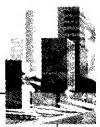
**3.8** **实施策略及风险控制**

客观地说，懂得安装数据库、建表、写存储过程，再会用展现工具开发报表基本就 算 BI入门了，但要实施企业级的BI项目，只满足于这几点是远远不够的。如果实施 策略不当，很可能造成投入巨资硬性推动后产生不了价值，或者东西很好并且有价 值，因为方法不得当最后还是得不到客户认可，甚至被否定。所以这里我们来讲一下 BI的实施策略，有代表性的恰恰是数据仓库支持OLAP 理念的两位创始人Bill Inmon 和 Ralph Kimball,Bill Inmon倡导自上而下实施策略(图3.22),而 Ralph Kim- ball 则提出不同看法，认为数据仓库应该实施自下而上的实施策略。但经过那么多 年的BI实施，对于一位有经验的BI实施人员我们认为可以采用第三种策略，那就是 整体规划、分步实施，这种方式可以融合两种策略的优点，又可以避免其中的短处。

**第** **三** **章** **商** **业** **智** **能** **在** **电** **子** **商** **务** **的** **解** **决** **方** **案**

41



**VS**

Ralph Kimball

Bil Inmon

自上而下

自下而上

优点

●投入小、周期短 ●见效快、风险小 ●业务理解高

缺点

●数据孤岛

●整体设计有挑战

优点

●数据仓库有整体设计

●数据一致性高

缺点

●投入大、周期长

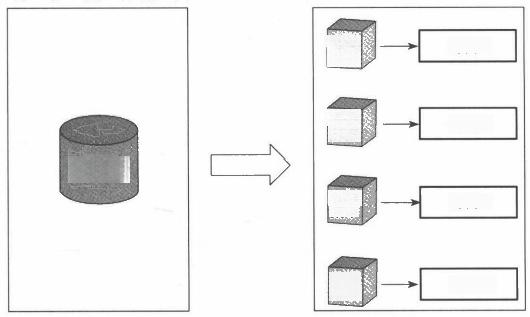
●业务理解低

●见效慢，风险大

**图3.22** **传统的两大** **BI实施策略：自上而下、自下而上**

**3.8.1** **自上而下策略**

自上而下实施策略(图3.23)最开始由 Bill Inmon 提出，其核心思想还是突出 前期的设计进而达到开发一步到位，实施的时候先统一规划业务，整体设计数据仓 库，在数据比较稳定后再面向各个部门及各个层级人员整体实施企业级 BI应用。



第一步：建立数据仓库 第二步：每个部门建立一个集市

数据

集市

数据

集市

数据金库

部门3

部门…

数据

集市

数 据

集市

部门1

部门2

**图3** **.** **23** **自** **上** **而** **下** **BI 实** **施** **策** **略**

这种策略优点是：

(1)数据仓库模型统一设计，结构简单易于管理及扩展，冗余数据少。

(2)虽然有数据集市，但是数据集市是基于数据仓库的基础上进行设计，数据 来源于数据集市，可以说是数据仓库派生出来面向特定主题的一个子集。

(3)数据仓库和数据集市在设计上属于两个不同层，只要数据仓库质量比较 高，增加主题应用会变得相对简单。



**商** **业** **智** **能** **在** **电** **子** **商** **务** **中** **的** **实** **践** **与** **应** **用**

其缺点是：

(1)初期就大规模实施数据仓库，投资大。

(2)需要较长时间才能使数据仓库开发到比较完善的程度，从而导致项目周 期长，见效慢。

(3)业务需求过于追求全面，导致项目难度大大提升，普遍情况是在 BI 实施 之前大部分用户是很难想象 BI 有什么功能，能对他们有什么帮助，更难提出合理 意见，这种时候唯一的办法就是引入高端咨询人员进行需求规划及设计。可想而 知对于塑造差异化竞争力的今天来说，企业采用自上而下的实施策略存在巨大的 风险。

从上面我们知道，自上而下的实施策略要发挥其优点避免风险必须具备如下 条件：

(1)有足够的财力。

(2)直接打造企业级数据仓库，需要高层领导支持，解决跨部门协调资源的 困难。

(3)业务形态较为常见，有强大的成熟的业务咨询团队和实施团队，有能力化 解整体设计及开发过程中的任何业务及技术问题。

(4)业务系统较为统一，数据仓库数据质量有足够保证。

42

另外 Bill Inmon 提出自上而下的实施策略同时要求数据集市能够上下钻取， 这种功能刚好数据立方体(Cube) 工具能够实现，最著名的有 Hyperion 公司的 Essbase, 还有就是 Cognos 的 Powerplay, 这些数据立方体工具笔者估计一开始应 该是为了实现Planning 应用而开发的，我们知道 Hyperion 和 Cognos 的 Planning 软件都是很知名的，Planning 软件在时间和组织架构上非常需要往上汇总、往下分 摊，后面这种上下钻取的功能因为刚好满足BI 的基本需求，过去十年被 BI 厂商广 泛采纳和推广，并取得了不错成绩；但是从解决方案和工具来看，现在技术和工具 只能说广义上的Bill Inmon的上下钻取，现在上下钻取在技术和架构上有如下三 大流派：

