



CD-ROM



WILEY

计算机技术
译林
精选系列

Microsoft OLAP

解决方案

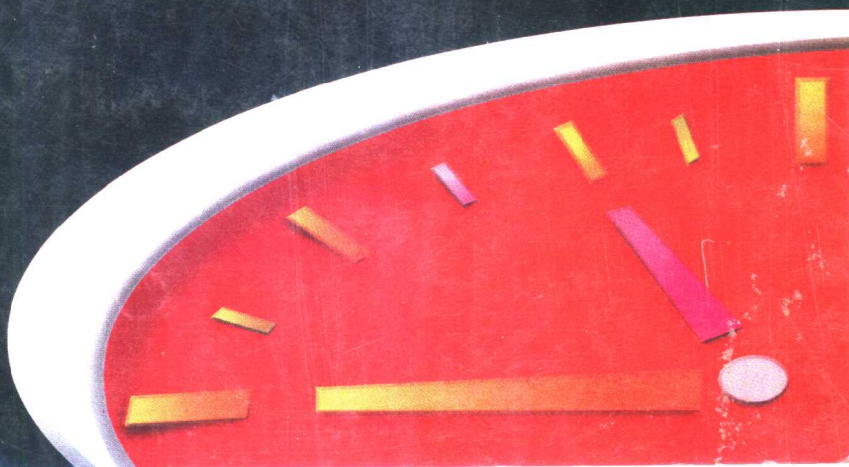
Erik Thomsen

〔美〕 George Spofford 著

Dick Chase

潇湘工作室 译

人民邮电出版社
www.pptph.com.cn



计算机技术译林精选系列

Microsoft OLAP 解决方案

Erik Thomsen

[美] George Spofford 著

Dick Chase

潇湘工作室 译

人民邮电出版社

计算机技术译林精选系列
Microsoft OLAP 解决方案

◆ 著 [美] Erik Thomsen George Spofford
Dick Chase

译 潇湘工作室

责任编辑 李 际

◆ 人民邮电出版社出版发行 北京市崇文区夕照寺街 14 号

邮编 100061 电子函件 315@ pptph.com.cn

网址 <http://www.pptph.com.cn>

北京汉魂图文设计有限公司制作

北京鸿佳印刷厂印刷

新华书店总店北京发行所经销

◆ 开本:787×1092 1/16

印张:28

字数:690 千字

2000 年 9 月第 1 版

印数:1-4 000 册

2000 年 9 月北京第 1 次印刷

著作权合同登记 图字:01-1999-2800 号

ISBN 7-115-08614-1/TP·1695

定价:60.00 元

版权声明

Erik Thomsen, George Spofford, Dick Chase: Microsoft
OLAP Solutions

Copyright © 1999 by Erik Thomsen, George Spofford,
Dick Chase. All Rights Reserved. Authorized translation
from the English language edition published by John
Wiley & Sons, Inc.

Erik Thomsen, George Spofford, Dick Chase 版权所
有。

本书授权自 John Wiley & Sons 公司出版的英文版
本翻译，中文简体字版由人民邮电出版社独家出版。

本书详细介绍了 OLAP 解决方案,其主要内容有: Microsoft OLAP Services 提供的逻辑 OLAP 结构、产品的物理和计算结构以及数据库设计和实现的基础;用于查询和计算的 MDX 语言,包括所有已有的 MDX 函数和运算符的参考说明;高级主题,例如维护、数据库优化和高级数据库设计;应用 OLAP 的实例和分析。

本书适用于数据库管理员和分析员。

译者序

在计算机技术中，需要用各种方法来表述现实世界中事物之间的关系。例如，在我们熟悉的关系数据库中，是以表格式的二维结构来表示信息的结构。而实际上，现实世界的构成是相当复杂的，在许多情况下，二维结构不足以描述实际的结构，我们需要的是三维甚至是多维的表达方式。

Microsoft OLAP Services 就是一个多维信息系统，其中的所有信息都是以 OLAP 结构方式建立模型的。它具有多维分析和查询能力。利用 OLAP，用户几乎可以立即从多维数据仓库中创建、更改和访问信息。OLAP 技术能够简化复杂的分析。以前看起来很困难的或不可能实现的任务，现在都可以轻松地实现了。

