

# 数据管控产品规格说明书

# 第一章 引言

## 1.1 编写目的

本文档作为数据管控产品体系(元数据和数据质量管理)的产品方案说明书，介绍元数据和数据质量的解决方案。

## 1.2 项目背景

经过多年建设，企业一般已经初步建成了各类管理信息系统，虽然在运用和分析数据支持经营决策方面已初见成效，但是对比战略发展要求和国内外最佳实践经验，还是存在数据管控水平不高，以及配套体系建设相对滞后的问题。为解决这些问题，企业会做种种努力，但由于未采用系统性的管控治理，数据问题并未能从根本上解决，阻碍了企业管理精细化的进程。

参考国内外同业先进理念、做法，一般会从数据标准、数据质量、数据模型、数据分布、数据安全、数据生命周期管理等6个方面，采用全面规划、分步实施的策略，分阶段有序推进数据管控项目建设，包括构建统一的数据管控制度体系、优化数据管控流程、实施有针对性的数据管控配套系统建设，实现对数据资产的全面管理和深度利用，进而提升数据资产管理水平和信息服务水平，形成差异化的竞争优势和核心竞争力。

# 第二章 方案概述

## 2.1 方案目标

本方案主要完成以下工作：

- ① 完成元数据管理系统和数据质量评估管理系统的引进、集成部署和客户化定制；
- ② 完成风险相关数据标准在元数据管理系统的落地；
- ③ 完成风险相关数据质量的评估及数据质量评估分析模型的开发。
- ④ 完成其他业务需求中明确的试点系统的数据标准发布和质量评估落地工作。

## 2.2 项目范围

### 2.2.1 实施内容

#### 1) 元数据管理系统和数据质量评估管理系统相关基础性工作

根据业务需求，结合企业现有系统的情况，制定具体项目实施方案，确保能完成相关咨询成果在系统内的部署和设置。完成系统接口设计、系统架构设计和形成实施所需的需求规格化文档等工作。完成产品的集成安装和初步调试工作。若提供的软硬件配置建议书不能完全满足企业软硬件选型需要的，需要协助事先完成必要的产品测试工作，确定最终的系统软硬件配置清单。

#### 2) 元数据管理系统和数据质量评估管理系统的客制化开发实施

根据企业的业务需求，对产品或应用进行客制化实施。根据对业务、技术元数据进行抽取和加载，在系统中建立元数据之间的映射关系。根据评分模型对评分卡进行客制化开发，并能将评分结果数据在第三方报表平台上展现。本阶段需要完成系统的概要设计、详细设计；完成数据接口设计和数据 ETL 工作；完成产品的配置，应用和报表的客制化开发等。产品客制化过程中，出现在场人员无法解决的技术或产品问题时，投标人应及时组织国内国外资源，快速解决问题，不能因此对项目进度造成影响。

#### 3) 元数据管理系统和数据质量评估管理系统的测试

完成元数据管理系统和数据质量评估管理系统的各类测试工作，包括：单元测试、集成测试、压力测试和协助用户测试，测试工作需要符合企业测试规范要求。单元测试和集成测试的主要工作是：制定测试策略和计划、准备测试案例和数据、执行各类测试案例(含功能测试、接口测试、报表测试、系统管理和非功能性需求测试等)、分析测试结果、针对测试问题进行回归测试。压力测试的主要工作是：制定测试策略和计划、准备测试案例和数据、执行性能测试案例、分析测试结果、针对测试问题进行回归测试，完成系统调优工作。用户测试的主要内容是：做好测试数据准备工作，并协助完成测试。

#### 4) 元数据管理系统和数据质量评估管理系统的上线和试运行

完成元数据管理系统和数据质量评估管理系统的上线和试运行工作。元数据管理系统和数据质量评估管理系统上线的相关事宜，主要包括：确定系统软硬件、完成系统整体部署(含集成部署)、调试和调优、完成系统的初始化工作。完成符合企业规范的上线和移交文档。编写用户操作手册、业务手册、系统操作手册等。元数据管理系统和数据质量评估管理系统试运行相关工作，主要包括：试运行实施规划、试运行过程中的问题整理和处理、试运行结果分析等。维护系统、明确系统应急处理机制并演练等。

## 2.3 方案概述

数据管控系统建设方案元数据产品 **MetaOne** 系统和数据质量系统 **QualityOne**，并结合专业的业务和技术团队，旨在建立一套符合企业数据管控体系建设需求的数据管控体系系统。

数据管控体系建设项目依托元数据和数据质量管理体系，旨在建立一套符合企业特色的数据管控体系系统，本项目实施主要完成数据管控系统咨询成果落地工作，通过数据标准配置；指标标准管理及分析预警；元数据获取和调度；数据质量分析和预警展示和数据探查分析和预警等模块，全方位全角度展示企业系统平台技术和业务元数据分析，最终，实现数据质量监控预警。

## 第三章 总体方案

### 3.1 系统架构

#### 3.1.1 总体系统架构

数据管控体系建设平台实施方案采用双模式进行调度配置，为保证第一时间有效监控业务风险相关业务指标监控情况，因此，总体方案中，首先采用企业数据仓库平台调度平台优先监控风险相关业务指标信息，其次元数据管理产品内部调度机制，加载相关平台，关系数据库(应用业务数据和技术业务数据)和数据模型等元数据信息。

