



Informatica® Multidomain MDM
10.3

基础架构规划指南

本软件和文档仅根据包含使用与披露限制的单独许可协议提供。未事先征得 Informatica LLC 同意，不得以任何形式、通过任何手段（电子、影印、录制或其他手段）复制或传播本文档的任何部分。

美国政府权利交付给美国政府客户的程序、软件、数据库及相关文档和技术数据是指适用的联邦采购条例和政府机构特定补充条例中定义的"商业计算机软件"或"商业技术数据"。因此，使用、复制、披露、修改和改编应遵循适用的政府合同中规定的限制和许可条款、政府合同条款的适用范围以及 FAR 52.227-19 商用计算机软件许可中规定的额外权利。

Informatica、Informatica 标志和 ActiveVOS 是 Informatica LLC 在美国和世界其他许多司法管辖区的商标或注册商标。欲获得 Informatica 商标的最新列表，请访问 <https://www.informatica.com/trademarks.html>。其他公司和产品名称可能是其各自所有者的商业名称或商标。

本软件和/或文档中的若干部分受第三方所拥有的版权约束。所需的第三方声明随产品一起提供。

本文档中的信息如有更改，恕不另行通知。如发现本文档中有什么问题，请通过以下电子邮件地址向我们报告：infa_documentation@informatica.com。

Informatica 产品根据对应协议的条款和条件进行担保。INFORMATICA 按"原样"提供本文档中的信息，无任何明示或暗示的担保，包括但不限于任何适销性和特定用途适用性担保，也没有任何非侵权担保或条件。

发布日期: 2019-05-28

目录

前言	5
Informatica 资源	5
Informatica Network	5
Informatica 知识库	5
Informatica 文档	5
Informatica 产品可用性矩阵	6
Informatica Velocity	6
Informatica Marketplace	6
Informatica 全球客户支持部门	6
第 1 章：基础架构规划简介	7
基础架构规划概述简介	7
安装要求表	7
第 2 章：业务和技术要求	9
业务和技术要求概述	9
确定安装组件	10
确定数据库环境	11
Oracle RAC	11
确定应用程序服务器环境	12
Java 虚拟机的逻辑分组	12
确定时间轴粒度	12
确定外部清理引擎	13
确定操作系统区域设置	14
确定 HTTPS 协议要求	14
确定密码哈希的安全配置	14
确定 Elasticsearch 的搜索配置	15
第 3 章：安装和部署注意事项	16
安装和部署注意事项概述	16
安装和部署目标	16
高可用性	16
可扩展性	17
负载均衡	17
可维护性	18
第 4 章：安装拓扑示例	19
安装拓扑示例	19
独立应用程序服务器实例拓扑	19
多应用程序服务器实例拓扑	21

应用程序服务器群集拓扑.....	24
------------------	----

前言

《Multidomain MDM 基础架构规划指南》可帮助您规划 Informatica^(R) MDM Hub 环境的基础架构和体系结构。本指南提供了一些安装拓扑示例，有助于您理解和决定安装拓扑。

《Multidomain MDM 基础架构规划指南》的目标读者是以下人员：

- 基础架构规划人员和主数据管理解决方案架构师
- 想了解基础架构和 MDM Hub 体系结构决策会如何影响业务的业务经理

本指南假定您了解 IT 基础架构的要求，也熟悉组织的数据管理需求。

Informatica 资源

Informatica Network

Informatica Network 囊括了 Informatica 全球客户支持部门、Informatica 知识库和其他产品资源。要访问 Informatica Network，请访问 <https://network.informatica.com>。

成员可以执行以下操作：

- 在一个位置访问您的所有 Informatica 资源。
- 在知识库中搜索文档、常见问题和最佳实践等产品资源。
- 查看产品可用性信息。
- 查看支持案例。
- 查找当地的 Informatica 用户组网络并与您的伙伴进行协作。

Informatica 知识库

使用 Informatica 知识库可在 Informatica Network 中搜索文档、入门知识文章、最佳实践和 PAM 等产品资源。

要访问知识库，请访问 <https://kb.informatica.com>。如果您对知识库有任何疑问、意见或建议，请与 Informatica 知识库团队联系，电子邮件地址为 KB_Feedback@informatica.com。

Informatica 文档

要获取有关产品的最新文档，请浏览 Informatica 知识库，网址为 https://kb.informatica.com/_layouts/ProductDocumentation/Page/ProductDocumentSearch.aspx。

如果您对此文档有任何疑问、意见或建议，请与 Informatica 文档团队联系，电子邮件地址为 infa_documentation@informatica.com。

Informatica 产品可用性矩阵

产品可用性矩阵 (PAM) 指明了产品版本支持的操作系统版本、数据库以及其他类型的数据源和目标。如果您是 Informatica Network 成员，您可以访问 PAM，网址为 <https://network.informatica.com/community/informatica-network/product-availability-matrices>。

Informatica Velocity

Informatica Velocity 收集了 Informatica 专业服务开发的一系列提示和最佳实践。Informatica Velocity 基于数以百计的数据管理项目的实际经验而开发，汇集了我们曾在世界各地组织就职的顾问在成功规划、开发、部署和维护数据管理解决方案方面的知识。

如果您是 Informatica Network 成员，您可以访问 Informatica Velocity 资源，网址为 <http://velocity.informatica.com>。

如果您对 Informatica Velocity 有任何疑问、意见或建议，请通过 ips@informatica.com 与 Informatica 专业服务联系。

Informatica Marketplace

Informatica Marketplace 是一个论坛，该论坛中提供的解决方案可补充、扩展或增强您的 Informatica 实现。您可以利用 Informatica 开发人员和合作伙伴提供的数以百计解决方案中的任何方案，提高生产率，加快项目的实现时间。您可以访问 Informatica Marketplace，网址为 <https://marketplace.informatica.com>。

Informatica 全球客户支持部门

您可以通过电话或 Informatica Network 上的联机支持与全球支持中心联系。

要查找您当地的 Informatica 全球客户支持部门电话号码，请访问 Informatica 网站，链接为：
<http://www.informatica.com/us/services-and-training/support-services/global-support-centers>。

如果您是 Informatica Network 成员，您可以使用联机支持，网址为 <http://network.informatica.com>。

第 1 章

基础架构规划简介

本章包括以下主题：

- [基础架构规划概述简介, 7](#)
- [安装要求表, 7](#)

基础架构规划概述简介

主数据管理 (MDM) 是一个提高组织中的数据可靠性和数据维护过程的可控流程。Multidomain MDM 也称为 MDM Hub。MDM 会作为更广泛的数据管理程序的一部分进行部署，该程序会将技术、人员、策略和流程结合在一起。请根据组织的数据策略定义、策略和目标规划 MDM Hub 部署的基础架构。

MDM Hub 含有核心和可选组件。决定要在 MDM Hub 环境中使用的组件。此外，您需要决定操作系统、数据库系统、应用程序服务器和负载均衡器等基础架构组件。安装和部署 MDM Hub 前，请拟定清晰的安装和部署目标。此外，您需要决定符合这些目标的安装拓扑。

为确保成功实施 MDM Hub，请在安装要求表中整理出 MDM Hub 实施人员要求提供的信息。

安装要求表

创建一个安装要求表，列出 MDM Hub 实施人员要求的信息，确保成功实施 MDM Hub。安装要求取决于您的业务和技术要求。此外，还要考虑安装和部署目标。

您可以将以下安装要求信息添加到安装要求表：

- 详细安装拓扑
- 要部署的可选 MDM Hub 安装组件
- 时间轴粒度
- 外部清理引擎
- MDM Hub 组件的操作系统区域设置
- HTTPS 协议
- 数据库类型
- 应用程序服务器类型
- 密码哈希的安全配置

