

计算机三级:动态数据仓库设计与应用浅谈计算机等级考试

PDF转换可能丢失图片或格式，建议阅读原文

https://www.100test.com/kao_ti2020/581/2021_2022__E8_AE_A1_

[E7_AE_97_E6_9C_BA_E4_c98_581792.htm](https://www.100test.com/kao_ti2020/581/2021_2022__E8_AE_A1_E7_AE_97_E6_9C_BA_E4_c98_581792.htm) 数据仓库技术的每次演进都以发掘企业数据中更多价值作为目标。而近期流行的动态数据仓库技术，不仅在灵活性、可视化方面有了长足进步，还能够对企业决策、合作伙伴及客户服务提供更为强大的支持。 数据仓库发展历程 数据仓库的发展历史具体可以划分为五个阶段。数据仓库系统发展之初，其主要作用是为企业内部的某些部门提供一些固定的报表。因此这一阶段通常被成为“报表”阶段。在该阶段，数据仓库的结构可以根据具体问题进行优化，即使数据查询人员要求访问的信息量极其巨大，处理这些资料的效率仍然可以很高。当企业用户的关注点从“发生了什么”转向“为什么会发生”，数据仓库进入了“分析”阶段。在这一阶段，决策者开始对数据进行分析，实质上是在了解报表数据的真实涵义。这就需要更详细地对数据进行多角度分析。为了解决数据查询的瓶颈，出现了联机分析处理（OLAP）环境。它可以使对数据进行多角度分析的反应时间以秒或分钟来计算。因为在OLAP环境中，很多聚合数据都是预先计算好的，而且数据的存储裕揭埠痛车墓叵敌褪菘个肪炒嬖谄局是稹?/p>. 拥有了量化的数据支持后，企业对经营的动态情况以及这种情况为什么发生都会有所体验，接下来就要将业务信息用于预测了。数据仓库也随之进入“预测”阶段，即数据挖掘阶段。数据挖掘能够预知企业即将发生的动向，帮助管理者更为积极地管理和实施企业战略。数据挖掘为用户提供丰富的数据采集工具，以

便利用历史数据创建预测模型。数据仓库演进的第4阶段即是动态数据仓库。第1到第3阶段的数据仓库技术都以支持企业内部战略性决策为重点。而第4阶段则侧重在“战术性”的决策支持，为“执行企业战略的员工”提供支持。我们将这一阶段称为“营运导向”阶段。动态数据仓库技术在企业环境成熟应用后，将引领企业“动态性”阶段。伴随着动态数据仓库在决策支持领域所扮演的角色越来越重要，企业实现决策自动化的积极性也在不断提高。在人工操作效果不明显时，为了寻求决策的有效性和连续性，企业会趋向于采取自动决策方式。数据仓库“动”起来“动态数据仓库”是一种创新理念，但其技术基础和架构思想还是来自传统数据仓库技术。关键的区别是动态数据仓库增加了“动态”特性，与传统数据仓库相比，它具有如下特点：1. 动态访问 动态访问是指一线用户可以动态、或实时地访问所需要的信息。传统的数据仓库用户只针对高端管理层，而如果要实现大量客户经理和客户代表同时访问，是一个很大的压力。动态数据仓库采用不同于传统数据仓库的技术手段，扩展了数据仓库系统的用户范围，实现动态访问。可运担荃挚馊靡幌咳惫心髯?ldquo.动了起来”。2. 动态数据加载 传统数据仓库保存的是历史的、相对静止的、集成的企业数据。其往往是先加载好数据，再去支撑业务查询。而动态数据仓库的数据加载却可以在加载数据的同时，满足用户的查询请求，而且动态加载的负荷不影响用户使用数据仓库。不仅如此，动态数据仓库的数据也是准实时加载的，这样就可以使用户能够访问几乎和生产环境时效相当的数据。3. 动态事件 传统数据仓库只是支持用户对企业的历史数据的分析，或者经过一些模型

对未来的一些发展进行预测。它无法支持一线员工在遇到一些“动态事件”（例如银行柜台向客户推荐理财产品）时，进行一些实时的业务操作。而动态数据仓库就可以支持一线员工在遇到这些“动态事件”时，及时做出响应，成功抓住业务机会，从而大幅提升业绩。因为，从事情发生到采取行动的时间越短，成功销售的命中率就越高，所获的价值也就越高。