信息交付层：主要展示元数据和数据质量成果信息，展示方式主要通过都必须通过 **AD**域验证(统一认证平台)，信息访问通过报表平台 **BIEE**和元数据系统内部展示平台展示元数据和数据质量分析预警信息。

应用层：主要提供企业 **BIEE**报表系统展示的数据接口，展示业务客户需要的业务元数据信息。

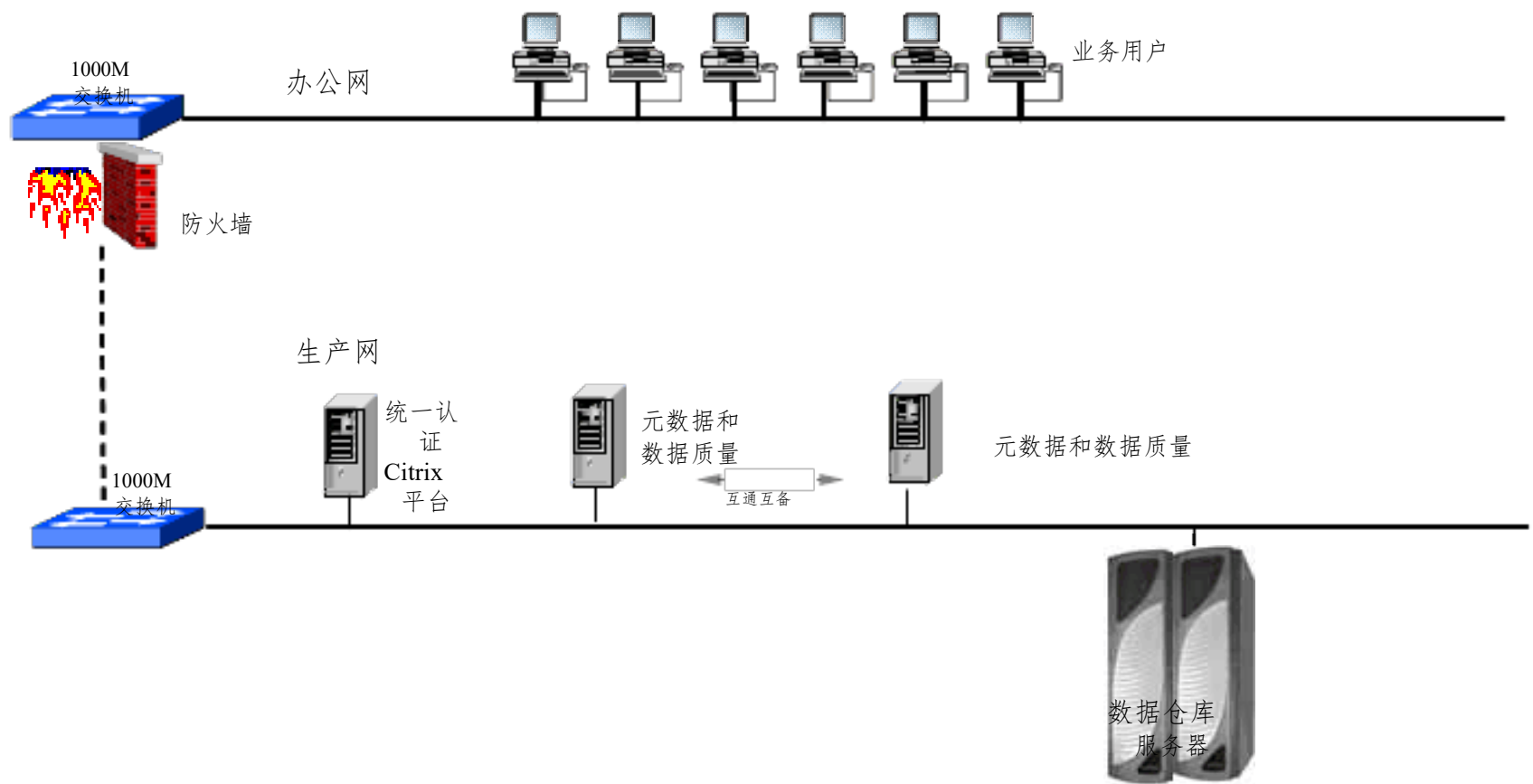
系统管理层：主要描述 **MetaOne**和 **QualityOne** 为用户提供基本功能完善的元数据仓储管理平台，并支持基于平台的应用服务定制。丰富的调用接口和强大的二次开发指南，保障产品良好的适用性。系统整体在功能上，划分为存储层、基本功能层、接口层、应用层。

数据获取层：主要描述系统获取元数据方式有两种触发方式，为实现监控风险相关业务指标，系统将采用部分任务配置在调度平台调度；其他数据获取方式将采用 **MetaOne** 定时触发和依赖满足触发任务，将所需的业务指标数据和平台技术数据安装规范加载至平台内部。

源数据层：主要描述数据管控体系的数据源信息，包括标准格的业务元数据；镜像业务系统关系数据库；ETL系统标准元数据信息；平台相关技术元数据信息；模型设计文件元数据和咨询涉及的其他标准元数据文件成果。

