多个作业服务器连接到一个单独的资源库(用于负载均衡)，或者每个作业服务器连接到多个资源库(其中一个是缺省的资源库)来支持单个资源库(例如，从生产系统分离测试)。你也可以制定一个作业服务器是SNMP激活的。服务器管理器也是为smtp_to

email 函数设置属性的地方。

第四节 Data Services操作系统平台

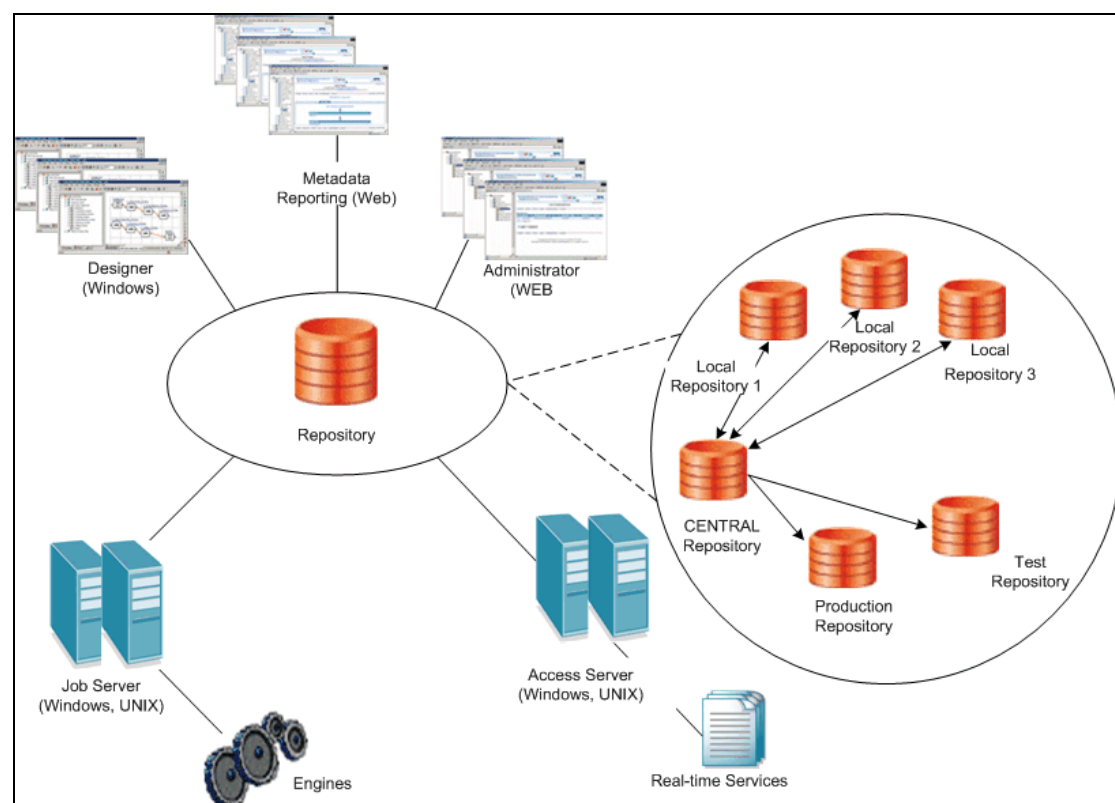
关于支持的环境和硬件需求的详细的列表参见Business Objects支持站点

<https://service.sap.com/bosap-support>.

上的支持平台文档。该文档包括了数据库，应用程序，Web应用程序服务器，Web浏览器和操作系统的特定版本和补丁版本要求。

第五节 Data Services分布式架构

Data Services有一个分布式的架构。访问服务器可以为多个作业服务器和资源库服务。多用户许可证的扩展允许多个Designer在同一个中央资源库中进行工作。下图展示了这两个特性：



可以在多台计算机之间按照下列规则分布Data Services组件：

- ◆ 引擎进程运行在产生这些引擎进程的作业服务器相同的计算机上
- ◆ 适配器需要一个本地作业服务器

在多台计算机上分布Data Services的组件以便更好地支持网络的流量和连接性的要求。可以创建一个最低限度的分布式系统为开发和测试而设计，或者创建非常分散的分布式系统为生产环境扩展需求而设计。

主机名和端口号

Web应用程序，Data Services访问服务器，Data Services作业服务器和实时服务之间的通信通过由IP地址(或者主机名)和端口号指定的TCP/IP连接才能发生。如果你的网络没用使用静态编址，那么使用计算机的名字作为主机名。如果连接到一个使用了静态地址的计算机，那么使用那个编号作为访问服务器和作业服务器配置中的主机名。为了允许更高的可扩展性，每个组件维护了它自己的一个链接列表。通过服务器管理器，Data Services管理员，资源库管理器和(来自Web客户端的)Message Client库调用来定义这些连接。