(1)Cube 数据集市支撑的物理 OLAP。

(2)关系型数据库加前端缓存支撑的逻辑型 OLAP。

(3)关系型数据库加内存支撑的逻辑型OLAP。

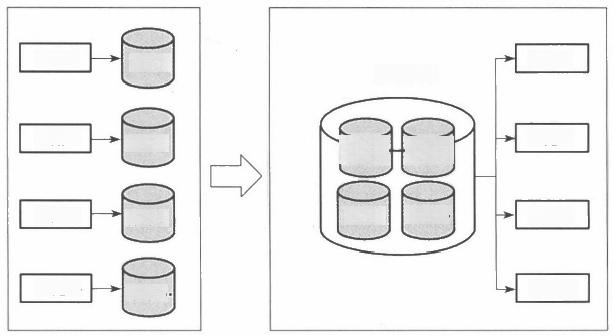
从上面的划分来看，当前数据模型设计和前端展现工具之间、软件和硬件之间 互相支撑，这也是一种趋势，特别是现在云计算的成熟也必将促使BI 在未来3～5 年里其技术及工具方面发生巨大变化；从这个层面笔者再次提醒我们开发者要真 正理解 BI的本质和问题，了解主流的工具和技术，只有正确认识前者，合理利用后 者才能做好企业级 BI实施，特别是具有大数据的电子商务行业更是如此。

**第** **三** **章** **商** **业** **智** **能** **在** **电** **子** **商** **务** **的** **解** **决** **方** **案**

43

**3.8.2** **自下而上策略**

**虽然都是数据仓库理念的创始人，但不同于** **Bill Inmon,Ralph Kimball 提倡** 自**下而上的数据仓库实施策略(图3.24),Ralph Kimball 认为“不同数据集市合起** **来就是我们要的数据仓库”,这种策略的优缺点也同样明显。**



第 一步：按主题建立集市 第二步：分主题集市合并为数据仓库

主题1

数据仓库

主题2

集市2+

主题3

主题…

集市

集市3

主题…

集市…

主题1

集市2

集市3

集市1

主题2

主题3

集市…

**图3.24** **自下而上** **BI 实施策略**

这种策略优点是：

(1)直接设计数据集市进行OLAP 分析，让BI 快速产生价值，项目见效快。

(2)不断叠加的数据仓库建设策略能保证数据仓库设计更接近实际需要，避 免仓库特别是模型的不合理设计，从而降低风险。

(3)另外在自下而上的策略里面Kimball提到维度建模的概念，从而产生了当 前最为流行的数据仓库按照维度和事实的设计，特别是Kimball 对维度建模的若 干建议对数据仓库设计具有非常高的指导价值。

其缺点是：

虽然前期集市设计速度快、风险小，但数据集市设计更多是在部门级或者比较 单一的业务领域，这种设计难于推广到其他业务部门和领域，所以前期只注重数据 集市设计是难于形成企业级数据仓库，具体来说有以下三个问题难于统一：

(1)数据粒度：数据仓库一般尽量保持必要的明细，而数据集市则是看当前业 务需要。

(2)数据范围：数据集市只选择当前需要的数据，哪怕是同一个表在不用应用

的时候也会在不同程度上进行过滤。

(3)数据定义：这个是问题最大的，业务部门的指标选择和定义往往有差别，

 **商** **业** **智** **能** **在** **电** **子** **商** **务** **中** **的** **实** **践** **与** **应** **用**

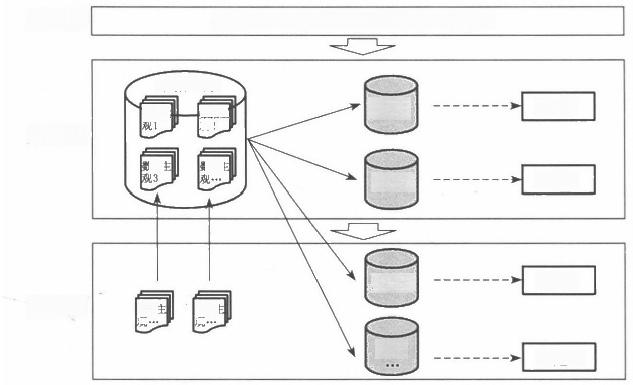
所以不同业务部门的需求在集市设计上必然会有区别。

这里不得不说的是 Bill Inmon 和 Ralph Kimball 提出的策略都没有绝对意义 上的自上而下或者自下而上，但在实际工作中，BI 项目组、IT 领导往往都喜欢绝对 的自上而下，能够获取到公司最大的资源支持，结果也容易导致投入大，但是见效 小而慢，甚至失败；而业务部门在久久看不到效果的时候又总是喜欢绝对的自下而 上，每个部门自成一派，最终企业级 BI 迟迟建立不起来。为此我们还是要从本意 上理解两种策略，特别是吸取其长处。