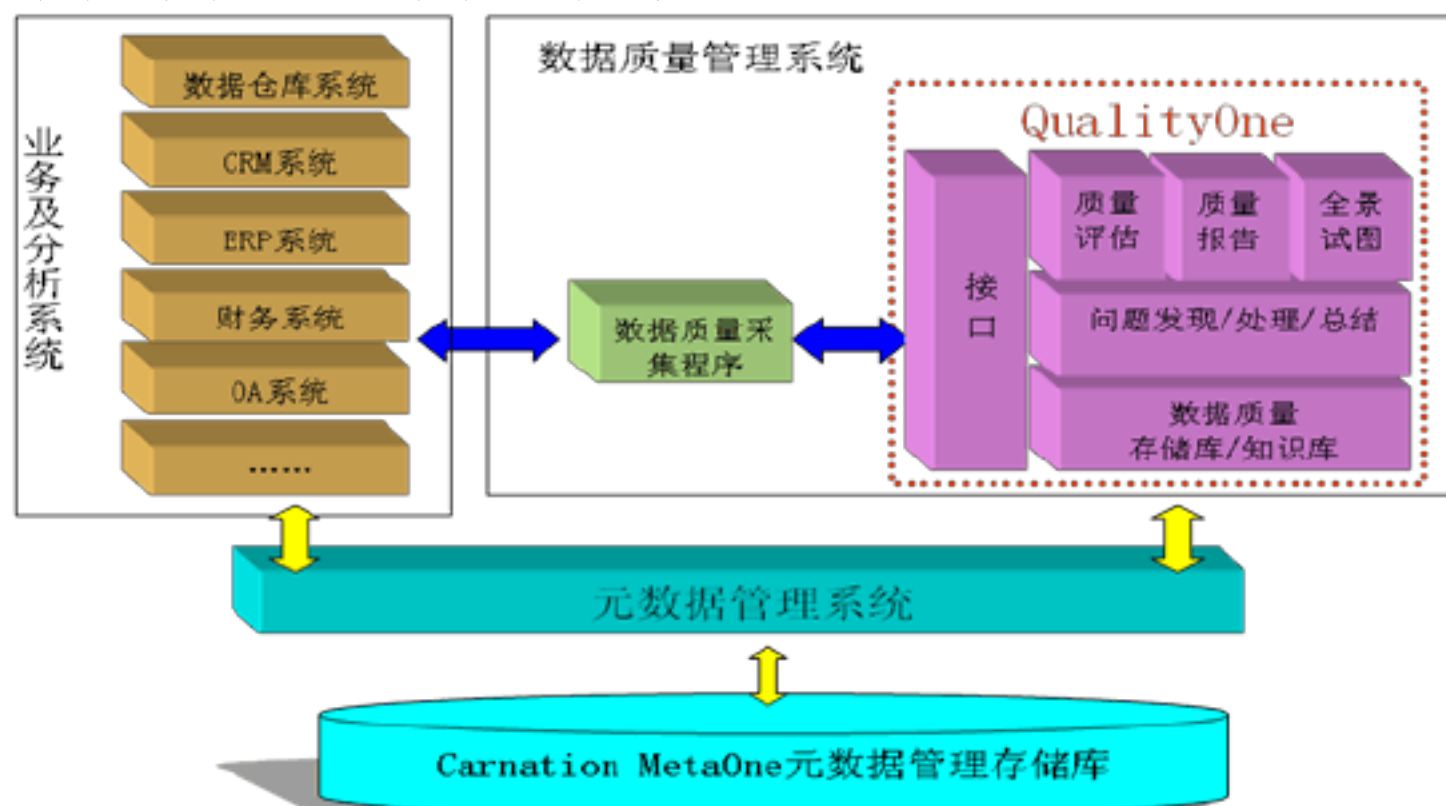
### 3.1.2 物理系统架构

数据管控体系系统物理架构采用应用服务和数据库服务器独立方式设计，同时，两台服务器具备内部互补自动切换功能，有效防止由其他因素导致的系统瘫痪。



### 3.2 系统数据流程

下面展示元数据和数据质量系统数据流程图示如下：



### 3.3 网络和安全

MetaOne元数据系统部署于企业的业务网段，通过代理将应用服务器地址反向映射办公网段，仅开通WEB服务端口，向位于办公网段的业务管理用户提供服务；其他业务系统的接入均在企业的业务网段内，与外部环境隔离。

#### 3.3.1 开放的 Internet 标准

系统建立在开放的 Internet 标准之上，例如：SSL, HTTP, XML, TCP/IP, SOAP 支持多层次的客户机/服务器体系结构，其通过 TCI 组件支持多种报文的传输、组包、解包；利用 webservice 实现与异构系统之间的数据交换。可以成功运行在支持 Java 虚拟机的各种操作系统及相应的各种硬件平台上，符合信息技术的发展方向，具备广泛的应用前景。

#### 3.3.2 系统层安全

系统层安全主要针对系统的操作系统和数据库等软件平台进行安全防护，其主要采用的措施如下：

安装系统的安全补丁（Patch）。

关闭不需要的进程服务和端口。

使用漏洞扫描产品，定期进行安全扫描及时发现问题并采取补救措施。

定期审核系统配置，避免因配置不当造成的权限管理混乱。

#### 3.3.3 业务安全控制

通过安全代理前置设备、防火墙等系统来保证系统的安全性，以及通过负载均衡来保证系统的高可用性，这只是从网络环境和系统结构的角度保证系统的安全，整个系统的安全性应该是一个多层次的概念。