- Elasticsearch 的搜索配置
- 详细安装拓扑

第 2 章

业务和技术要求

本章包括以下主题：

- [业务和技术要求概述, 9](#)
- [确定安装组件, 10](#)
- [确定数据库环境, 11](#)
- [确定应用程序服务器环境, 12](#)
- [确定时间轴粒度, 12](#)
- [确定外部清理引擎, 13](#)
- [确定操作系统区域设置, 14](#)
- [确定 HTTPS 协议要求, 14](#)
- [确定密码哈希的安全配置, 14](#)
- [确定 Elasticsearch 的搜索配置, 15](#)

业务和技术要求概述

为 MDM Hub 环境规划基础架构时，请考虑业务和技术要求。您可能需要与组织中对 MDM Hub 环境有兴趣的其他利益相关方协商业务和技术要求。

实施 MDM Hub 前，实施人员需要了解需要部署的 MDM Hub 组件。此外，实施人员还需要了解业务和技术要求，例如时间轴粒度、操作系统区域设置以及安全通信需求。

确定安装组件

MDM Hub 增强了数据可靠性和数据维护过程。可以通过 MDM Hub 访问 Hub 控制台 功能。MDM Hub 含有核心和可选安装组件。请根据业务需求，决定要安装的可选组件。

核心组件

下表介绍了核心安装组件：

组件	说明
MDM Hub 主数据库	存储和合并 MDM Hub 业务数据的架构。包含 MDM Hub 环境配置设置，例如，用户帐户、安全配置、操作引用存储注册表以及消息队列设置。可以从 MDM Hub 主数据库访问和管理操作引用存储。MDM Hub 主数据库 的默认名称为 CMX_SYSTEM，但您可以使用自定义名称。
操作引用存储	存储和合并 MDM Hub 业务数据的架构。包含主数据、内容元数据以及处理和管理主数据所采用的规则。可以为不同地理区域、不同组织部门以及开发和生产环境配置单独的操作引用存储数据库。可以在多个服务器计算机中分发操作引用存储数据库。操作引用存储的默认名称为 CMX_ORS。
Hub 服务器	在应用程序服务器上部署的 J2EE 应用程序。Hub 服务器会处理 MDM Hub 中存储的数据并将 MDM Hub 与外部应用程序集成。Hub 服务器为 MDM Hub 管理核心和通用服务。
进程服务器	在应用程序服务器上部署的 J2EE 应用程序。进程服务器处理批处理作业（如负载、重新计算 BVT 和重新验证），并执行数据清理和匹配操作。进程服务器与配置用于标准化和优化数据的清理引擎相连接，以进行匹配和合并。
置备工具	构建业务实体模型以及为 Data Director 配置 Entity 360 框架的工具。构建业务实体模型后，可以将配置发布到 MDM Hub。
Informatica ActiveVOS ^(R)	MDM Hub 内部需要用来处理数据的业务流程管理 (BPM) 工具。Informatica ActiveVOS 支持自动化业务流程，包括数据更改-审批流程。也可以使用 Informatica ActiveVOS 确保对主数据的更改要经历审核和审批流程，然后才能加入最佳数据版本 (BVT) 记录。 在 Hub 服务器 安装过程中安装 ActiveVOS Server 时，将安装 ActiveVOS Server、ActiveVOS Console 和 Process Central。此外，请安装预定义 MDM workflow、任务和角色。
Data Director (IDD)	用来控制和管理存储在 MDM Hub 中的数据的用户界面。在 IDD 中，数据按业务实体组织，例如客户、供应商和员工。业务实体指的是对组织具有重大意义的数据组。

可选组件

下表介绍了可选安装组件：

组件	说明
资源工具包	将 MDM Hub 集成到应用程序和工作流中的示例、应用程序以及实用程序集。可以选择要安装的资源工具包组件。
Informatica platform	包含用来清理源数据和将源数据传输到 MDM Hub 的 Informatica 服务和 Informatica 客户端的环境。可以使用 Informatica platform，而非 MDM Hub 中提供的清理功能来清理数据。 在安装 Hub 服务器 的过程中安装 Informatica 平台时，将安装数据集成服务、模型存储库服务以及 Informatica Developer（即 Developer tool）。

组件	说明
Dynamic Data Masking	在 MDM Hub 和数据库之间运行的数据安全工具，可防止对敏感信息的非授权访问。Dynamic Data Masking 可拦截发送至数据库的请求，并对请求应用数据屏蔽规则，然后再将数据发送回 MDM Hub。
Informatica 数据控件 (IDC)	适用于仅基于主题区域数据模型的 Informatica Data Director (IDD)。IDC 是一组用户界面控件，用于在业务用户所使用的第三方应用程序中展示 MDM Hub 数据。
零停机 (ZDT) 模块	<p>确保应用程序在 MDM Hub 升级期间仍能访问 MDM Hub 中数据的模块。在 ZDT 环境中，请复制源数据库和目标数据库。在 MDM Hub 升级期间，ZDT 模块会将源数据库中发生的数据更改复制到目标数据库。</p> <p>要购买 ZDT 模块，请联系 Informatica 销售代表。有关安装零停机环境的信息，请参阅数据库的《<i>Multidomain MDM Zero Downtime 安装指南</i>》。</p>

确定数据库环境

您可以在以下数据库环境中存储 MDM Hub 数据：

- Oracle 数据库
- IBM DB2
- Microsoft SQL Server

请根据业务需求，决定要设置的数据库环境。有关支持的数据库环境的详细信息，请参阅“产品可用性表格 (PAM)”。可以在 <https://network.informatica.com/community/informatica-network/product-availability-matrices> 访问 PAM。

Oracle RAC

Oracle RAC 环境可提高性能、改进容错能力并提供可扩展性。确定您的业务是否需要在 MDM Hub 中使用 Oracle RAC 的好处。

有关 Oracle RAC 的信息，请参阅 Oracle 文档。

使用 Oracle RAC 的注意事项

根据以下注意事项，决定 Oracle RAC 环境设置：

- 安装 Oracle RAC 时使用 Oracle 服务名称代替 Oracle SID。提供指定连接和动态重新分配数据库服务器的灵活性。
- 在 tnsnames.ora 文件中配置所有 Oracle RAC 节点。
- 使用 Oracle RAC 负载平衡连接。在群集中的所有可用节点上分布工作负载。如果节点出现故障，MDM Hub 批处理作业将失败，但该作业可以在群集中的可用节点上启动。

MDM Hub 支持为以下操作和组件实现负载平衡：

- Hub 控制台操作
- 除“生成匹配标志”作业以外的所有批处理作业

注意：MDM Hub 支持为“匹配”作业的模糊匹配部分以及为“暂存”作业的清理工进部分实现负载平衡。

当从 Hub 控制台调用批处理作业时，批处理作业无法使用负载平衡。

- 服务集成框架 (SIF)
- 传出 JMS 消息队列
- Hub 服务器
- 进程服务器
- 置备工具
- 存储库管理器

注意: 如果需要 DDL，则由于 DDL 使用直接 JDBC 连接，存储库管理器 将无法使用负载平衡。

MDM Hub 不支持为 Data Director 实现负载平衡。

确定应用程序服务器环境

您可以在以下应用程序服务器环境中部署 MDM Hub：

- JBoss
- WebLogic
- WebSphere

请根据业务需求，决定要设置的应用程序服务器环境。有关支持的应用程序服务器环境的详细信息，请参阅“产品可用性表格 (PAM)”。可以在 <https://network.informatica.com/community/informatica-network/product-availability-matrices> 访问 PAM。