本书的主要内容如下：

首先，我们要学习一些新的概念，如维、量度、层次、级别、立方体、单元、公式和链接等，正是这些概念构成了 OLAP 数据库的逻辑结构及其支持的计算。

随后，作者以逻辑方式循序渐进地说明有关维护、数据库优化和高级数据库设计中的关键内容，如 Microsoft OLAP 服务器的存储和运作结构、使用 MDX 语言、使用 Microsoft OLAP 服务器来访问和更新数据、指定计算、优化服务器性能以及服务器的安全保护等。

最后，作者给出了两个实例，一个是健康保险分析，另一个是杂货店的常客分析程序，来深入说明如何建立 OLAP 应用程序。

另外，本书还配有光盘，其中的内容有在本书中介绍的全部样例应用程序和数据库，还有使用 Microsoft OLAP 服务器的其他软件。

参加本书翻译工作的有柯海晰、符东浩、张定军、黄素芳、李明、李重华等，另外，龚亚平、王雷、贺民、樊鹏、吴国栋、周伟松等人也做了大量的工作，全书最后由贺军进行统稿。

本书如有错误和疏漏之处，请读者不吝指教。

潇湘工作室

2000年6月

1996 年，当我们在微软公司开始“柏拉图（Plato）”工程的时候，我们的目标是把 OLAP 带给大家。从一开始预算雇佣三位全职的雇员，到从 Panorama 公司获得技术和人员，再到以后的 1998 年将 Microsoft SQL Server OLAP Service 交付使用，就是这种简单的思想推动了我们的工程不断前进。

开始，我们只有一小组相信 OLAP 的人们。Corey Salka 和我先前为 IRI Software 公司（Express 的生产商）工作，目前是 Oracle 产品生产线的成员。我们一个公司接一个公司亲眼看到 OLAP 是如何体现其价值的。Corey 目前在 Excel 小组工作，致力于研究数据透视表的特性。Gunnar Mein 和 Goetz Graefe 是另外两个真正的信徒。这个很小的小组早在 1995 年秋天就开始了关于 OLAP 的讨论。

说服微软公司进行 OLAP 的研究则是不同的经历。我们在 1996 年 4 月提交了关于“柏拉图”的最初预算。那时，SQL Server 6.5 刚刚进行了测试版的发行。这个版本的特点是“立方体（cube）”算子和“上卷（rollup）”算子，微软公司的许多人都觉得，对于数据分析来说，这已经足够了。Corey 和 Gunnar 写了一份出色的为 OLAP 辩护的证明，我们把这份证明作为总体的中间件 FY97 预算请求的一部分提交了上去。最后，我们有了足够的预算，可以成立一个探索 OLAP 的队伍。

Corey 和 Gunnar 在制定创建 OLAP 工程的计划时，他们调查了大量的公司和技术部门，看看有没有可以直接拿过来用的东西。在访问某一家公司时，他们遇到了两个以后在很重要的方面影响了“柏拉图”的人：Tom Conlon 和 George Spofford，他们的贡献我将在稍后介绍。Tom Conlon 后来加入了微软公司，并接替了 Corey 在 Excel 小组中数据透视表程序管理员的位置。Tom 在“柏拉图”中的关键角色是作为向 Excel 中引进 OLAP 的发起人。Excel 在早期能够接受“柏拉图”，立即就增加了我们的努力的可信性。他们对于质量和过程的高标准要求帮助我们的队伍成熟起来。

在“柏拉图”早期的岁月里，具有创新意义的事情是发现了 Panorama Software 公司和获得了关键的人和技术。不幸的是，我错过了“柏拉图”历史上的这一整个阶段。在“柏拉图”一生的头一年半中，我要么处在先前公司（Oracle）的非竞争状态之下，要么全部投入到移植 Windows NT 4.0 Option Pack 的工作中。长话短说，来自 Panorama 公司的 Rony Ross 和 Amir Netz 正在努力与微软公司订立合同。他们遇到了 Microsoft Israel 小组，后者指引他们去找 Corey。几乎是在同时，OLAP 报道组的 Nigel Pendse 告诉 Tom Conlon 关于小公司 Israeli 的事，说他们有一个有创新意义的 OLAP 产品。Corey 和 Tom 把这两件事联系到一起，并在 Rony 和 Amir 到达美国时安排会见他们。

当 Corey 和 Gunnar 终于见到 Panorama 产品时，他们意识到，这就是他们一

直在寻找的东西。在不到 3 个月的时间里，他们以应有的勤奋确定了交易的框架，并完成了这宗交易。1997 年 1 月，Panorama 公司的 6 位职员来到华盛顿的 Redmond，成为微软公司的雇员。他们是 Amir、他的兄弟 Ariel、他的妻子 Merav、Mosha Pasumansky、Sasha Berger 和 Edward Melomed。Merav 加入了 Office 小组。其他人跟随 Stephen Bremmer 和 Corey，成为“柏拉图”小组的核心。所有的人现在仍然是“柏拉图”小组的成员。