主要考虑一下几个方面：

关键操作有详细的操作记录可查。

登录系统日志的安全审计

多维的交易权限管理机制，关键操作提供多重授权功能。对于管理端的关键数据设置提供严格的复核机制，以保证数据的正确性、合法性。

会话管理 (Session)：与应用服务器的会话管理结合，实现多种会话的建立和管理，让不同的会话采用统一的管理机制。同时实现会话的超时管理，有效防范避免操作人员使用已经失效的会话进行误操作，同时防止垃圾会话数据占用内存，影响系统性能甚至使系统无法工作。

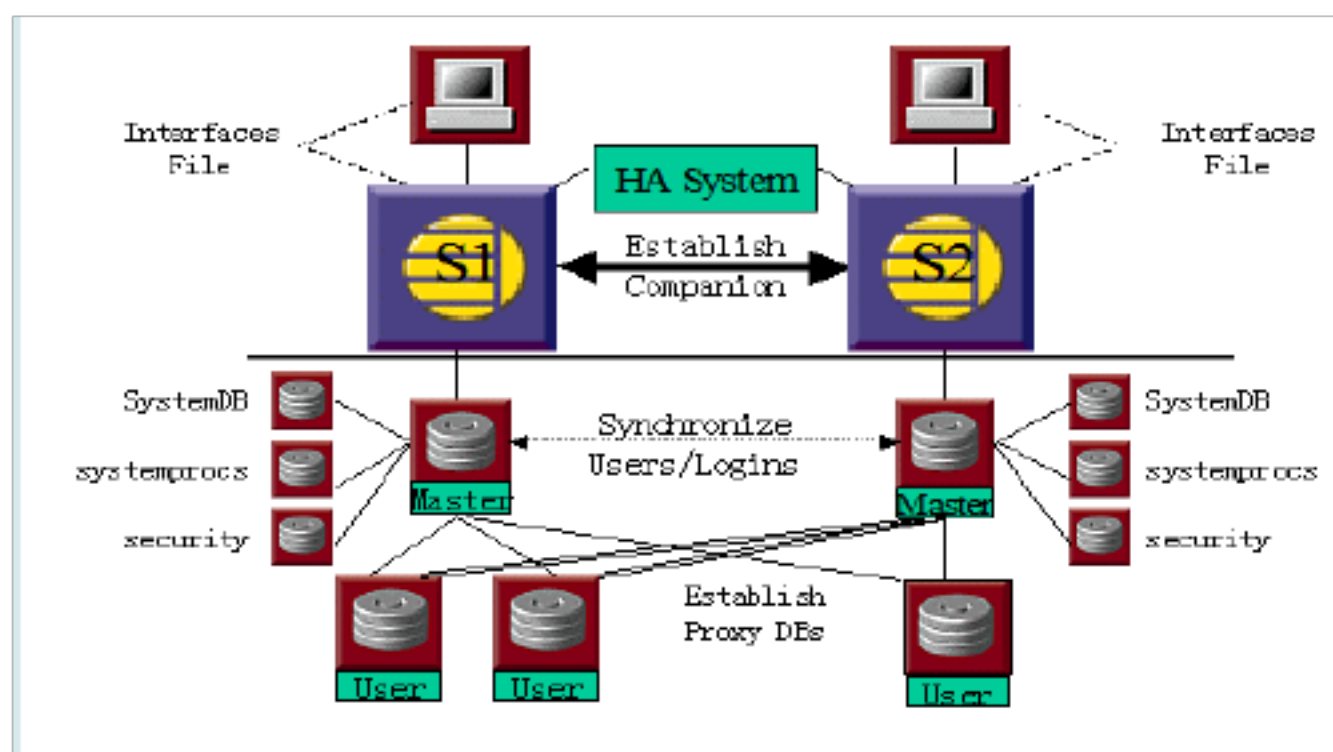
### 3.3.4 统一身份认证

针对统一身份认证系统的要求，系统采取一种适用（针对）于 WebLogicServer 的、简洁的、有效可行的统一认证即单点登陆 (SSO) 解决方案。这个方案以集中统一的用户信息为基础，但不包括上节说明的权限管理，而仅仅提供一个身份验证的服务。实现 SSO 的本义。

在 SSO 帮助用户完成身份验证后，由系统本身的权限管理进行用户行为的进一步控制。这样不仅完成了用户帐号的集中管理，又不失原有系统灵活的权限控制。而且方案简单可行，不需要对原有应用做较大更改，适合快速解决 Single Sign-On 问题，符合行内系统的要求。其主要采用 SAM Browser Post Profile 方案，通过推模式获取 SAML Token 由 Identity Provider 推 Assertion 到 Service Provider. 而后验证 Assertion 将用户重新定向到目标 URL

## 3.4 系统灾备

为了保障对数据处理的高可靠性和高可用性，系统需要具有较好的实时备份和灾难恢复能力。系统采用 SQL SERVER 2008 数据库服务器支持的双机热备容错方案来实现系统的备份策略，由两台机器共享一个磁盘阵列（共享磁盘组），当其中一台机器发生故障时，另外一台机器接管整个磁盘阵列，从而实现双机容错。如下图：