Java 虚拟机的逻辑分组

确定业务是否需要创建逻辑 Java 虚拟机 (JVM) 组。部署逻辑 JVM 组中的 Hub 服务器和进程服务器时，应用程序之间的所有通信都将留在组内。要对 JVM 进行分组，请在 MDM Hub 环境中为每个 JVM 分配一个组 ID。

您可以在以下场景中使用逻辑 JVM 组：

- 需要多个 JVM，并要将一些 JVM 用作主节点，一些用作辅助节点。可以使用辅助节点完成特定操作，或在主节点超载时使用辅助节点。例如，可以配置一个逻辑 JVM 组完成主进程，配置另一个组完成辅助或备份进程。
- 要并行运行多个批处理组作业，并要控制线程等可用资源。例如，可以为在“参与方”表和“地址”表上运行的批处理组作业创建逻辑 JVM 组。
- 要将类似的进程分组。例如，可以为 SIF API 调用创建一个逻辑 JVM 组，为批处理作业创建另一个组。

确定时间轴粒度

时间轴粒度是用于定义记录版本有效期的时间度量单位。例如，可以选择将有效期设置为以年、月或秒为单位。决定时间轴粒度并向 MDM Hub 实施人员提供此信息。

可以将时间轴粒度配置为年、月、日、时、分或秒，以指定 MDM Hub 实现中的数据有效期。您可以在创建或更新操作引用存储时配置所需的时间轴粒度。

重要说明: 您配置的时间轴粒度无法更改。

如果指定任意时间轴粒度的有效期，系统会对有效期使用数据库时间区域设置。要创建一个有效期为一个时间轴度量单位的版本，则开始日期和结束日期必须相同。

下表介绍了可配置的时间轴粒度选项：

时间轴粒度	说明
年	如果时间轴粒度为年，可以采用年格式 yyyy（例如 2010）指定有效期。记录的有效开始日期从年初开始，有效结束日期在年末结束。例如，如果有效开始日期为 2013 年，有效结束日期为 2014 年，则记录的有效期将从 01/01/2013 到 31/12/2014。
月	如果时间轴粒度为月，可以采用月格式 mm/yyyy（例如 01/2013）指定有效期。记录的有效开始日期从一个月的第一天开始。记录的有效结束日期在一个月最后一天结束。例如，如果有效开始日期为 02/2013，有效结束日期为 04/2013，则记录的有效期将从 01/02/2013 到 30/04/2013。
日	如果时间轴粒度为日，可以采用日格式 dd/mm/yyyy（例如 13/01/2013）指定有效期。记录的有效开始日期从午夜 12:00 点开始。记录的有效结束日期在 23:59 结束。例如，如果有效开始日期为 13/01/2013，有效结束日期为 15/04/2013，则记录的有效期将从 13/01/2013 的午夜 12:00 到 15/04/2013 的 23:59。
时	如果时间轴粒度为时，则有效期包括年、月、日和时。时间轴格式为 dd/mm/yyyy hh，例如 13/01/2013 15。记录的有效开始日期从一天的某一小时开始。记录的有效结束日期在您指定的小时结束。例如，如果有效开始日期为 13/01/2013 15，有效结束日期为 15/04/2013 10，则记录的有效期将从 13/01/2013 的 15:00 到 15/04/2013 的 10:59。
分	如果时间轴粒度为分，则有效期包括年、月、日、时和分。时间轴格式为 dd/mm/yyyy hh:mm，例如 13/01/2013 15:30。记录的有效开始日期从某一分钟开始。记录的有效结束日期在您指定的分钟结束。例如，如果有效开始日期为 13/01/2013 15:30，有效结束日期为 15/04/2013 10:45，则记录的有效期将从 13/01/2013 的 15:30:00 到 15/04/2013 的 10:45:59。
秒	如果时间轴粒度为秒，有效期将包括年、月、日、时、分和秒。时间轴格式为 dd/mm/yyyy hh:mm:ss，例如 13/01/2013 15:30:45。记录的有效开始日期从某一秒钟开始。记录的有效结束日期在您指定的秒钟结束。例如，如果有效开始日期为 13/01/2013 15:30:55，有效结束日期为 15/04/2013 10:45:15，则记录的有效期将从 13/01/2013 的 15:30:55:00 到 15/04/2013 的 10:45:15:00。

确定外部清理引擎

如果有意将清理引擎与 MDM Hub 集成，请确定清理引擎。您可以将地址验证等清理引擎与 MDM Hub 集成起来。

下表列出了 MDM Hub 支持的清理引擎以及可与这些清理引擎配合使用的 Informatica MDM 适配器：

清理引擎	Informatica MDM Hub 适配器
IDQ	Informatica IDQ 适配器
Informatica 地址验证	Informatica 地址验证适配器

清理引擎	Informatica MDM Hub 适配器
FirstLogic Direct	FirstLogic Data Quality 适配器
Trillium	Trillium Director 适配器
SAP Data Services XI	SAP Data Services XI 适配器

有关可以与 MDM Hub 集成的清理引擎的详细信息，请参阅《*Multidomain MDM 清理适配器指南*》。

确定操作系统区域设置

操作系统区域设置用于定义用户语言及区域。请根据业务需求，为 Hub 服务器、Hub 存储和 Hub 控制台设置相同的操作系统区域设置。

请为 MDM Hub 组件选择以下区域设置之一：

- en_US
- fr_FR
- de_DE
- ja_JP
- ko_KR
- zh_CN
- ES
- pt_BR

确定 HTTPS 协议要求

可以为 MDM Hub 通信配置 HTTPS 协议。此外，您可能也希望为 ActiveVOS 和 MDM Hub 之间的通信使用 HTTPS 协议。

是否为 MDM Hub 使用安全通信取决于业务需求。您需要向 MDM Hub 实施人员说明是否需要为 MDM Hub 使用安全通信。

确定密码哈希的安全配置

密码哈希是一种通过加密哈希函数对密码进行加密的方法。MDM Hub 使用密码哈希方法保护用户密码的安全，并确保密码绝不会以明文形式存储在数据库中。MDM Hub 管理员在安装 Hub 服务器过程中配置密码哈希选项，例如使用的算法和证书。

确定实施人员在安装 MDM Hub 期间需要指定的密码哈希安全配置选项。

实施人员需要为密码哈希指定以下安全配置选项：

- 是否要创建自己的客户哈希键作为哈希算法的一部分。
- 是使用默认的 SHA3 哈希算法还是创建自定义哈希算法。
- 是使用默认的证书提供程序还是使用自定义证书提供程序。

确定 Elasticsearch 的搜索配置

要为业务用户和数据管理者提供快速、现代、全文本的搜索体验，必须配置 Elasticsearch 搜索。Solr 搜索已被弃用，由 Elasticsearch 搜索替代。

Elasticsearch 是一个提供分布式索引和搜索的开源全文本搜索引擎。Elasticsearch 可以安装在安装了 MDM Hub 组件的任何计算机上，也可以安装在单独的计算机上。基于 MDM Hub 拓扑和要建立索引的数据量，确定要为 Elasticsearch 群集配置的节点数。每个节点又可以具有多个索引。您可以决定将索引拆分为多个分片。

