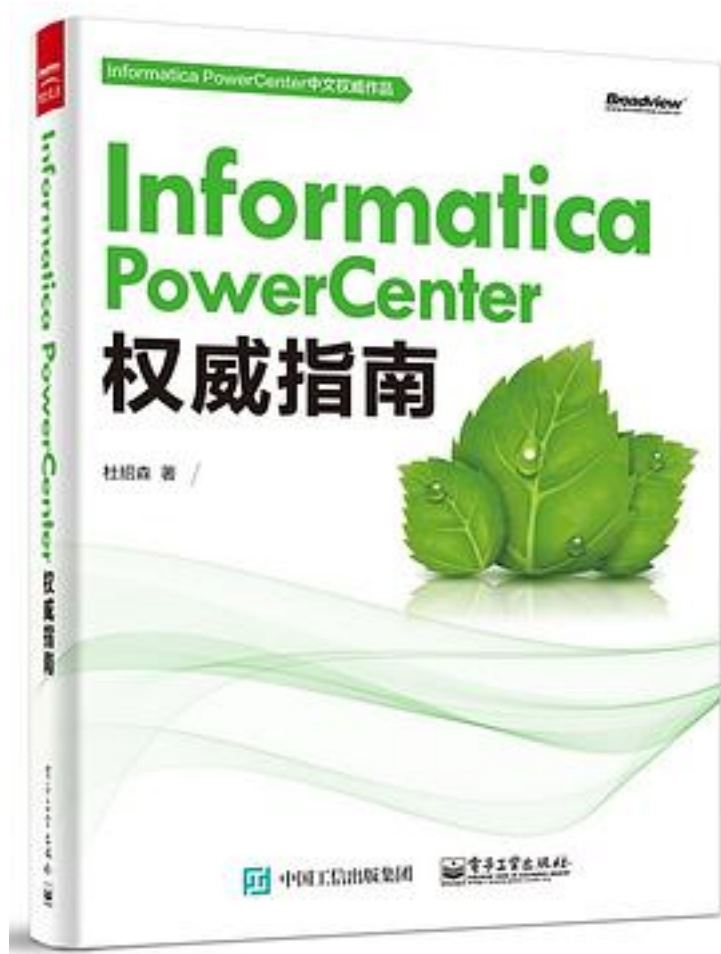


Informatica PowerCenter权威指南



[Informatica PowerCenter权威指南_下载链接1_](#)

著者:杜绍森

出版者:电子工业出版社

出版时间:2015-9-1

装帧:平装

isbn:9787121270451

大数据时代，掌控数据首先需要掌握数据的处理能力。俗话说，“工欲善其事，必先利其器”。InformaticaPowerCenter作为业界广泛使用的数据处理工具之一，被全球多数

大型机构、组织认可并采用。《InformaticaPowerCenter权威指南》全面地介绍了InformaticaPowerCenter的主要功能及高级特性。

《InformaticaPowerCenter权威指南》分为3个部分：第一部分——基础篇，包括第1～4章系统介绍了PowerCenter的基础组件和常用功能，并在其中穿插了大量实践案例；第二部分——高级篇，包括第5～8章，系统介绍了PowerCenter并行、集群、性能调优和字符集管理等高级内容；第三部分——扩展篇，包括第9章，简要介绍了CDC，与Hadoop、MPP集成，以及非结构化和半结构化数据处理能力。

《InformaticaPowerCenter权威指南》适合PowerCenter的入门者及有一定PowerCenter使用经验的用户参考，也可作为各数据仓库、大数据专业培训机构的培训教材。

作者介绍:

杜绍森，任Informatica大中国区资深顾问。从PowerMart 5开始使用Informatica产品，自PowerCenter 8.1.1版本发布就一直任职于Informatica公司，服务于Informatica近10年时间，帮助Informatica用户设计和构建系统，或者对系统进行重构和调优。之前曾作为系统架构师、项目经理等服务于惠普及本地公司，从事电信行业计费及数据仓库项目。

