

数据仓库设计的21条原则计算机等级考试 PDF转换可能丢失
图片或格式，建议阅读原文

https://www.100test.com/kao_ti2020/644/2021_2022__E6_95_B0_E6_8D_AE_E4_BB_93_E5_c98_644624.htm

高效实现数据仓库的七个步骤 数据仓库和我们常见的RDBMS系统有些亲缘关系，但它又有所不同。如果你没有实施过数据仓库，那么从设定目标到给出设计，从创建数据结构到编写数据分析程序，再到面对挑剔的用户的评估，整个过程都会带给你一种与以往的项目完全不同的体验。一句话，如果你试图以旧有的方式创建数据仓库，那你所面对的不是预算超支就是所建立的数据仓库无法良好运作。在处理一个数据仓库项目时需要注意的问题很多，但同时也有很多有建设性的参考可以帮助你更顺利的完成任务。开放思维，不断尝试新的途径，对于找到一种可行的数据仓库实现方法来说也是必需的。

1. 配备一个全职的项目经理或你自己全面负责项目管理 在通常情况下，项目经理都会同时负责多个项目的实施。这么做完全是出于资金和IT资源方面的考虑。但是对于数据仓库项目的管理，绝对不能出现一人身兼数个项目的情况。由于你所处的领域是你和你的团队之前没有进入过的领域，有关数据仓库的一切 - 数据分析、设计、编程、测试、修改、维护 - 全都是崭新的，因此你或者你指派的项目经理如果能全心投入，对于项目的成功会有很大帮助。
2. 将项目管理职责推给别的项目经理 由于数据仓库实现过程实在是太困难了，为了避免自虐，你可以在当前阶段的项目完成后就将项目管理职责推给别的项目经理。当然，这个新的项目经理一定要复合第一条所说的具有全职性。为什么要这么做呢？首先，从项目经理的

角度看，数据仓库实施过程的任何一个阶段都足以让人身心疲惫。从物理存储设备的开发到Extract-Transform-Load的实现，从设计开发模型到OLAP，所有阶段都明显的比以前接触的项目更加困难。每个阶段不但需要新的处理方法、新的管理方法，还需要创新性的观点。所以将管理职责推给别的项目经理不但不会对项目有损害，还可以起到帮助作用。

3. 与用户进行沟通 这里所讲的内容远比一篇文章本身要重要的多。你必须明白，在数据仓库的设计阶段，那些潜在用户自己也不清楚他们到底需要数据仓库为他们做什么。他们在不断的探索和发现自己的需求，而你的开发团队也在和客户的接触中做着同样的事情。更加频繁的与客户接触，多做记录，并让你的团队更关注于项目需求讨论的结果而不是讨论的过程本身。既然你和客户的交流是为了了解存储的数据是何种类型以及如何有效存储数据，你也许需要（和你的用户一起）采用一种新的方法观察数据，而不是直接处理数据。你可以尝试从中找出隐藏的信息，比如在一段时期内的数字涨落等。不要试图追寻项目需求的答案，而是要让答案找上门来。

4. 以技术/信息库作为领导 由于数据仓库实施的各个阶段都有很大不同，因此你需要有人能起到维持整个项目的连续进行的作用，不过这个职责并不需要那种全职性。项目实施有三个重要方面：架构、技术和业务。将架构作为重点可以保证在整个项目中，数据仓库的架构从物理层往上，都会受到良好的维护。而我们应该将技术作为重点，因为开发团队和关键用户都在使用他们以前从未用过的工具，必须有人监督开发过程以及工具使用的一致性。最后，在数据仓库的应用过程中浮现出来的业务需求必须被详细分析和记录，以促

机开发过程持续下去。如果用户不能很好的开发人员以及其它用户沟通，那么数据分析和度量方面的开发进程就会延期，所以必须有人关注业务方面的开发，推动开发进入更高级别。

5. 跳出反复修改程序的陷阱 第一次实现的数据仓库肯定不会是最终交付的版本。为什么呢？实际上在真正见到产品前，你无法确定的知道自己的目标是什么。或者说，最终用户只有在使用数据仓库产品一段时间后，才能明确告诉你这个产品是不是他所希望的。与你以往处理的项目不同，业务智能还处于发展的初期，每个公司对业务智能都有不同的解释，因此你的项目决不会一次成功。为了以正确的格式获得数据，你需要在不断变化的状况中摸索前进。BI具有很强的个性，不同的环境、不同的市场以及不同的企业都有不同的BI。这又代表什么呢？这表示你需要把数据库管理员放在一个消息相对封闭的环境中，不要让他知道数据仓库的数据结构以及ETL程序在不断的改变。对此没有别的办法。这样可以减轻你和DBA所承受的压力。

6. 对大量的前端资源进行数据源分析 在数据仓库实现过程中，你不得不在旧有的数据中艰难跋涉，这些数据来自老的数据库、老的磁带机以及远程的数据。它们中的大部分都凌乱不堪，并且难以获取。你要对这些数据进行大量处理，并且还要设计ETL程序来寻找其中的有用信息。如果你希望整个项目做起来比较顺利，并且找到一种方法能够一次成功，那就需要你的开发人员必须花费足够的时间来充分研究这些旧有数据，将凌乱的数据规则化，并尽力设计和实现强壮的数据采集和转换过程。数据仓库的ETL部分会占用整个项目资源的百分之八十，所以一定要确定你的资源都用在刀刃上了。

7. 将人际关系处理放在

首位 在数据仓库实现过程中真正的地狱不是来自技术或者开发方面，而是来自你周围的人。你也许会遇到一个对项目并不乐观而又没时间听你陈述的领导。你也许会遇到一些开发人员将进度拖延太长时间还抱怨为什么不能用老方法实施。你也许还会遇到一些抱有不切实际的幻想的用户，他们希望轻点鼠标就能实现想象中的功能，却不愿在他们那边多做些智力投资，更好的培训他们自己的员工。而你也已经疲惫不堪，鼓励投资，以及在开发团队和用户（甚至老板）中推广新的开发技巧。总之你要保持微笑。当一切搞定，你的烦恼也就一扫而空了，笑到最后才笑得最轻松。 100Test 下载频道开通，各类考试题目直接下载。详细请访问 www.100test.com