1997 年，小组安顿下来，并开始设计和创建“柏拉图”。在那段日子里，我在大部分时间里都处在非竞争的状态之下，或者投入到移植工作中去，对这只队伍没有多加注意。本书的某些读者会记得，在那一年的 9 月由这个小组移植的 α 发行版。1998 年，小组推出了称为 $\beta 2$ 的发行版。我们称它为 $\beta 2$ ，是因为它是用 SQL Server 7.0 的第二个 β 发行版移植的。7 月发行了 $\beta 3$ ，并在 1998 年的 12 月开始生产最终的版本。我全天与这个小组一起工作并参与产品开发，也是从 1998 年的 1 月份开始的。

移植“柏拉图”需要花相当大的力气。在 1997 年的大部分和 1998 年的全年，工作时间为 80 个工作周，每周工作 100 个小时，这成了一种规范。我们有 20 多人的小组多次超出了公司的伙食预算。我们的口号是“没有吃的就没有 OLAP!”。除了工作时间之外，工作还是相当愉快的。一个有着明确任务的（大众化的 OLAP）、年轻富有活力的小组有着出色的领导、充足的伙食，能够完成任何事情。作为一名勤奋工作的老手，我很高兴地看到这个小组在自信心和能力方面都成长起来。

就在写这篇序言的时候，有人会问，我们到底做得有多好？我们是否真的把 OLAP 带给大众了呢？答案肯定是各种各样的。我确信我们已经做了大部分工作，使全部的 OLAP 产品把复杂的数据分析功能带给更多的大众。OLAP 管理器中的向导和与 Microsoft Office 的紧密集成，使得能够创建并使用 OLAP 数据库和应用程序的人比以前更多。“柏拉图”已经改变了 OLAP 的面貌。

这也就是说，我们有更多的工作要做。我们对于下一个版本有专门的计划，我们要进一步改善 OLAP 的易学、易用和易集成功能。但是，这只不过是故事的一半。实现 OLAP 的大众化很重要的一个因素是大多数的作者、顾问和培训人员，他们使整个市场总体上采用这种技术，特别是采用我们的产品。这使我认识到本书的重要性，以及作为 SQL Server OLAP Service 开发者之一的重要性。

Corey 和 Gunnar 在 1996 年遇到了 George Spofford。他们立即意识到，George 就是属于那种在任何层次上都了解 OLAP 的人。他们下决心邀请他参与“柏拉图”最初的以及以后全部的初步设计和复审。他总是毫无保留地提出对特性的建议，还帮助我们整理无数的选项以及添加、去掉合适的特性。George 接受并且测试了“柏拉图”产品的每一个发行版，并在工业界大多数人完全了解“柏拉图”之前发出故障报告。

他对于软件的评论使得小组在许多方面改进了这个产品。我们发现 George 总是实事求是，该怎样说就怎样说，好就是好，坏就是坏，这偶尔令我们沮丧，但我们必须这样，以免大家认为 George 是个“柏拉图”的“托儿”。对于广大读者来说幸运的是，他告诉大家症结所在以及如何对付它们。

Erik Thomsen 也因为 OLAP 和更广泛的数据库商务而被我们熟知。他的著作《OLAP Solutions》(Wiley 1997)是第一本影响广泛、易于接受的有关这一主题的著作。Erik 通过

文章、简报、演讲以及咨询活动，做了大量推动 OLAP 被工业界接受的工作。

对于每一种能够被大众接受的关键技术，有一本有创新意义的著作总是其成功的一部分。例如 Charles Petzold 关于 Windows 程序设计的著作。本书在 OLAP Services 方面非常有可能成为这种著作。我们对本产品及其文献倾注了极大的精力。George、Erik 和 Dick 会让你领略到超出我们所能讲的范围之外的东西，并帮助你找到使 OLAP Services 在你的商务中为你服务的方法。与那些简单重复产品文献的书不同，George 和他的伙伴们解释了幕后进行的一些东西。例如，创建维数时对于成员键值和成员名称的讨论。