从上面我们知道自上而下、自下而上两种策略都有可取之处，也各有致命弊 端，显然实际开发及应用过程中，谁都不会简单使用其中一种，都会两者兼备，这里 最为提倡的是整体规划，逐步实施，下面进行详细介绍。

**3.8.3** **整体规划、逐步实施**

简单概述不难发现 Bill Inmon 重视数据仓库的整体设计，从而保证系统的完 整性，而 Ralph Kimball 重视快速推进，快速出效果，重视项目的实用性。应该说两 者意图都值得肯定，所以我们提倡两者兼备的整体规划、逐步实施，如图3.25 所示。



数据分解(把数据分解成相对独立的主题)

数据仓库

主题1

数据主

第二步

主题2

集市3 主题3

数据主

观

主题…

第三步

第一步

集市2

集市1

数据主

观2

集市

数据主 观 ·

数据主

据主

**图3.25** **整体规划、分步实施**

这里我们建议把 BI 实施分成三步走。

第一步：把具有相互独立性的业务数据分成不同的主题，特别是要明确核心业 务主题及维度信息表。

第二步：选定1～2个分析主题，对核心业务进行数据仓库设计并且实现数据 入库；维表设计和开发尽可能一步到位，满足主题分析的集市数据要求； 一个分析

**第** **三** **章** **商** **业** **智** **能** **在** **电** **子** **商** **务** **的** **解** **决** **方** **案**

主题一个集市。

第三步：根据用户使用情况及规划可以不断增加分析主题，扩展主题集市，如 果数据仓库数据不够，那么需要不断补入数据， 一般尽可能同一主题数据一起 补入。

分步实施在推进整体规划上需要遵循下面的原则。

(1)基本原则：分步实施虽然要满足整体的规划，体现其整体性，但更要重视 每一期项目的实效性，项目实施过程中可以根据实际情况对规划及计划进行合理 调整。

(2)制订实施计划：基于整体规划和主题划分，对业务主题进行重要性分级及 差距分析，从而形成分步实施计划。

(3)基本任务：针对不同的业务需求建立独立的集市，以满足不同的数据分析 需求及性能需求；对数据仓库基本数据进行逐步完善，重视模型的完善和数据质量 管理。

(4)初期目标：BI 初期开发偏向于系统的基础性建设及快速出效果。在初期 一般都会把数据仓库建设到70%左右，完成公共维度和基本事实数据的入库；BI 实施效果得到体现，可以让企业对于 BI价值有直观认识，便于后面工作开展。

(5)中远期目标：中远期重视数据质量管理和 BI 整体应用体系。数据质量问 题对 BI应用是致命的，长期来讲需要建立起数据质量管理体系；还要成体系的推 进 BI应用，让BI 支持的分析点互相支撑，形成统一的企业级决策分析平台。

45

优点：

(1)整体规划、分步实施很好地融合了自上而下及自下而上BI 建设思路的各 自优点。

(2)注重数据仓库的整体设计，70%左右的前期设计可以建立起数据仓库的 基本框架。

(3)分步实施的业务应用可以让 BI不断得到回报，从而让企业更有信心。

(4)同步推进的业务应用可以不断检验数据仓库模型的正确性，从而避免在 仓库模型建设上出现大的反复。

缺点及应对措施：

这种策略的不足就是怎么来保证项目的整体规划能力和分步实施效果。

(1)业务规划能力：BI 应用规划不得不面临不同主题间的横向区别，同一主题 业务对于不同级别用户的需求也是有区别的，这就要求 BI 规划人员既要懂业务又 要懂BI 应用，且要有较高的 BI行业咨询能力。

(2)技术方案能力：因为数据仓库的完整性，BI 分步实施是对技术方案的一个 挑战，如果没有领先的技术方案是难以保证项目的独立性和设计的整体性，所以团



**商** **业** **智** **能** **在** **电** **子** **商** **务** **中** **的** **实** **践** **与** **应** **用**

队必须要有领先的 BI 技术方案。

(3)项目实施能力：分步实施重视不断出效果，如果前期效果连续失败，那整体

规划还是一句空谈，因此保证项目实施效果，特别是初期效果尤其关键，所以要求 项目组有较高的项目实施能力；另外可以选择对报表分析需求较大的核心业务部 门开始，并获得业务部门的支持，这样成功可能性要大。