**Lotus文档数据库概述**

美国Lotus公司（已被[IBM](http://www.csai.cn/incsearch/search.asp?key=IBM)公司兼并）于八十年代后期推出的群组工作软件（群件）——Lotus Notes/Domino是一种基于Internet/Intranet技术为构架的群件系统，运行模式不但有客户机／[服务器](http://www.csai.cn/incsearch/search.asp?key=%B7%FE%CE%F1%C6%F7)（Client/Server），而且有浏览器／服务器（Browser/Server），是构造企业信息网主要工具之一。现已被广泛应用于办公自动化、工作流控制、协同工作环境等方面，它提供了功能强大的多媒体共享文档数据库和电子邮件系统，是工业界的群件[标准](http://www.csai.cn/incsearch/search.asp?key=%B1%EA%D7%BC)。

　　在SQL Server、DB2、Sybase、[Oracle](http://www.csai.cn/incsearch/search.asp?key=Oracle)等大型关系数据库得到普遍应用的大环境下，当前我国的院校学历[教育](http://www.csai.cn/incsearch/search.asp?key=%BD%CC%D3%FD)课程中，有关数据库系统的教学内容主要集中在关系数据库上，关于文档数据库的知识介绍较少。尤其是Notes文档数据库在我国大、中型行政和企、事业单位应用日趋普及的情况下，对于每一个转向Notes的开发者来说，必须有比较深刻的认识和理解。本文根据在实际开发中的经验体会，结合对Notes应用中的需要来对文档[数据库原理](http://www.csai.cn/incsearch/search.asp?key=%CA%FD%BE%DD%BF%E2%D4%AD%C0%ED)和模式结构进行比较分析，加以讨论。

**一．关于Notes数据库与关系数据库的概念比较：**

　　我们知道在数据库理论中，有外模式、模式和内模式之分。外模式（又称子模式或用户模式）是数据库用户使用的局部数据的逻辑结构和特征的描述，是数据库用户的数据视图，如视图和表的结构、类型、长度、保密级别等。模式是数据库中全体数据的逻辑结构和特征的描述，是所有用户的公共数据视图，如不同表的结构、类型、长度、保密级别等，以及表和表之间的各种逻辑关系。内模式（又称物理模式）是数据物理结构和[存储](http://www.csai.cn/incsearch/search.asp?key=%B4%E6%B4%A2)方式的描述，是数据在数据库内部的表示方式，如存储、索引方式以及数据压缩、加密和聚簇等。为了对Notes文档数据库和关系数据库进行比较，下面让我们来了解一下Notes文档数据库的相关概念。

　　1．Notes 数据库文件 (.NSF)：数据库文件包容应用程序的数据，其结构由表单、域、文件夹、视图及其他显示功能部件（诸如导航器和数据库图标）组成。具体分两个部分：设计元素、数据文档。设计元素就是表单（Form）、域（Field)、视图（View）和文件夹（Folder）等等，它们构成了应用的基础。数据文档（Document）是用户提交的数据信息。形象来说，它就是用户保存在Form中的文档，例如：一份公文。

　　2．表单（Form）：NSF数据库文件中，一个NSF数据库中可定义多个表单，表单由多个表单元素组成。表单元素是用于创建表单外观和功能的组件。和页面一样，表单可以包含：文本、水平基准线、表格、区段、图形、图像映射、链接、自动功能、小程序、附件、嵌入元素和[HTML](http://www.csai.cn/incsearch/search.asp?key=HTML)。表单可由多个子表单组成，子表单是表单元素集合，它作为一个单独的对象来保存。子表单可以包含与常规表单相同的元素。使用子表单可以节省重新设计的时间。如果更改了子表单的某个域，则用到该子表单的每个表单都会更新。子表单的一般用法包括：向商务文档中添加公司标志，或向邮件和便笺表单中添加邮寄标签信息。子表单可以作为表单的永久部件，或者可以根据公式的结果有条件地显示。例如：可以给用户提供一种选择，允许用户使用不同的图像和样式来为各种类型的邮件消息定制邮件表单，例如便笺、警告或信函。在子表单中用到的域名不能在表单的其他地方使用。对子表单所做的更改将影响到所有使用该子表单的表单和文档。

　　3．域（Field)：域是构成表单的重要元素，对一个Notes数据库来说，外部数据的录入要通过域，库内存放数据显示也要依靠域。域类型确定了域可以包含的信息类型。用户可以在“域属性”框中定义域类型。域类型是：文本、日期/时间、数字、对话框列表、复选框、单选按钮、列表框、组合框、RTF 文本、作者、姓名、读者、口令和公式。域的主要属性是：  
　　可编辑：数据可以通过按钮执行公式（Formulas）或脚本（Script）来产生。  
　　隐含的：只作存储，没有显示作用，显示文档时不显示该域的数据。  
　　准许多值：单一域准许用户存入多个值。