我真诚地希望大家会喜欢这本书，并能从书中的建议和内部细节中获益。我们快速创建了 OLAP Services。George、Erik 和 Dick 看起来很喜欢在产品和编写本书中获得的经验。我们所有的人都希望大家在配置自己的 OLAP 解决方案时有同样的感受。

Bill Baker

产品部经理

Microsoft SQL Server 决策支持产品部

本书面向数据库管理员 (DBA) 或功能分析员 (power analyst), 我们将介绍如何在 Microsoft Server 7 OLAP Services (以后我们将其简称为 Microsoft OLAP Services 或 OLAP Services) 中创建高效的分析应用程序系统。我们的目标是通过一种实用的、彻底的方式介绍 OLAP Services 的特性和能力, 提供一种方法和指南去理解如何把 OLAP Services 应用到你自己的应用程序中去, 并提供了实际的例子, 说明如何借助于前人的工作实现你自己的任务。因为多维建模和分析的领域与事务数据库和应用程序的领域不同, 所以, 在本书中我们自始至终都会为你提供如何考虑建模和分析的指导。但是, 这并不意味着本书是一本有关 OLAP 技术的综合性指南。如果你对这种书感兴趣的话, 我们向你推荐 Erik Thomsen 所著的《OLAP Solutions》(Wiley, 1997)。我们假设大多数读者会利用微软公司的产品开发自己的第一个 OLAP 应用程序, 所以, 我们在适当的地方介绍了一些基本的 OLAP 概念。

OLAP Services 既是产品, 也是工具。在其他环境中, 例如 Java、C++、Delphi 和 Visual Basic (甚至还有 Perl) 中, 它可以生成用户自己的代码, 创建许多有趣的应用程序, 这突出了它作为工具的一面。某些基本的 OLAP Services 任务要求操作程序员级别的界面, 我们将在本书中描述这些操作。但是, 我们在此没有介绍 OLAP Services 服务器和客户程序设计的细节, 如果对此有兴趣的话, 可以参看其他书。

有一些主题尽管会很有用处, 但我们不能都包含在本书中。我们的范围只限于 OLAP 数据结构与处理, 主要集中在逻辑设计、物理设计和实现。实际上, 我们并不讨论有关软件开发的题目。OLAP Services 的基本管理界面实际上是按照 ActiveX 自动操作对象来定义的, 随本产品附带的 GUI 管理工具只显示了其中一部分有用的功能。基于 OLAP Services 的复杂的系统可以拥有在相当大的程度上由用户定制的软件, 即使是简单的系统也有可能包含由其他团体提供的软件。由于这个原因, 我们讨论了大多数应用程序需要的最基本的程序设计。

注意

为 OLAP 这样的新产品提供详细的指南可能会有些问题, 我们希望这不会影响你。正像所有的软件一样, Microsoft OLAP Services 有一些错误。正当我们准备这本书的时候, 第一个发行版已经生产出来, 而且, 它的服务包已经在使用中了。因此, 在你读到本书的时候, 我们提供的一些有关如何执行任务和短语查询的信息可能发生了改变。例如, 我们可能提供了如何组成一条查询语句的信息, 而你正在使用的版本已经提供了更为直接的途径。或者, 我们教给

你做某些事情的方法也许不能用，这是因为我们预期某种误操作会以某种方式得到补偿，而实际上并不是这样。

本书的组织

本书分为 4 部分。第一部分（第 1 章至第 4 章）介绍了 Microsoft OLAP Services 提供的逻辑 OLAP 结构、产品的物理和计算结构以及数据库设计和实现的基础。第二部分（第 5 章至第 9 章）描述了用于查询和计算的 MDX 语言，包括所有已有的 MDX 函数和运算符的参考说明，它们相当详细，而且带有图解说明。第三部分（第 10 章至第 14 章）描述了更高级的主题，例如维护、数据库优化和高级数据库设计。本书的最后一部分（第 15 章至第 16 章）是比较复杂的学习教程。最后，本书的 3 个附录提供了主要特性的参考文献，它们是查询（附录 A）、词汇表（附录 B）和随书配套光盘的参考资料（附录 C）。

配套光盘

本书配套光盘包含了学习教程中的源数据库、与 OLAP Services 一起使用的附带软件的两个部件，以及这个软件的 Visual Basic 源代码。这些程序填充了 OLAP Services 管理 GUI 的第一个发行版没有提到的功能部分，使你能够检验和操作重要的数据库组件。你还能够把这些程序作为代码范例，而这些代码还涉及到我们在本书中没有讨论到的领域。