对 MDM Hub 环境而言，Elasticsearch 搜索的性能优于 Solr 搜索。此外，在 MDM Hub 环境中，使用 Elasticsearch 建立数据索引的性能要好得多。

为 MDM Hub 环境配置 Elasticsearch 时，用户可以搜索特定业务实体中的记录。这与使用 Solr 搜索不同，在使用 Solr 搜索时，您可以在所有业务实体中或特定业务实体中进行搜索。

当配置为使用 Elasticsearch 搜索时，您可以使用星号通配符 (*) 执行搜索。与 Solr 不同的是，Elasticsearch 查询解析器可以在搜索字符串中灵活使用各种字符。您可以使用 AND 和 OR 等运算符搜索记录。

有关如何确定要为 Elasticsearch 群集配置的节点数量以及如何配置索引的信息，请参阅 *Elasticsearch 文档*。

第 3 章

安装和部署注意事项

本章包括以下主题：

- [安装和部署注意事项概述, 16](#)
- [安装和部署目标, 16](#)

安装和部署注意事项概述

安装和部署 MDM Hub 前，请首先考虑安装和部署目标。接着，您可以决定将 MDM Hub 安装并部署在独立应用程序服务器实例还是应用程序服务器群集上。

Informatica 建议将 MDM Hub 安装并部署在独立应用程序服务器实例上。要实现高可用性，您可为实时 API 调用使用外部负载均衡器。要扩展 MDM Hub 实施，您可添加独立应用程序服务器实例并部署附加 MDM Hub 组件。您可通过 MDM Hub 的内部机制为批处理作业实现负载平衡。

如果在应用程序服务器群集中安装和部署 MDM Hub，会话不会复制或故障转移到群集中的其他节点。您可以使用 SIF API 或 Hub 控制台运行 MDM Hub 批处理作业，但批处理作业不会故障转移，因为会话不会复制。但是，如果批处理作业请求通过 Hub 控制台发出，该请求（而不是批处理作业）会故障转移到群集中的活动节点。此外，如果某个应用程序服务器节点发生故障，Informatica Data Director (IDD) 会话会丢失，也不会复制。

安装和部署目标

安装和部署 MDM Hub 前，请考虑以下安装和部署目标：

- 高可用性
- 可扩展性
- 负载平衡
- 可维护性

高可用性

高可用性描述的是系统在一台或多台服务器发生故障后仍能正常工作的能力。您可以利用多台独立的应用程序服务器实例或使用应用程序服务器群集来实现 MDM Hub 高可用性。

应用程序服务器无法提供 MDM Hub 高可用性，因为 MDM Hub 使用无状态会话 Bean。无状态会话 Bean 无法保持与客户端的会话状态，因此应用程序服务器无法在各个应用程序服务器群集节点之间同步 Bean 状态。

为实现高可用性，MDM Hub 使用内部元数据缓存机制。该元数据缓存机制会同步 MDM Hub 环境中的元数据，并在整个 MDM Hub 实施中提供这些元数据。如果一台计算机上的应用程序发生故障，联机的应用程序可以使用缓存中的元数据。该元数据缓存机制使用 Infinispan，这是一种复制的缓存，可处理任意应用程序服务器环境中的元数据缓存要求。

假设您受到以下因素影响，倾向于决定使用高可用环境：

- MDM Hub 实施含有多个 Hub 服务器实例，但发生故障时，Hub 控制台操作无法故障转移到活动节点。为确保 Hub 控制台操作能够故障转移到活动节点，Hub 服务器实例必须是应用程序服务器群集的一部分。
- 如果批处理作业请求通过 Hub 控制台发出，该请求会故障转移到群集中的活动节点。如果批处理作业请求通过服务集成框架 API 发出，该请求不会故障转移到群集中的活动节点。批处理作业不会故障转移，因为批处理作业不会复制。
- 如果 MDM Hub 实施含有多个 Hub 服务器实例，而且您使用 JMS 消息，则可以将 Hub 服务器实例部署在群集中。如果未将 Hub 服务器实例部署在群集中，传出 JMS 消息将无法供所有使用者使用。此外，您可以考虑使用合适的 JMS 服务器部署策略来管理此场景。
- 如果使用 Informatica Data Director (IDD)，IDD 会话会绑定到为会话提供服务的应用程序服务器节点。如果应用程序服务器节点发生故障，IDD 会话将终止。IDD 会话不会复制。IDD 用户需要重新登录。

可扩展性

可扩展性描述的是系统应对未来对资源和处理能力更高需求的能力。您可以利用独立应用程序服务器或者使用应用程序服务器群集来实现 MDM Hub 可扩展性。

MDM Hub 实施具有以下功能，使其非常便于扩展：

MDM Hub 缓存实施

MDM Hub 缓存实施使用独立于应用程序服务器的分发机制。

多线程进程服务器实例

进程服务器实例为多线程，可以并发处理多个请求。MDM Hub 支持对 Hub 控制台操作、批处理作业和服务集成框架 (SIF) 请求进行多线程处理。

多个进程服务器实例

您可以为 MDM Hub 中的每个操作引用存储运行多个进程服务器。

MDM Hub 的扩展性功能不需要额外的外部组件。随着数据量增加，如果需要扩展 MDM Hub 实施，您可以添加更多进程服务器实例。要跨多个 CPU 分发处理负载，同时并行运行批处理作业，请在多个主机上部署进程服务器。

负载均衡

负载均衡指的是系统将工作负载分发到可用资源的能力。将进程服务器部署到独立应用程序服务器实例时，您可以实现负载均衡。

MDM Hub 实施的进程服务器实例使用内部负载均衡机制。在应用程序服务器群集中，您无需使用负载均衡功能。您可以将 MDM Hub 安装并部署到独立应用程序服务器，并使用进程服务器的负载均衡功能。将 MDM Hub 安装并部署到独立应用程序服务器时，请使用进程服务器的负载均衡功能。

注意：要使用传出 JMS 队列或对 Hub 控制台操作进行负载均衡，请在应用程序服务器群集中部署 Hub 服务器实例。请勿将进程服务器实例部署到应用程序服务器群集。

可维护性

可维护性指的是灵活更改或升级 MDM Hub 实施的能力。您可以维护独立应用程序服务器或应用程序服务器群集中的 MDM Hub 实施。

应用程序服务器群集中可以实现协调一致的多服务器管理，这是独立应用程序服务器实例所无法实现的。在应用程序服务器群集中，管理和维护对 MDM Hub 配置和部署的更改要比在多个独立应用程序服务器实例上简单。

请考虑开展维护任务的频率，并据此决定是否使用可维护性高的环境。在独立应用程序服务器实例上安装或升级 MDM Hub 期间，您需要安装或升级并部署每个应用程序服务器实例。此外，每个计算机需要更新多个配置。在应用程序服务器群集上，安装或升级以及部署相对来说不那么繁琐。

注意：仅当可维护性预计能带来很大的好处时，才应该在应用程序服务器群集上部署进程服务器实例。

第 4 章

安装拓扑示例

本章包括以下主题：

- [安装拓扑示例, 19](#)
- [独立应用程序服务器实例拓扑, 19](#)
- [多应用程序服务器实例拓扑, 21](#)
- [应用程序服务器群集拓扑, 24](#)

安装拓扑示例

决定安装拓扑时，您需要在高可用性、可扩展性和负载平衡要求等系统特性之间取得平衡。为确保使用理想的安装拓扑，您需要了解您的特定使用场景。安装拓扑示例介绍了有关规划安装拓扑的方案。