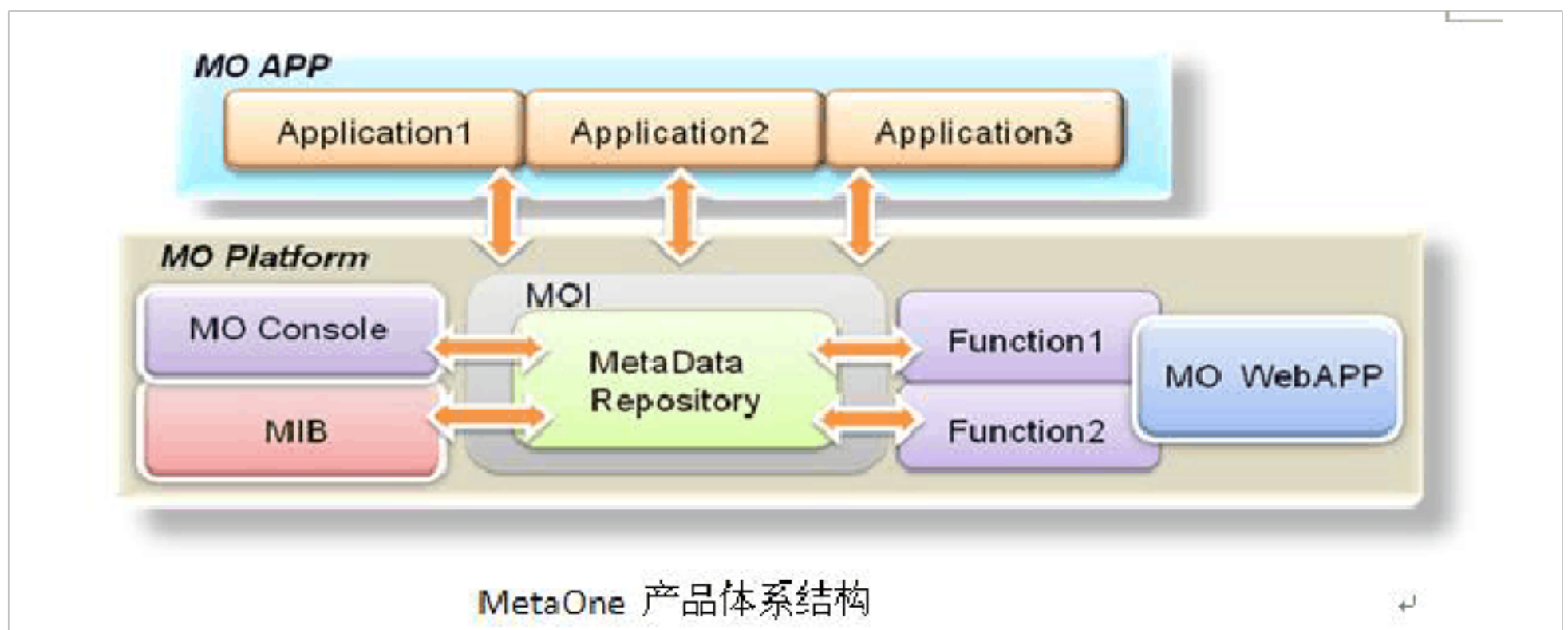


## 第四章 方案细化

### 4.1 产品功能

#### 4.1.1 元数据

元数据产品 **MetaOne** 采用高内聚、低耦合的组件式产品架构，丰富功能组件，为客户搭建功能强大的元数据管理平台。向用户全面开放元数据功能调用接口，并提供整套实施方法论。使用户在完成元数据管理、维护等基础功能的同时，方便实现二次开发，满足行业应用的针对性需求。