格式约定

在本书中，表示代码的文字，例如查询或编程语言，将以如下字体表示：

```
This is some code
```

开始

本书对 Microsoft OLAP Services 的探索是很有趣的，甚至很好玩。OLAP 技术主要的好处是它能够简化复杂的分析。以前看起来很困难的或不可能实现的任务，现在看来是很有可能，甚至会是一目了然的。

第 1 章 OLAP 结构和 OLAP Services	1
1.1 常规 OLAP 结构	1
1.1.1 多维域结构图表	2
1.2 OLAP Services 维	3
1.2.1 概述：维的结构	3
1.2.2 OLAP Services 中的多维层次	5
1.2.3 All 级别和 All 成员	7
1.2.4 维中的成员	8
1.2.5 成员的性质（属性）	8
1.2.6 维类型	9
1.2.7 时间与常规维和级别	11
1.3 立方体和量度	12
1.3.1 常规立方体	13
1.3.2 虚拟立方体	15
1.3.3 立方体量度和级别属性之间的相似与差异	17
1.4 公式	18
1.5 结构与 RDBMS 大纲的链接	19
1.6 小结	21
第 2 章 Microsoft OLAP Services 结构	23
2.1 核心组件结构和平台	23
2.2 存储和查询处理结构	25
2.2.1 OLAP Services 存储聚合基础	26
2.2.2 计算结构	27
2.2.3 分区结构、存储和查询	28
2.2.4 定义、存储和处理	33
2.2.5 维和处理过程	35
2.2.6 立方体和处理过程	36
2.3 客户/服务器处理	37
2.4 小结	38

第 3 章 数据库设计基础	39
3.1 创建 OLAP Services 数据库的步骤	39
3.2 了解数据和目的	40
3.2.1 所需数据都已存在，并可通过 SUM、COUNT、MIN 或 MAX 函数进行聚合	41
3.2.2 维数据必须以合适的表的形式存在	42
3.2.3 需要的维数据必须存在	46
3.2.4 各维必须在所有聚合级别中都有成员	48
3.2.5 表中所有事件必须和维成员的同一级别相关	52
3.2.6 本节小结	55
3.3 设计 OLAP Services 维和立方体	55
3.3.1 维和立方体名称	55
3.3.2 共享与私有维	56
3.3.3 多重层次	56
3.3.4 成员属性	57
3.3.5 虚拟维	58
3.3.6 虚拟立方体	59
3.3.7 分区	61
3.4 小结	63
第 4 章 用 OLAP Manager 建立数据库的基本步骤	65
4.1 创建 OLAP Services 数据库	65
4.2 创建维：级别、成员和层次	66
4.2.1 创建标准维	69
4.2.2 编辑维属性	74
4.2.3 创建时间维	80
4.2.4 成员属性	82
4.2.5 处理维	82
4.2.6 虚拟维	83
4.3 创建立方体：综合起来	85
4.3.1 选择事实表、量度和维	85
4.3.2 改进立方体定义	87
4.3.3 创建计算成员和私有维	93
4.3.4 设计聚合和处理立方体	95
4.3.5 添加分区	97
4.4 小结	101

第 5 章 MDX 语言简介	103
5.1 基础知识	104
5.2 在 MDX 中命名元数据实体	106
5.2.1 定界名称和多部分名称	106
5.2.2 名称与唯一名称的比较	106
5.3 MDX 数据模型	108
5.3.1 数据模型：元组和集合	108
5.3.2 数据模型：查询	110
5.4 简单的 MDX 结构	111
5.4.1 , (逗号) 和: (冒号)	111
5.4.2 .Member 运算符	112
5.4.3 CrossJoin() 函数	113
5.4.4 Filter() 函数	115
5.4.5 Order() 函数	116
5.5 作为计算成员的维度计算	117
5.5.1 查询中的计算成员和 WITH 区域	118
5.5.2 公式首选项	120
5.6 命名集合	124
5.7 MDX 查询细节	126
5.7.1 轴的编号和排序	126
5.7.2 从查询轴中删除空数据片	127
5.7.3 查询成员的属性	128
5.7.4 WITH 节	130
5.7.5 查询单元属性	130
5.7.6 MDX 立方体分片表达式	131
5.7.7 复杂的分片表达式	132
5.7.8 分片表达式里的集合	133
5.8 有关 MDX 的严重不足	133
5.9 小结	134
第 6 章 MDX 计算的上下文和数据	135
6.1 查询中的单元上下文内容和解析顺序	135
6.1.1 集合函数里的单元上下文	138
6.1.2 无限递归：与计算上下文相关的常见问题	139
6.2 非数据：无效数值、空和无效成员	140
6.2.1 无效计算：除零和数值错误	141
6.2.2 空单元的语义	141