安装拓扑示例说明了在一个 MDM Hub 实施中设置 MDM Hub 组件时可以采用的多种方式。您可以根据需要定制这些示例。

规划安装拓扑时，请考虑将以下安装拓扑之一作为起点：

- 独立应用程序服务器实例拓扑
- 多应用程序服务器实例拓扑
- 应用程序服务器群集拓扑

注意：MDM Hub 实施的所有组件版本必须相同。如果有多个 MDM Hub 版本，请将每个版本都安装在一个单独的环境中。

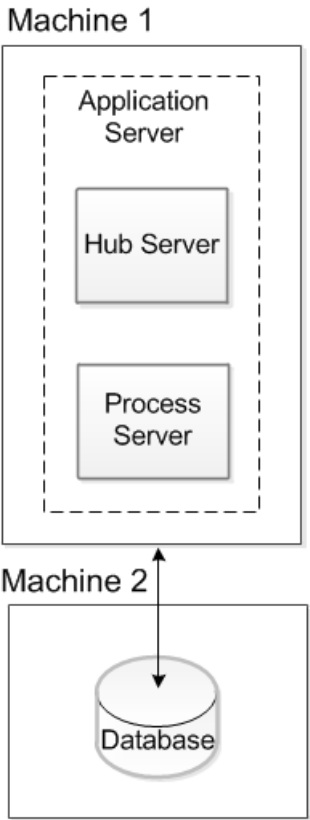
独立应用程序服务器实例拓扑

在独立应用程序服务器实例拓扑中，您需要将所有 MDM Hub 组件安装在独立应用程序服务器实例中。独立应用程序服务器实例拓扑是最基本的拓扑。在独立应用程序服务器实例上部署可简化 MDM Hub 组件之间的通信。

独立应用程序服务器实例拓扑无需在计划内停机时间或计划外停机时间置备。但它不具可扩展性，或可扩展性十分有限，仅限于扩展 Hub 服务器和进程服务器部署到的计算机的处理能力。该拓扑维护起来十分方便。请将该拓扑用于小数据量场景。

该安装拓扑示例含有两台计算机。应用程序服务器实例安装在一台计算机上，数据库服务器安装在另一台计算机上。Hub 服务器和进程服务器部署在含有应用程序服务器实例的计算机上。Hub 存储在安装了数据库服务器的计算机上配置。

下图显示了独立应用程序服务器实例安装拓扑示例：



下表介绍了独立应用程序服务器实例拓扑的功能：

功能	可用性
高可用性	无。
可扩展性	是。 要扩展 MDM Hub 以支持大量数据，请为进程服务器配置多线程。您可以通过添加更多处理能力，纵向扩展 MDM Hub 环境。 MDM Hub 支持为以下操作和组件使用多线程处理： <ul style="list-style-type: none">- Hub 控制台操作- 批处理作业- 服务集成框架 (SIF)
负载均衡	无。
可维护性	方便维护，因为所有 MDM Hub 组件都部署在含有应用程序服务器实例的一台计算机上。

多应用程序服务器实例拓扑

在多应用程序服务器实例拓扑中，可以将 MDM Hub 组件安装分布在多个应用程序服务器实例上。

要设置多应用程序服务器实例拓扑，您需要多台具有应用程序服务器实例的计算机。该拓扑具有可扩展性优势。为了扩展 MDM Hub 实施的处理能力，可以在额外的应用程序服务器实例上部署更多进程服务器实例。

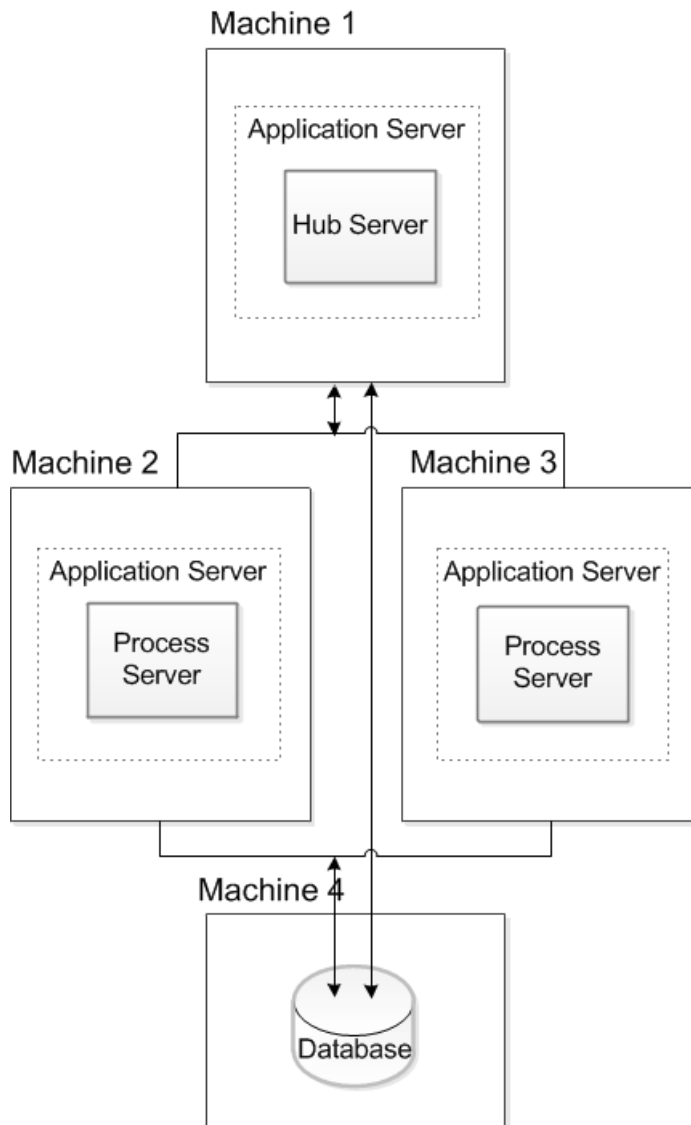
如果进程服务器失败，在进程服务器上运行的批处理作业也会失败。批处理作业不会故障转移至联机的进程服务器实例继续完成。您必须重新启动批处理作业。MDM Hub 的内部负载均衡机制会在联机的进程服务器实例之间分发批处理作业请求。

数据量较大时，您可以使用多应用程序服务器实例拓扑。该拓扑跨您配置的多个进程服务器分发负载，可以处理数据量较大的批处理作业。

注意: 如果 MDM Hub 实施含有多个 Hub 服务器实例而且使用了 JMS 消息队列，为了使用传出 JMS 消息，可以将 Hub 服务器实例部署在群集中。否则，每个应用程序服务器实例都有不同的传出 JMS 目标。