#### □ 基本管理平台（MetaOne Platform）

由元数据仓储、Web应用程序、元数据接口层、统一数据源管理客户端、元模型操控台以及元数据基本维护管理功能组成的组件集合，是 MetaOne 产品的基本功能集。

#### □ 元数据存储库（MetaData Repository）

基于关系数据库的元数据存储库，用于实现元数据的物理存储。

#### □ 元数据基本维护界面（MetaOne WebAPP）

B/S 结构的 Web 管理程序，实现用户对元数据的一般管理、维护操作。

#### □ 对外服务接口（MetaOne Interface（MOI））

一组基于 Java RMI 的 API 集合，内容包括对元数据、元模型的所有基本操作，配套提供使用方法论，使用户方便调用，实现二次应用开发。

#### □ 元数据获取器（MetaData Integration Bridge（MIB））

数据源管理平台，对所有连接系统的数据源进行统一管理，实现数据源元数据的抽取和元数据向 **MetaData Repository** 的导入。

元模型管理控制操作台（MetaOne Console）

元模型操控台，用户通过操控台实现元模型建模。

基础应用分析

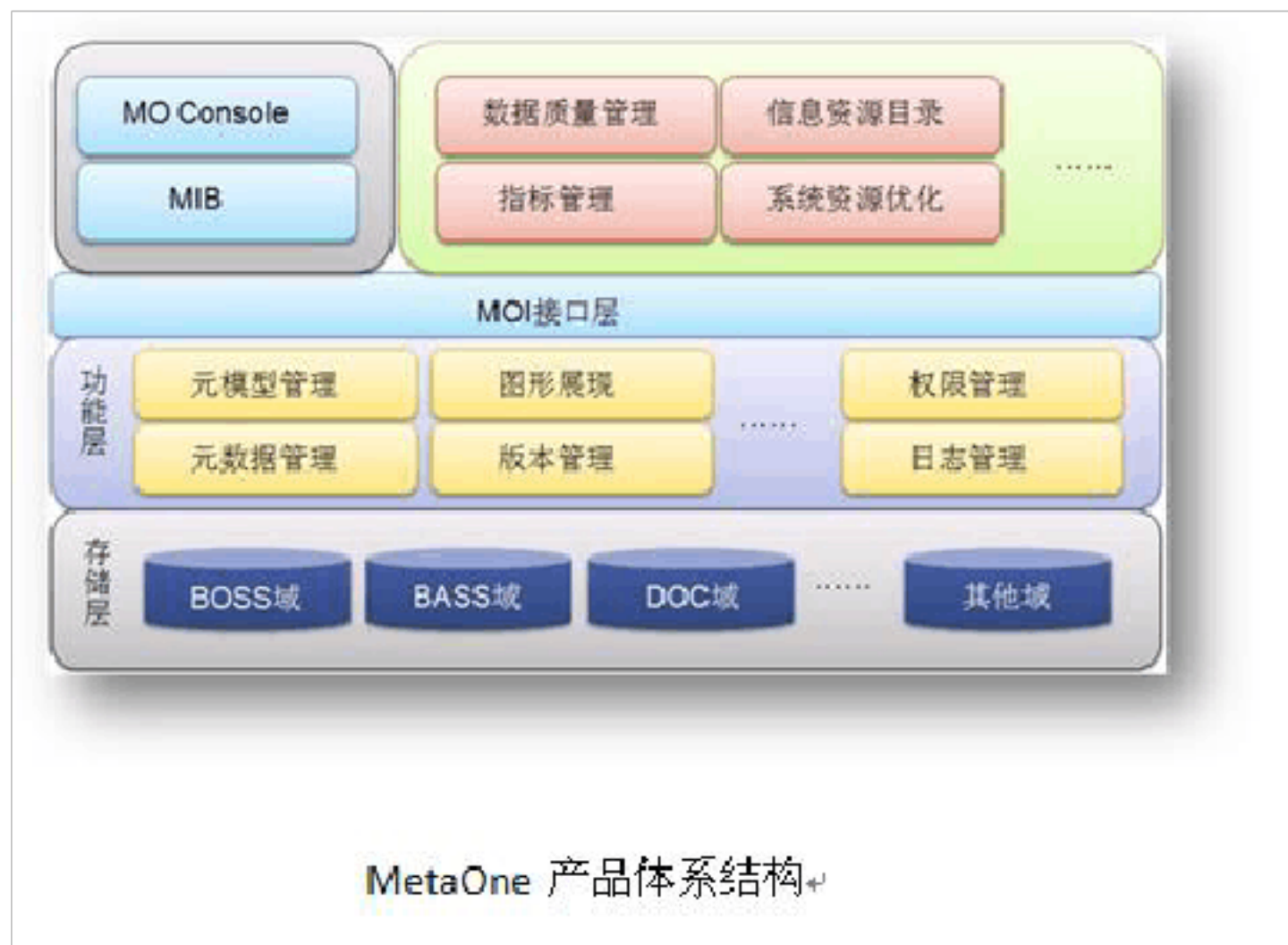
元数据的基本维护管理功能

高级应用分析

通过元数据接口，调用元数据及相应功能，针对行业需求开发的应用。