目录: 目录

第1章 PowerCenter Hello World世界 1

1.1 Informatica Hello World 1

1.2 PowerCenter架构和客户端简介 3

1.2.1 PowerCenter架构 3

1.2.2 PowerCenter客户端 6

1.3 PowerCenter Hello World 7

第2章 PowerCenter基础组件 27

2.1 Source 27

2.1.1 数据库源 28

2.1.2 文本文件源 30

2.2 Target 33

2.2.1 数据库目标 33

2.2.2 文本文件目标 34

2.3 Expression表达式 35

Expression中的变量端口（Variable Port） 40

2.4 Filter 41

2.5 Source Qualifier 43

2.5.1 Source Qualifier的作用 43

2.5.2 数据库数据源的Source Qualifier 44

2.5.3 Source Qualifier自定义SQL 47

2.5.4 Source Qualifier复杂关联 48

2.6 Sorter 49

2.7 Joiner 51

2.7.1 关联类型 52

2.7.2 Sorted Joiner 54

2.7.3 Joiner有什么独特作用 55

2.7.4 自关联（Self-Join） 56

2.8 Lookup 57

2.8.1 Lookup Caching enabled 59

2.8.2 非连接的Lookup 61

2.8.3 Lookup SQL Override 63

- 2.8.4 共享Lookup Cache 65
- 2.8.5 Dynamic Lookup 65
- 2.8.6 Lookup、Source Qualifier和Joiner的对比 69
- 2.9 Stored Procedure 70
 - 2.9.1 Connected Stored Procedure 70
 - 2.9.2 Unconnected Stored Procedure 72
 - 2.9.3 Pre- or Post-Session Stored Procedure 74
- 2.10 Union 76
- 2.11 Transaction Control 78
 - 2.11.1 Transaction Control有效性问题 79
 - 2.11.2 Transaction Control组件 80
- 2.12 Sequence 80
 - 2.12.1 Sequence的常规用法 80
 - 2.12.2 共享Sequence 82
 - 2.12.3 可重用的Sequence 83
- 2.13 Aggregator 84
 - 2.13.1 条件聚合 85
 - 2.13.2 使用Aggregator进行行列转换 86
- 2.14 Rank 88
- 2.15 Update strategy 90
 - 2.15.1 Treat source rows as属性的使用 91
 - 2.15.2 Update strategy使用 93
 - 2.15.3 如何实现Update else Insert 94
 - 2.15.4 Update Stagy案例：缓慢变化维 98
- 2.16 SQL Transformation 104
 - 2.16.1 Script Mode 104
 - 2.16.2 Static Query Mode 106
 - 2.16.3 Dynamic Query Mode 108
- 2.17 Java Transformation 109
 - 2.17.1 Java Transformation简介 109
 - 2.17.2 Passive Java Transformation 114
 - 2.17.3 Active Java Transformation 121
 - 2.17.4 常见错误说明 123
- 2.18 Normalizer 124
- 2.19 Router 126
- 2.20 Custom Tranformation 128
- 2.21 HTTP Transformation 129
- 2.22 XML组件组 132
- 2.23 Transformation中的一些概念 135
 - 2.23.1 Connect与Unconnect 135
 - 2.23.2 Active与Passive 136
- 第3章 Workflow执行、监控 138
 - 3.1 Session 139
 - 3.1.1 Reusable Session 139
 - 3.1.2 非Reusable Session 141
 - 3.2 最简单、最常用的Workflow 143
 - 3.2.1 并行执行 143
 - 3.2.2 串行执行 144
 - 3.2.3 调度 146
 - 3.3 Worklet 147
 - 3.4 Command 148
 - 3.5 Control 150
 - 3.6 发送E-mail 151
 - 3.6.1 配置发送E-mail 151

