

DB2V9pureXML在企业中的典型应用 PDF转换可能丢失图片或格式，建议阅读原文

https://www.100test.com/kao_ti2020/253/2021_2022_DB2V9pureX_c97_253038.htm 我相信，XML 对于多数读者来说已经是一个耳熟能详的术语了。由于XML具有自我描述性(易理解性)、灵活性、平台无关性和标准化等优点，因此，许多 IT 精英和组织都在探索如何将 XML 这一“基因”优秀的技术广泛应用到企业各个领域。要让 XML 在企业级层面广泛应用的第一个问题就是技术规范和标准。1998 年 2 月，W3C 正式发布了有关 XML 的规范，并一直不断地完善其标准。到目前为止，XML 的技术规范和标准方面应该说是相当成熟了。然而，XML 广泛应用于企业应用还有一些一直困扰着 IT 人的基础问题：如何有效地存储 XML?如何更好地管理 XML?如何提高 XML 的信息处理效率，包括查询、部分更新等等?这些基础问题不解决，XML 在企业应用中就“无法落地”，或者至少说是“无法令人满意地进行落地管理”。例如，XML 在前些年已经广泛应用于数据交换领域，从此，IT 系统与系统之间就有了“公共语言”。然而，人们一直没有找到一种方式，将这些“公共语言”进行高效率地原生态存储和管理，就像原始人有了语言以后长期没有找到能够原生态记录和管理语言的方式“文字”。人类是充满智慧的，IT 人更是如此。原始人在没有发明文字的时候，采用了“结绳记事”等等简陋的或是间接的方法来记录和管理人与人之间交流的语言。而 IT 人在过去几年里，也尝试用了一些简陋的或是间接的方式来存储和管理 XML，简单归纳如下：传统方式1：将 XML 简单地保存在文件系统中。这种方式可以说是非常简陋的，基

基本上没有管理性和效率可言。XML 文件数量不多的情况还勉强可以接受，到了一定规模的企业应用就是难以忍受了。我亲眼目睹了一家公司的应用将数千个 XML 文档放在一个文件夹中，然后通过 Java 应用进行信息查询、更新和删除。其效率、应用维护的复杂度、权限管理等方面几乎可以用“惨不忍睹”来形容。

传统方式2：将 XML 存入关系数据库(DBMS)的大对象字段中。实际上，这种方式只不过是“传统方式1”中 XML 存储的位置从简单的文件系统挪到数据库大字段中而已，并没有实质性地提高 XML 的可管理性和可操作性。而且不仅如此，我们知道，管理大对象并非 DBMS 的强项。例如在多数 DBMS 中，大对象是不能够像一般结构化数据那样进入数据库内存缓冲池的，因此，效率上大打折扣。

传统方式3：将 XML 分解到关系数据库的多张关系型表中。如果 XML 文档本身比较简单还好，如果稍微复杂一点，则可能一个 XML 文档就要映射到好几张关系表，而如果不幸的，XML 的格式经常要发生一些变化的话，那这个映射关系维护起来就更复杂了。这是问题之一。其二，这种方式会让 XML 文档本身的完整性荡然无存，对于一些需要保证 XML 文档完整性的应用就非常不合适。其三，这种分解的方式会耗费数据库服务器的大量 CPU 和内存资源，有可能使数据库系统面临系统资源不足的风险。

传统方式4：将 XML 存放到 XML-Only 的数据库中。这种特殊的 XML-Only 的数据库大大提升了 XML 的可管理性和可操作性。然而将这类非主流的、尚未经过业界验证的数据库应用到企业系统中，未免让人心惊胆战。其次，这种 XML-Only 的数据库基本上不支持关系型数据，而二十几年的关系型数据历史让关系数据库已经渗

入到企业应用的各个方面，因此这种XML-Only数据库在保护投资方面也是倍受批评的。DB2 V9 的推出开辟了 XML 存储和管理的新天地。在存储方面，DB2 V9 将 XML 进行高效地原生态的存储，在应用访问方面，支持 SQL 和 XML 两种访问方式。不仅如此，所有关系型的操作方式和数据库工具都可用于 XML 之上，例如索引机制、导入导出、关联查询、高速批量加载、系统优化等等。IBM 将这一技术称之为“pureXML”。有了 pureXML 技术，XML 在关系数据库中再也不是“二等公民”了。DB2 V9 成为一个既支持关系数据，又原生态支持 XML 的名副其实的双引擎数据库。图 1. DB2 V9 中的“双引擎”处理方式

那么，在应用了 DB2 V9 pureXML 技术，排除了 XML 应用的基础性技术障碍之后，我们可以用 XML 来为企业应用做些什么事情？事实上，我无法，也无意将 pureXML 在企业中的应用价值进行穷举。下文只是将我到目前为止所能够看到应用领域做一些小结。真正的目的在于抛砖引玉，引发读者对 DB2 pureXML 的价值进行不断探索。

应用领域之一：信息交换和共享 XML 最先被应用的企业领域是信息交换和信息共享。由于普遍意识到 XML 的优点，许多行业开始制定本行业的基于 XML 的数据交换和信息共享标准。例如：ACORD(保险业 XML 标准)、FIXML(基于 XML 的金融信息交换协议)、FPML(金融产品 XML)、HL7(医疗卫生 XML 标准)、IXRetail(零售行业 XML 标准)、XBRL(业务报告和会计 XML)、NewsML(新闻和发行 XML).....当然，信息交换和共享并非因为 XML 才出现的，然而有了 XML，企业与企业之间或者企业内部的 IT 系统之间的信息交换和共享就更具标准化，同时具有可理解性和灵活性的特点，具有

了“共同语言”。以 FIXML 为例，老的 FIX 标准是基于简单文本的，几乎没有可理解性和灵活性可言。而新的 FIXML 标准由于采用了 XML，具有很好的可理解性和灵活性，如图 2 所示：图 2. FIX 与 FIXML 对比

现在，有了 DB2 pureXML 技术，这些数据交换平台的能力大大增强了，这些交换的 XML 可以被 DB2 数据库原生态地管理起来。应用可以非常便捷地将 XML 以灵活的格式送入“信息高速公路”（企业总线），或者从企业总线中获得 XML。

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