#### 4.1.1.1 功能特性

**MetaOne** 的产品理念，是为用户提供基本功能完善的元数据仓储管理平台，并支持基于平台的应用服务定制。丰富的调用接口和强大的二次开发指南，保障产品良好的适用性。系统整体在功能上，划分为存储层、基本功能层、接口层、应用层。



元数据存储层

包含元数据存储库，负责承载不同领域的元模型以及相关的元数据。

元数据基本功能层

实现系统对元数据的基本功能，针对元数据自身的维护和管理。包括元数据的维护管理、元模型

维护管理，影响分析、血统分析、差异分析等图形展现功能，元数据版本管理、权限管理、日志管理、质量管理等。

接口层

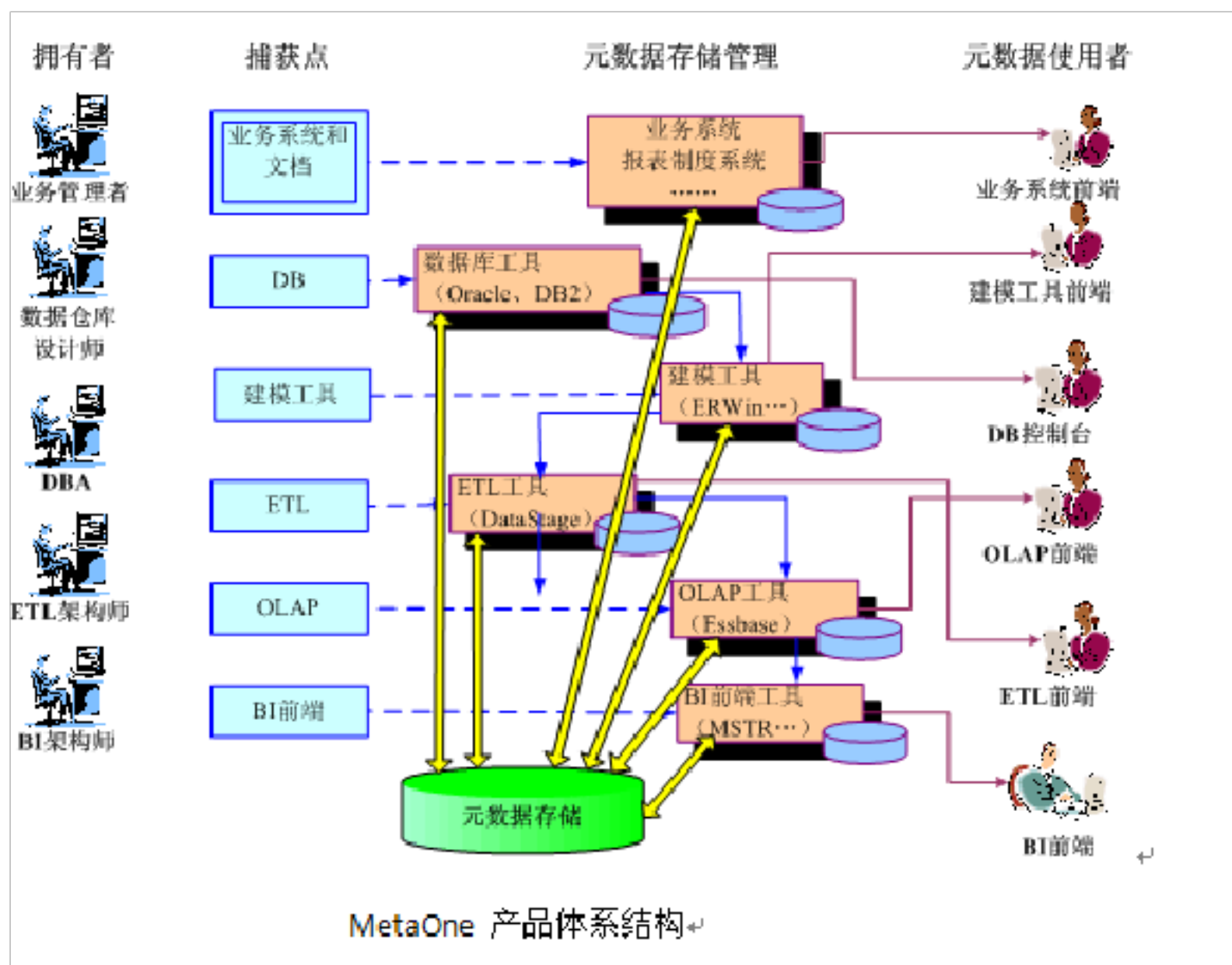
MOI为 MetaOne提供的元数据接口集合，供二次开发对系统各类数据、功能的调用。

□ 应用层

通过 MOI接口层与元数据存储层实现连接的部分属于应用层。包含两个方面，Metaone 提供的功能组件 MO Console和 MIB 以及针对不同的具体需求开发的应用功能。

#### 4.1.1.2 数据流程

MetaOne元数据管理系统将涵盖系统数据处理流程各个环节的数据结构等信息读取进来，包括源系统信息、ETL过程、数据库结构、数据模型、业务应用、前端展示和门户管理等，并加入业务规则的描述和业务量值的内涵，下面图示详细介绍数据获取流程信息：