4. 动态负载管理 传统数据仓库在负载管理方面，没有什么特殊要求。而动态数据仓库则包含策略动态负载管理和操作动态负载管理。其中，操作动态负载管理一般是比较简单的访问，不需要看太多的信息；而策略动态负载管理则需要实施复杂的数据挖掘。

5. 动态企业集成 传统数据仓库主要用来支持企业用户对数据进行分析。而动态数据仓库则可以将企业所有系统都很好整合在一起，形成一个闭环，从而实现流程的自动化，而不是一个独立的系统。

6. 动态可用性 因为动态数据库已经不是一个纯粹的后台系统，而是业务运营的一部分，因此对可靠性、稳定性的要求更为严格。

动态数据仓库设计 前面提到，“动态数据仓库”在技术设施上采用得都是已有技术，而实现的难点在于“动态”，即如何实现数据仓库的“动态”特性，这也是部署“动态数据仓库”系统的关键所在。动态数据仓库中包含了传统数据仓库的各种元素，例如元数据管理、数据分发、对外服务、调度管理、代码自动化、数据质量管理等。要实现动态数据仓库中的动态数据加载，有多个实施方案供用户选择。目前，市场中存在多个接近实时的数据同步解决方案。例如，在专有工具方面，可以借助IBM WRS SQL复制和IBM WRS基于Q的复制实现；ETL工具包含Informatica PowerExchange、IBM

DataStage加CDC组件；数据库工具包括，甲骨文数据库的复制技术、DB2基于CD/CCD的SQL复制，以及SQL Server的出版社订阅复制技术；另外，数据捕获器与消息中间件的集成可以提供动态数据仓库的客户化集成解决方案。上述方案基本上是基于现有的数据库复制技术进行。除此之外，还有一些专业的数据同步软件，例如GoldenGate就是进行数据实时同步的一个非常好的工具。有报道称，美国领先卫星电视服务提供商DIRECTV就是采用GoldenGate和Teradata来部署动态数据仓库，并获得成功的。动态数据仓库的应用现阶段，动态数据仓库的实践者主要为速递服务公司和金融保险类企业。软件解决方案方面，NCR teradata的ADW（Active Data Warehouse）、Sybase IQ能够为企业的动态数据仓库应用提供比较好的支持。其中，Sybase IQ的特色在于采用了按列存储的创新技术和专有的数据压缩技术。近期速递服务行业纷纷开始构建邮件速递的实时动态查询系统。作为这一系统的主要支撑技术，动态数据仓库将帮助速递服务企业构建覆盖业务分析、时限控制、财务结算等内容的业务应用平台。日前，国内某速递服务公司就以Teradata平台作为其特快专递系统的硬件平台，并采用Teradata ADW解决方案搭建起其第三代速递跟踪查询系统。系统上线后，该公司的速递邮件实时动态跟踪查询系统每天处理超过千万笔信息，全天邮件查询量从上线之初的几十万件次增加到了上百万件次，高峰访问时段可达到每小时十几万件次。借助动态数据仓库，该公司速递邮件的收寄、投递信息在处理完成几分钟内即可发送上网，而对于采用无线手持终端方式上传的邮件状态，可以在几秒钟内即体现出邮件的最新状态，大大提高了查询的时效性。

，更好地满足了用户的需求。而在金融行业，现阶段的动态数据仓库应用主要是主动预防和管控某些业务风险，而不是等到风险发生后再报告。该过程将原来的被动管理转化为现在的主动防范。而除了风险控管，企业还可以通过对数据的分析和整理来判定自己的交易行为或客户的交易行为是否符合法律法规的要求。特别推荐：2009年9月全国计算机等级考试时间及科目预告 2009年上半年全国计算机等级考试参考答案请进入计算机考试论坛 2009年全国计算机等级考试报名信息汇总 2009年NCRE考试有新变化 2009年全国计算机等级考试大纲 2009年上半年全国计算机二级考试试题及答案 2009年上半年全国计算机等级考试试题答案汇总 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