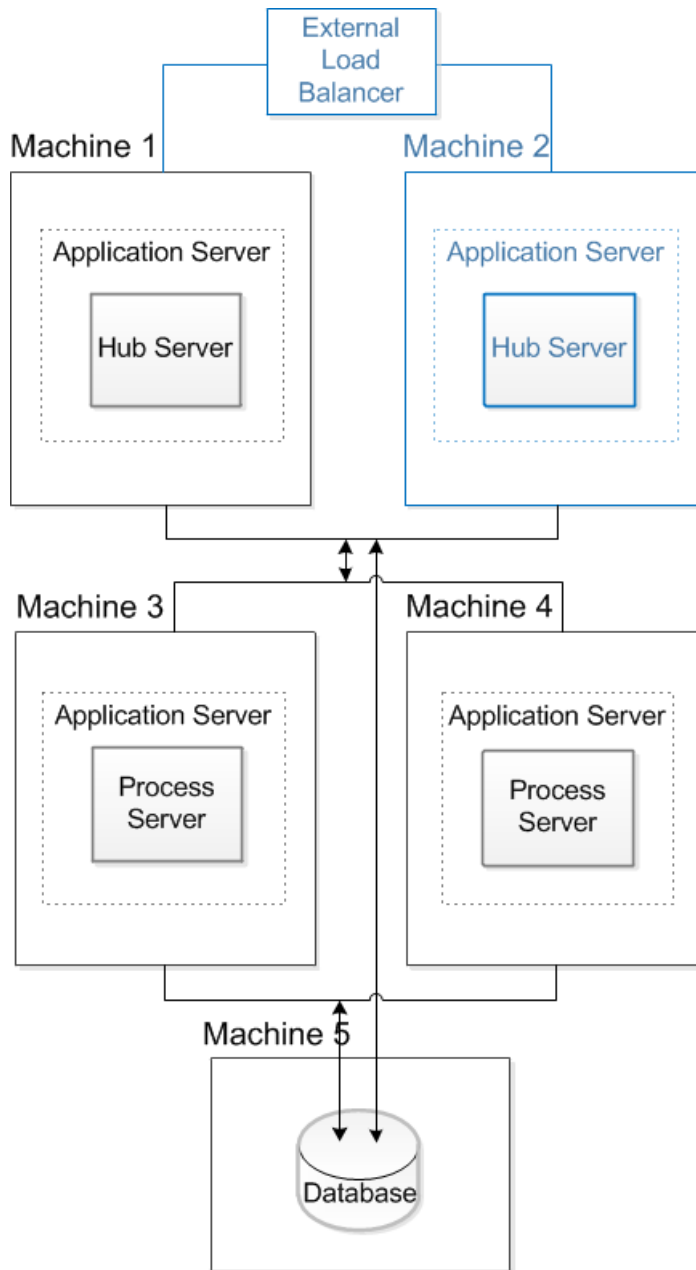
该安装拓扑示例含有四台计算机。四台计算机中有三台安装了应用程序服务器实例。Hub 服务器部署在其中一台计算机的应用程序服务器实例上。进程服务器实例部署在另外两台计算机的应用程序服务器实例上。Hub 服务器会在两个进程服务器实例之间分发批处理作业的处理负载。如果一个进程服务器实例发生故障或脱机，Hub 服务器会将处理请求发送至联机的另一个进程服务器。Hub 存储在安装了数据库服务器的第四台计算机上配置。

下图展示的是一个不具备高可用性的多应用程序服务器实例安装拓扑示例：



如果需要高可用性，则可配置额外的 Hub 服务器实例，并在 Hub 服务器实例之间配置外部负载均衡器。

下图展示的是一个具备高可用性的多应用程序服务器实例安装拓扑示例：



下表介绍了多应用程序服务器实例拓扑的功能：

功能	可用性
高可用性	无。 注意: 如果需要高可用性，则可配置额外的 Hub 服务器实例，并在 Hub 服务器实例之间配置外部负载均衡器。
可扩展性	是。 要扩展 MDM Hub 以支持大量数据，请添加更多 MDM Hub 组件。此外，要并发处理多个请求，请为进程服务器配置多个线程。 MDM Hub 支持为以下操作和组件使用多线程处理： <ul style="list-style-type: none">- Hub 控制台操作- 批处理作业- 服务集成框架 (SIF)
负载均衡	是。MDM Hub 使用内部负载均衡机制，在可用进程服务器实例之间分发负载。 MDM Hub 支持为以下操作实现负载均衡： <ul style="list-style-type: none">- Hub 控制台操作- 除“生成匹配标志”作业以外的所有批处理作业 注意: MDM Hub 支持为“匹配”作业的模糊匹配部分以及为“暂存”作业的清理进程部分实现负载均衡。 MDM Hub 不支持为以下组件实现负载均衡： <ul style="list-style-type: none">- Informatica Data Director- 服务集成框架 (SIF)- 传出 JMS 消息队列
可维护性	难以维护，因为 MDM Hub 组件部署在多台计算机上。在需要频繁更改的环境中，您需要在每台计算机上执行部署和配置。

应用程序服务器群集拓扑

在应用程序服务器群集拓扑中，MDM Hub 组件都安装在应用程序服务器群集中。由于可使用多种组合方案，应用程序服务器群集拓扑计划可能会比较复杂。应用程序服务器群集拓扑的主要优势是易于部署。

要设置应用程序服务器群集拓扑，您需要具备多台含有应用程序服务器实例的计算机，并将这些计算机构成一个应用程序服务器群集。在不同的应用程序服务器群集上部署 Hub 服务器和进程服务器实例。应用程序服务器群集拓扑可应对计划内停机或计划外停机。向群集添加更多节点并部署额外的 MDM Hub 组件即可实现可扩展性。

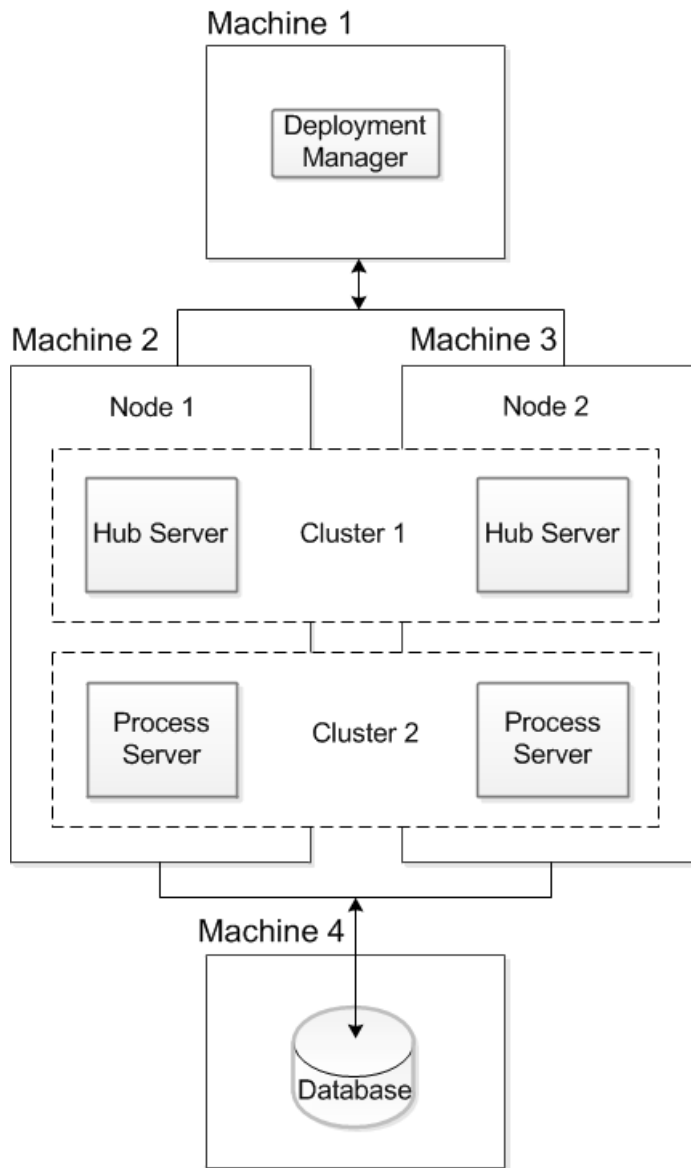
WebSphere 群集拓扑

安装拓扑示例含有构成了两个 WebSphere 群集的四台计算机。WebSphere 部署管理器可安装在任意计算机上，但在本例中，它安装在一台单独的计算机上，用来提供安全的 WebSphere 管理。每个 WebSphere 群集都包括两个节点。Hub 服务器实例部署在一个群集中的每个节点上，因此一个节点发生故障时，群集中的另一个节点会接管。进程服务器实例部署在另一个群集中的每个节点上，因此一个节点发生故障时，群集中的另一个节点会接管。