6.2.3 无效位置	143
6.3 计算成员单元属性的先后次序	145
6.3.1 计算成员的显示格式化优先权	145
6.3.2 计算单元的数据类型	146
6.4 小结	147
第 7 章 MDX 函数和运算符引用	149
7.1 数字计算	149
7.2 比较、逻辑和条件运算符	154
7.2.1 指定元组和集合	156
7.2.2 集合相关的基本运算符	159
7.2.3 建立在数据基础的集合筛选	164
7.2.4 集合排序	166
7.2.5 量度值	169
7.2.6 属性值	170
7.3 OLAP 元数据	171
7.4 字符串和命名函数	173
7.5 创建数组	174
7.6 MDX DDL 语句	175
7.7 小结	179
第 8 章 MDX 的维引用函数	181
8.1 成员的上下文无关集合和上下文相关集合	181
8.1.1 上下文无关的单独成员	184
8.1.2 相同级别	184
8.1.3 引用低级级别	190
8.1.4 更高级别	195
8.2 MDX 传播函数	195
8.3 小结	204
第 9 章 MDX 应用	205
9.1 常用的简单表达式	205
9.1.1 层次各个级别之间的比值	206
9.1.2 平均值	207
9.1.3 不同维上的不同聚合	209
9.1.4 在一个级别里移转聚合	210
9.1.5 填充空格	212
9.2 高级操作：查询和表达式	215

9.2.1	报表里各成员在不同级别之间的比值	216
9.2.2	在层次里跳过级别的层次排序	220
9.2.3	用集合里的约束等级包括所有元组	221
9.2.4	Pareto 分析以及累积求和	222
9.2.5	在不同位置对不同内容计算的公式	226
9.2.6	逻辑聚合 (For Any, For Each, For None)	229
9.3	样本分析	230
9.4	小结	238
第 10 章	维护应用程序	239
10.1	变化的特征	239
10.1.1	与数据源同步更新	239
10.1.2	增加新的数据	242
10.1.3	改变现有的数据	242
10.2	实现改变	244
10.2.1	维	244
10.2.2	立方体空间	246
10.3	管理改变的技术	251
10.3.1	更正处理	252
10.4	设计改变	255
10.5	小结	255
第 11 章	优化	259
11.1	数据源	259
11.1.1	索引	260
11.1.2	视图与表	260
11.2	维和立方体设计	261
11.2.1	表联接	261
11.2.2	键数据类型和大小	261
11.2.3	唯一的成员	262
11.2.4	立方体定义中的成员键列	263
11.2.5	级别聚合	263
11.2.6	计算成员	264
11.3	OLAP 存储选项	265
11.4	聚合	267
11.4.1	OLAP Services 如何决定聚合	267
11.4.2	优化聚合	271
11.5	查询处理	276

11.6 MDX 表达式	277
11.7 服务器设置	280
11.8 小结	281
第 12 章 高级立方体空间和数据库设计	283
12.1 行式项目维	283
12.1.1 常规维和量度维	284
12.1.2 表示行式项目	286
12.2 闭合回路: 将分析结果写回至 RDBMS	290
12.2.1 MDX 查询的 OLE DB 和列表结果	290
12.2.2 DTS 基本原理	293
12.2.3 查询和流示例	295
12.3 屏蔽虚拟立方体中量度的应用范围	297
12.4 小结	299
第 13 章 安全性	301
13.1 安全体系	301
13.2 安全性设计	302
13.2.1 隐藏量度	303
13.2.2 隐藏详细资料的级别	303
13.2.3 隐藏数据片	304
13.3 实现安全	306
13.4 小结	308
第 14 章 使用外部函数扩充 OLAP Services 系统	309
14.1 MDX 中的外部函数使用综述	309
14.2 参数和返回类型详述	311
14.3 用于外部函数的 MDX 函数	315
14.3.1 SetToStr() 和 TupleToStr()	316
14.3.2 Members(), StrToSet() 和 StrToTuple()	316
14.4 外部函数的其他注意事项	318
14.5 加载和使用函数库	318
14.5.1 USE LIBRARY 语法结构	318
14.5.2 加载函数库	319
14.5.3 澄清重复的函数名	319
14.6 小结	320
第 15 章 指南 1: 健康保险分析	321