3.6.2 Workflow中使用E-mail	151
3.7 Event Tasks	155
3.7.1 用户自定义事件使用	156
3.7.2 预定义事件使用	158
3.8 Timer	159
3.9 Decision	159
3.10 Assignment	160
第4章 常用功能汇集	163
4.1 Debugger	163
4.2 Mapplet/Reusable Transformation	165
4.2.1 Reusable Transformation	165
4.2.2 Mapplet	167
4.3 使用Shortcut	169
4.3.1 Local Shortcut	170
4.3.2 Global Shortcut	171
4.4 Session相关属性	173
4.4.1 Properties Tab相关属性	173
4.4.2 Config Object Tab相关属性	174
4.5 参数和变量	176
4.5.1 Mapping参数	176
4.5.2 Mapping变量	180
4.5.3 系统/Session参数与变量	184
4.5.4 Workflow/Worklet变量	189
4.5.5 Local 变量 (Local Variables)	191
第5章 PowerCenter高级应用	193
5.1 任务分区 (Partition)	193
5.1.1 Database Partitioning	196
5.1.2 Hash Partitioning	201
5.1.3 Key Range Partitioning	204
5.1.4 Pass Through Partitioning	205
5.1.5 Round-robin Partitioning	211
5.2 内存管理	214
5.2.1 DTM内存	215
5.2.2 Transformation Cache	216
5.3 网格计算	219
5.3.1 Grid架构	219
5.3.2 Grid负载均衡	221
5.3.3 Grid与任务分区 (Partition)	224
5.4 高可用性 (HA)	227
5.4.1 PowerCenter自带的HA方案	228
5.4.2 依托第三方厂商的HA方案	229
5.4.3 两种HA方案对比	230
5.5 Web Service 应用	230
5.5.1 Web Service Hub	231
5.5.2 Web Service调度/监控接口	232
5.5.3 Web Service Provider	234
5.5.4 Web Service Consumer	246
5.6 Pushdown Optimization	251
5.6.1 Pushdown优化是什么	252
5.6.2 Pushdown优化类型	252
5.7 版本控制及部署	256
5.7.1 Check In/Check Out	256
5.7.2 Team-Based开发的一些有用功能	258
5.7.3 Label与Deployment Group	260

5.7.4 复制对象从开发Repository到生产Repository 264

第6章 PowerCenter实战汇总 266

6.1 PowerCenter字符集 266

6.1.1 Oracle数据库 267

6.1.2 DB2字符集 268

6.1.3 AS/400字符集 268

6.1.4 ODBC字符集 269

6.1.5 文本文件字符集 270

6.1.6 Repository Service字符集 271

6.1.7 Integration Service字符集 272

6.1.8 Data Movement Mode 273

6.2 UNIX ODBC配置 274

6.2.1 ODBC常规配置 274

6.2.2 My SQL社区版ODBC配置 276

6.3 使用Mapping动态分发文件 277

6.4 超越EDW，商品自动价格跟踪 279

6.5 pmcmd命令 283

6.6 pmrep命令 284

6.7 infasetup命令 284

6.8 Mapping Architect for Visio 286

6.9 MX View语句 293

6.10 PowerCenter与其他工具集成 294

第7章 性能调优 297

7.1 性能调优过程 298

7.2 发现瓶颈 299

7.2.1 定位目标写瓶颈及调优 301

7.2.2 定位源读瓶颈及调优 302

7.2.3 定位Mapping/Session瓶颈 303

7.2.4 定位系统瓶颈 305

7.3 Mapping调优 305

7.3.1 Transformation优化 305

7.3.2 列级别的优化 310

7.3.3 其他方面的优化 312

7.4 Session调优 313

7.4.1 内存调优 313

7.4.2 PowerCenter高级特性支持高性能 313

7.4.3 其他手段 314

7.5 SQL Override调优 316

第8章 PowerCenter Troubleshooting 317

8.1 安装、启动过程的错误 317

8.2 开发过程的错误 319

8.3 Session运行错误 320

8.4 源读或者目标写的错误 321

第9章 PowerCenter扩展能力 322

9.1 PowerExchange CDC（变化数据捕捉） 322

9.1.1 PowerExchange CDC的3种模式 323

9.1.2 开放数据库CDC基本原理 325

9.1.3 CDC常见的一些讨论 326

9.1.4 CDC Real-Time for Oracle安装配置（实例） 327

9.1.5 CDC定义注册组和添加捕获注册（实例续） 331

9.1.6 CDC Mapping开发及运行（实例） 334

9.2 PowerCenter与SAP 336

9.2.1 R/3、mySAP、ECC 337

9.2.2 PowerCenter与BW 338

9.3 PowerCenter与MPP数据库 339

9.4 PowerCenter与Hadoop 340

9.4.1 接口能力 341

9.4.2 PowerCenter on Hadoop 344

9.5 元数据管理与业务术语管理 345

9.5.1 元数据的血缘分析 346

9.5.2 元数据影响分析 347

9.5.3 业务数据管理 347

9.6 B2B Data Transformation 348

• • • • • (收起)

[Informatica PowerCenter权威指南_下载链接1](#)

标签

数据处理

金融

数据仓库

@编程/项目管理

评论

[Informatica PowerCenter权威指南_下载链接1](#)

书评