Hub 服务器负责在进程服务器实例之间分发处理负载。如果进程服务器实例发生故障或脱机，Hub 服务器会将处理请求发送至联机的进程服务器实例。Hub 存储在安装了数据库服务器的第四台计算机上配置。

注意: 您无需将进程服务器实例部署在群集中。使用 JMS 消息队列时，为了使用传出 JMS 消息，可将 Hub 服务器实例部署在群集中。否则，每个应用程序服务器实例都有不同的传出 JMS 目标。

下图展示的是一个 WebSphere 群集安装拓扑示例：



下表介绍了应用程序服务器群集拓扑的功能。

功能	可用性
高可用性	<p>是。</p> <p>MDM Hub 支持为以下操作和组件提供高可用性：</p> <ul style="list-style-type: none"> - Hub 控制台操作 - 服务集成框架 (SIF) - 传出 JMS 消息 <p>MDM Hub 不支持为以下操作和组件提供高可用性：</p> <ul style="list-style-type: none"> - 批处理作业 <p>注意: 如果群集中的某个节点故障，通过 Hub 控制台发出的批处理作业请求会故障转移到活动节点，但批处理作业本身不会故障转移。</p> <ul style="list-style-type: none"> - Informatica Data Director
可扩展性	<p>是。</p> <p>要扩展 MDM Hub 以支持大量数据，请添加更多 MDM Hub 组件。此外，要并发处理多个请求，请为进程服务器配置多个线程。</p> <p>MDM Hub 支持为以下操作和组件使用多线程处理：</p> <ul style="list-style-type: none"> - Hub 控制台操作 - 批处理作业 - 服务集成框架 (SIF)
负载平衡	<p>是。要实现负载平衡，您无需将进程服务器实例部署到应用程序服务器群集。进程服务器实例使用内部负载平衡机制。</p> <p>MDM Hub 支持为以下操作和组件实现负载平衡：</p> <ul style="list-style-type: none"> - Hub 控制台操作 - 除“生成匹配标志”作业以外的所有批处理作业 <p>注意: MDM Hub 支持为“匹配”作业的模糊匹配部分以及为“暂存”作业的清理进程部分实现负载平衡。</p> <ul style="list-style-type: none"> - 服务集成框架 (SIF) - 传出 JMS 消息 <p>注意: Informatica Data Director 在应用程序服务器群集中不支持负载平衡。群集环境中的负载平衡可能会带来意外结果。要增强 MDM Hub 环境的性能，您可使用外部负载平衡器。</p>
可维护性	<p>比独立应用程序服务器实例拓扑更复杂，但相比分布式应用程序服务器拓扑而言更易于部署和维护。使用 WebSphere 部署管理器时，您可以在群集中的节点之间轻松部署 MDM Hub 组件。</p>

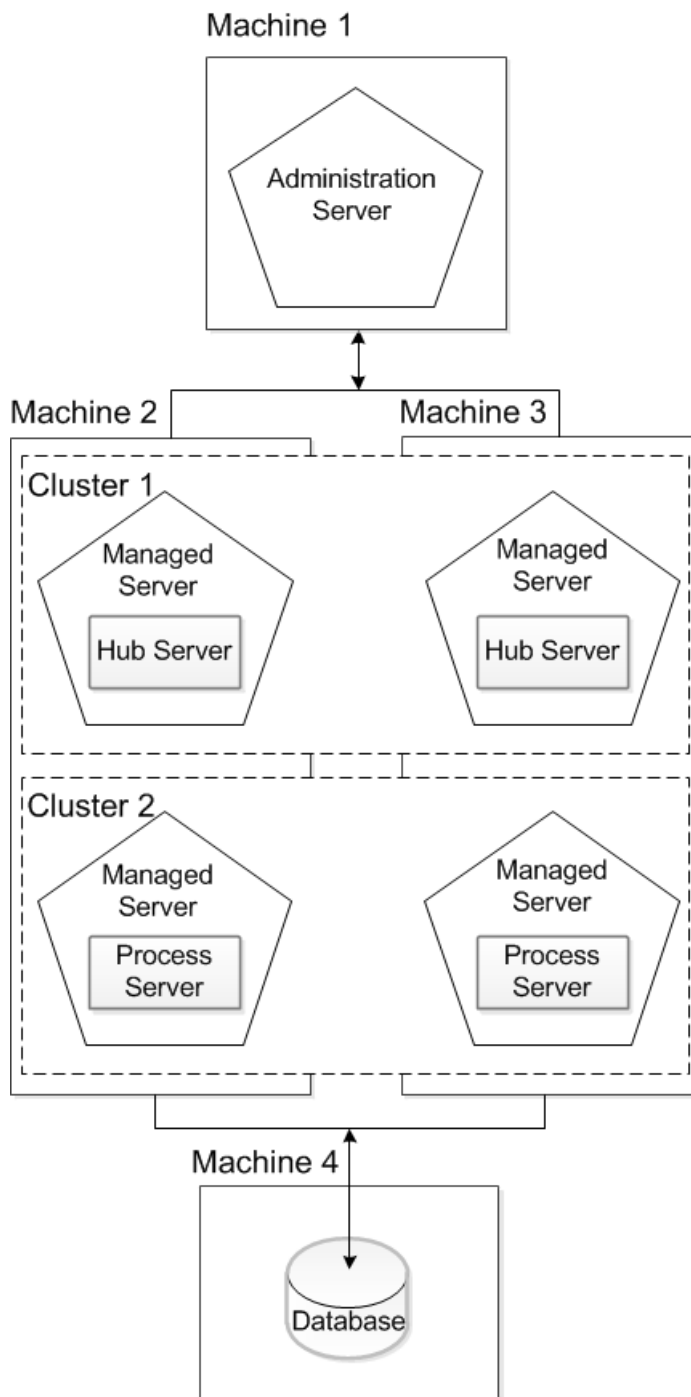
WebLogic 群集拓扑

安装拓扑示例含有构成了两个 WebLogic 群集的四台计算机。WebLogic 管理服务器可安装在任意计算机上，但在本例中，它安装在一台单独的计算机上，用来提供安全的 WebLogic 管理。每个 WebLogic 群集都包括两个受管服务器。Hub 服务器实例部署在一个群集中的每个受管服务器上，因此一个受管服务器发生故障时，群集中的另一个受管服务器会接管。进程服务器实例部署在另一个群集中的每个受管服务器上，因此一个受管服务器发生故障时，群集中的另一个受管服务器会接管。

Hub 服务器负责在两个进程服务器实例之间分发处理负载。如果进程服务器实例发生故障或脱机，Hub 服务器会将处理请求发送至联机的进程服务器实例。Hub 存储在安装了数据库服务器的第四台计算机上配置。