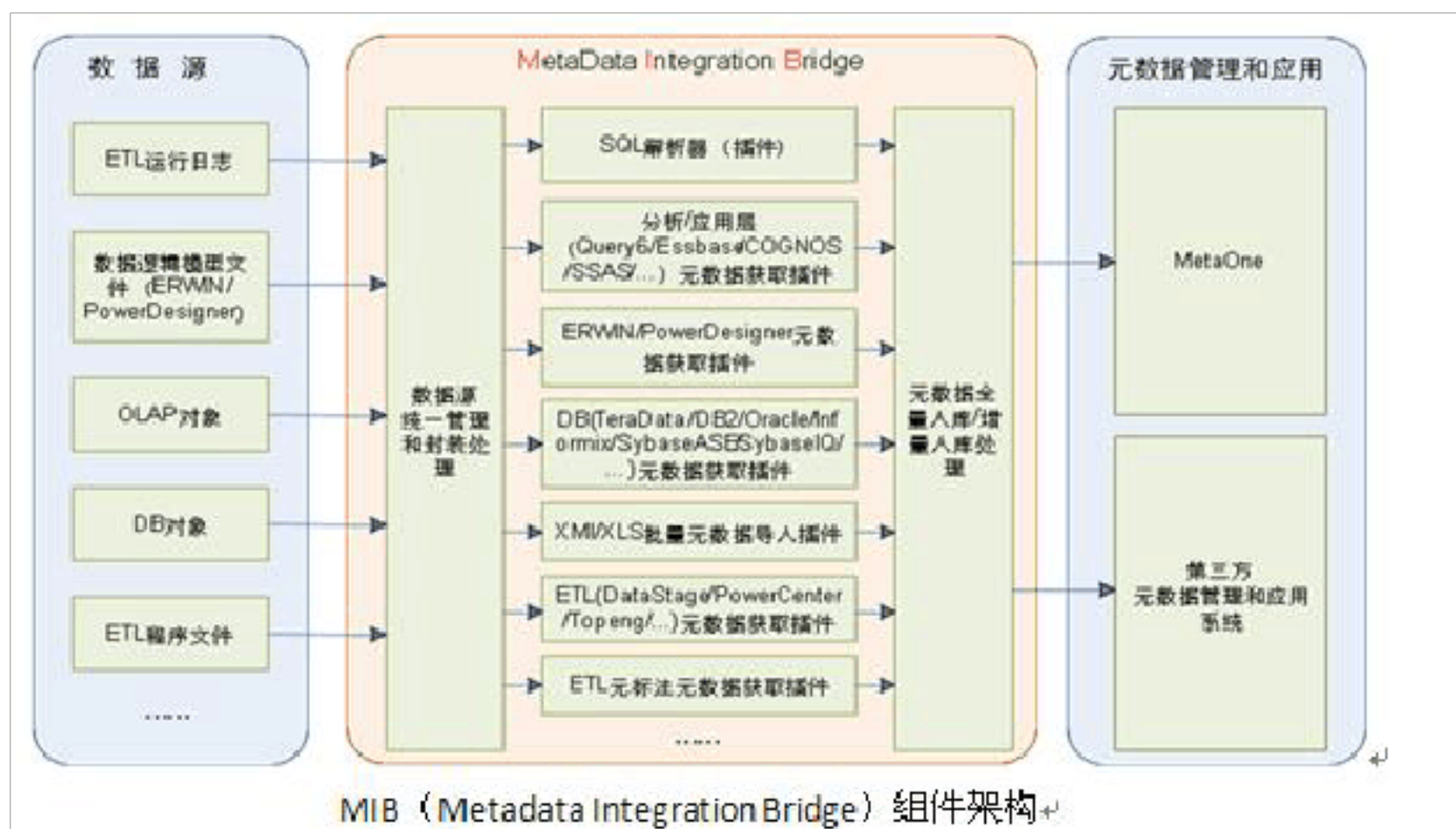


### 4.1.1.3 数据获取

MetaOne 元数据管理系统的元数据管理范围，涵盖系统数据处理流程各个环节，各个环节中数据源所能提供的数据类型、结构都需要相对应的元数据获取方法。MetaOne 元数据管理系统提供统一的数据源管理平台，平台中集成了针对不同数据类型的元数据获取功能插件，可以将包括源系统信息、ETL过程、数据库结构、数据模型、业务应用、前端展示和门户管理等数据源进行统一管理，实现元数据的自动获取。

MetaOne 提供元数据整合获取平台元数据集成桥（Metadata Integration Bridge，元数据集成桥），提供自动获取、手工获取两种元数据获取功能，自动获取主流 BI 工具的元数据或手工获取 XMI/Excel 文件中的元数据。

MIB具备强大的管理能力和可扩展性，能够对自动获取的数据源进行管理，能够提供自动获取的框架，在此框架可以扩展自动获取的适配器；对元数据的自动获取提供持续稳定的调度支持，能够按预设的调度策略触发相应的元数据自动获取过程，同时支持对自动获取过程的日志记录。MIB体系架构如下图所示：



#### MIB数据接口

元数据集成桥能够自动获取主流 BI 工具的元数据，并支持通过对 SQL脚本、TCL脚本的词法、语法和语义分析，生成满足 CWM规范要求的元数据，存储到 MetaOne 中进行统一管理。通过 MIB MetaOne 可自动获取系统项目的相关工具所涉及的元数据，确保元数据及时更新与业务系统同步，减少人为的手工梳理过程。

## MIB技术特点

元数据集成桥通过把不同数据源的元数据抽取、转换到统一的元数据系统中，能够实现企业级的元数据整合、集中和统一，提升系统的元数据管理水平，其技术特点如下：

1. 跨平台：服务器支持 Windows Unix、Linux(Windows , Linux, IBMAIX, SUNSolaris) , RCP Client 支持 Windows Linux、Mac OS X系统；
2. 兼容性：数据存储于主流关系型数据库系统，支持 Oracle , DB2 SQL Server 等；
3. 灵活及易用性：元数据集成桥提供了友好的图形界面，用户只需要配置相应的参数就可以完成本次导入。用户可以选择需要导入的对象列表，设置过滤规则，选择元数据的根节点，无需要编程就能够完成各省的特性化需要。
4. 主流开放性：100%JAVA按 J2EE 模式构建、基于 RCPClient 的客户端架构，采用 XMLRPC和 JSON RPC通信方式，支持各类编程语言开发的分布式应用程序间的通信，如 C、JAVA Perl 、PHP .NET 等；
5. 全中文文化和国际化：界面操作全部显示中文。可以根据需要按照区域化的配置显示客户熟悉的语言。
6. 后台自动调度： 通过后台调度自动定期执行导入作业。
7. 支持增量导入和自动关系维护： 同一个数据源多次导入时，可以支持以增量方式进行导入。减少数据处理量，提高效率。
8. 自动关系维护： 通过元数据的属性或者解析 SQL的方式，能够自动的建立元数据之间的关系。
9. 支持大数据量： 为了能够支持超大数据量 (>200M) 的元数据导入，系统优化了 XML解析和数据比较，并且支持多个 Java 虚拟机 (JVM) 的数据导入，最大限度的利用了硬件系统的资源。

### 4.1.1.4 数据模型

元数据管理系统具有独立的元模型管理控制平台，实现元模型的建立、关系的维护等功能。元模型架构支持国际 CWM规范、提供行业预定义元模型、支持用户按需完全定制，从而保障用户随着时间推移对更多种类元数据的管理需要，以及完善现有元数据定义的需要，能满足企业中长期发展的需求，提高投资回报率。

元模型管理功能包括：

提供图形化元模型管理控制台 (Console)

以上内容仅为本文档的试下载部分，为可阅读页数的一半内容。如要下载或阅读全文，请访问：<https://d.book118.com/198025005105006123>