注意: 您无需将进程服务器实例部署在群集中。使用 JMS 消息队列时，为了使用传出 JMS 消息，可将 Hub 服务器实例部署在群集中。否则，每个应用程序服务器实例将具有不同的传出 JMS 目标。

下图显示了 WebLogic 群集安装拓扑示例：



下表介绍了应用程序服务器群集拓扑的功能。

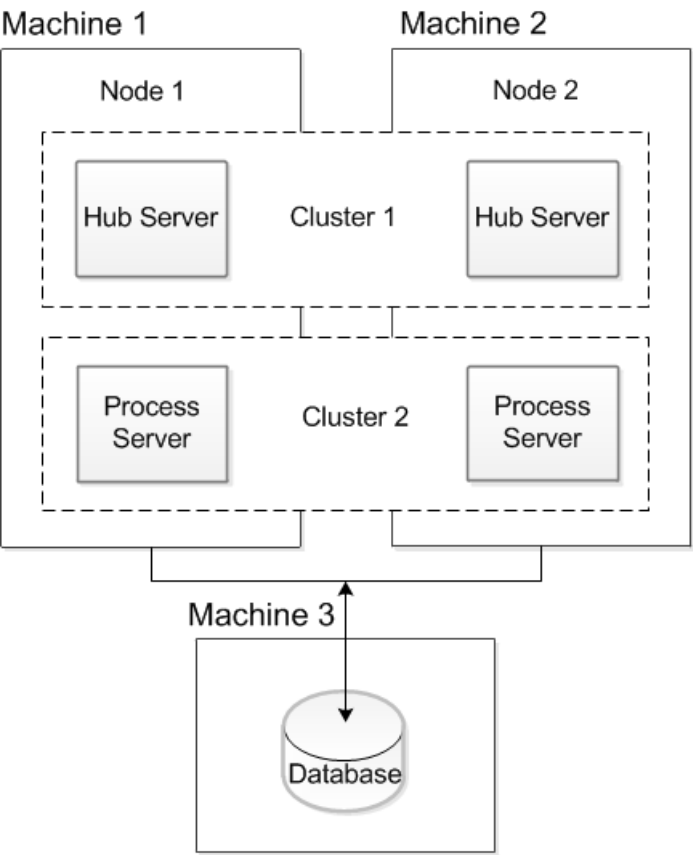
功能	可用性
高可用性	<p>是。</p> <p>MDM Hub 支持为以下操作和组件提供高可用性：</p> <ul style="list-style-type: none"> - Hub 控制台操作 - 服务集成框架 (SIF) - 传出 JMS 消息 <p>MDM Hub 不支持为以下操作和组件提供高可用性：</p> <ul style="list-style-type: none"> - 批处理作业 <p>注意: 如果群集中的某个节点故障，通过 Hub 控制台发出的批处理作业请求会故障转移到活动节点，但批处理作业本身不会故障转移。</p> <ul style="list-style-type: none"> - Informatica Data Director
可扩展性	<p>是。</p> <p>要扩展 MDM Hub 以支持大量数据，请添加更多 MDM Hub 组件。此外，要并发处理多个请求，请为进程服务器配置多个线程。</p> <p>MDM Hub 支持为以下操作和组件使用多线程处理：</p> <ul style="list-style-type: none"> - Hub 控制台操作 - 批处理作业 - 服务集成框架 (SIF)
负载均衡	<p>是。要实现负载均衡，您无需将进程服务器实例部署到应用程序服务器群集。进程服务器实例使用内部负载均衡机制。</p> <p>MDM Hub 支持为以下操作和组件实现负载均衡：</p> <ul style="list-style-type: none"> - Hub 控制台操作 - 除“生成匹配标志”作业以外的所有批处理作业 <p>注意: MDM Hub 支持为“匹配”作业的模糊匹配部分以及为“暂存”作业的清理进程部分实现负载均衡。</p> <ul style="list-style-type: none"> - 服务集成框架 (SIF) - 传出 JMS 消息 <p>注意: Informatica Data Director 在应用程序服务器群集中不支持负载均衡。群集环境中的负载均衡可能会带来意外结果。要增强 MDM Hub 环境的性能，您可使用外部负载均衡器。</p>
可维护性	<p>比独立应用程序服务器实例拓扑更复杂，但相比分布式应用程序服务器拓扑而言更易于部署和维护。使用 WebLogic 管理服务器时，您可以在群集中的 WebLogic 受管服务器之间轻松部署 MDM Hub 组件。</p>

JBoss 群集拓扑

安装拓扑示例含有构成了两个 JBoss 群集的二台计算机。每个 JBoss 群集都包括两个节点。Hub 服务器实例部署在一个群集中的每个节点上，因此一个节点发生故障时，群集中的另一个节点会接管。进程服务器实例部署在另一个群集中的每个节点上，因此一个节点发生故障时，群集中的另一个节点会接管。Hub 服务器负责在两个进程服务器实例之间分发处理负载。如果进程服务器实例发生故障或脱机，Hub 服务器会将处理请求发送至联机的进程服务器实例。Hub 存储在安装了数据库服务器的第三台计算机上配置。

注意: 您无需将进程服务器实例部署在群集中。您可能希望将进程服务器实例部署在群集中以方便部署，但每个进程服务器实例都必须向 Hub 服务器注册。使用 JMS 消息队列时，为了使用传出 JMS 消息，可将 Hub 服务器实例部署在群集中。否则，每个应用程序服务器实例都有不同的传出 JMS 目标。

下图显示了 JBoss 群集安装拓扑示例：



下表介绍了应用程序服务器群集拓扑的功能。

功能	可用性
高可用性	<p>是。</p> <p>MDM Hub 支持为以下操作和组件提供高可用性：</p> <ul style="list-style-type: none">- Hub 控制台操作- 服务集成框架 (SIF)- 传出 JMS 消息 <p>MDM Hub 不支持为以下操作和组件提供高可用性：</p> <ul style="list-style-type: none">- 批处理作业 <p>注意: 如果群集中的某个节点故障，通过 Hub 控制台发出的批处理作业请求会故障转移到活动节点，但批处理作业本身不会故障转移。</p> <ul style="list-style-type: none">- Informatica Data Director
可扩展性	<p>是。</p> <p>要扩展 MDM Hub 以支持大量数据，请添加更多 MDM Hub 组件。此外，要并发处理多个请求，请为进程服务器配置多个线程。</p> <p>MDM Hub 支持为以下操作和组件使用多线程处理：</p> <ul style="list-style-type: none">- Hub 控制台操作- 批处理作业- 服务集成框架 (SIF)

功能	可用性
负载平衡	<p>是。要实现负载平衡，您无需将进程服务器实例部署到应用程序服务器群集。进程服务器实例使用内部负载平衡机制。</p> <p>MDM Hub 支持为以下操作和组件实现负载平衡：</p> <ul style="list-style-type: none"> - Hub 控制台操作 - 除“生成匹配标志”作业以外的所有批处理作业 <p>注意: MDM Hub 支持为“匹配”作业的模糊匹配部分以及为“暂存”作业的清理进程部分实现负载平衡。</p> <ul style="list-style-type: none"> - 服务集成框架 (SIF) - 传出 JMS 消息 <p>注意: Informatica Data Director 在应用程序服务器群集中不支持负载平衡。群集环境中的负载平衡可能会带来意外结果。要增强 MDM Hub 环境的性能，您可使用外部负载平衡器。</p>
可维护性	<p>无。</p> <p>注意: MDM Hub 支持为 JBoss 群集使用独立模式。不同于域模式群集，独立模式群集无法管理配置和部署。</p>