　　4．文档（Document）：在Lotus Notes中，信息是以文档的形式保存在数据库中的，一个文档相当于关系型数据库中的一个记录。

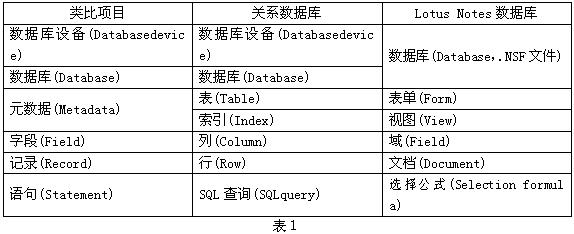
　　5．视图（View）：视图是 Lotus Notes中文档的主要浏览窗口，当一个视图的选择条件给定以后，通过该视图所看到的文档就是符合条件的文档，每一视图都包含符合一定条件的文档。视图除了有选择条件外，还可以按不同的特性将文档进行分类和排序，使得我们可以及其快捷地导航到要查找的文档。对于简单的查询，可以不编写任何程序，而通过把视图按合理的方式进行分类和排序就可以了。对用户来说是直观的，可打印输出屏幕上显示的视图。如某一视图的选择条件为:Select form="通知"；则打开该视图后，我们所看到的文档都是通知。

　　6．文件夹（Folder）： 文件夹也是文档的浏览窗口之一，但是和视图不同的是，文件夹没有选择条件。由于Lotus Notes的文档是可以共享的，文件夹也可以共享，也可以私有。里面的文档通过Putinfolder放进去，必须通过RemoveFromFolder来将其中的文档移开。

　　不难看出，在Notes的文档数据库当中，模式也是表格的定义表单（Form）和域（Field)，它用来向用户显示和保存数据。而外模式通常是一个用户的能看到视图（View）或是文件夹（Folder）。

　　从某种角度来看，一方面表单（Form）提供用户浏览保存在NSF文件中的文档信息的方法，Form也是一种内模式的机制，我们可以实现为不同的用户定制不同的Form，以达到不同的用户他们能看到的文档信息(Form)是不同的。另一方面由于在Notes当中，表单（Form）本身并不保存在文档里，只在显示文档时才起作用，Form也是一种数据库的外模式的体现。另外不像关系数据库，一个Table中的一行数据就必须是符合Table定义的。Notes文档数据库中的一份文档可以不符合任何的格式定义，也没有相应的数据约束，数据约束只在用户透过Form输入数据时是有效的。

　　我们可以认为，Notes中视图（View）或是文件夹（Folder）是一个外模式，它提供用户浏览和查找文档数据的方法，我们可以为不同的用户定制不同的视图或文件夹。例如：在常见的公文处理应用中，文件审批者和承办人，能看到视图的内容是不一样的。Notes的文档数据库与关系数据库的类比关系（见表１）



**二．NSF数据库与关系数据库比较的技术特点：**

　　Lotus Notes/Domino是一个较优秀的群件工作[平台](http://www.csai.cn/incsearch/search.asp?key=%C6%BD%CC%A8)，拥有较强的应用开发能力和编程能力，适应于企业级的管理和与[互联网](http://www.csai.cn/incsearch/search.asp?key=%BB%A5%C1%AA%CD%F8)络的集成等性能。并且它有很好的电子邮件系统，领先的全文检索和复制功能。对信息的查询就是对文档的查询，信息检索技术作为Lotus Notes/Domino的一个主要技术，为用户提供了包括全文检索、按关键字查询、视图和文件夹等多种方式。它还具有极强的[安全](http://www.csai.cn/incsearch/search.asp?key=%B0%B2%C8%AB)机制，可以可靠地保证信息安全性。但Notes不支持关系，NSF中的文档不支持加锁，Lotus Notes数据库作为非关系数据库管理系统，它的特点是：①．基于文档组织模式；②．使用非结构化的数据元素；③．通过定时复制刷新数据；④．使用视图定位数据；⑤．通过全文检索访问数据。关系数据库管理系统的特点是：①．基于事物处理模式；②．使用结构化数据元素；③．提供对数据的实时访问；④．使用实时查询定位数据；⑤．通过SQL语言访问数据。

**三．访问Notes数据库的方法：**

　　Lotus Notes群件系统，它不仅仅是个应用软件，更是一个开发平台。开发Notes[应用系统](http://www.csai.cn/incsearch/search.asp?key=%D3%A6%D3%C3%CF%B5%CD%B3)，除了Notes自身的开发环境外，开发者还可根据自己的技术、经验以及工作环境来决定使用第三方厂商提供的开发工具。

　　1．通过ODBC（Open Database Connectivity，开放式数据库互接）方式访问Notes数据库，该方式需要安装NotesSQL（ODBC Driver for Domino），驱动程序可到IBM公司的技术支持网站（<http://www-142.ibm.com/software/sw-lotus/products/product4.nsf/wdocs/notessqlhome>）[下载](http://www.csai.cn/incsearch/search.asp?key=%CF%C2%D4%D8)，其最新版本是NotesSQL 3.02g。值得注意的是Notes尽管可以通过Notes ODBC访问，但它并不支持全文检索。

　　2．使用COM（The Component Object Model，组件对象模型）和OLE（Object Link and Embed，对象链接和嵌入）访问Notes数据库。

　　①．OLE作为[Windows](http://www.csai.cn/incsearch/search.asp?key=Windows)的一个特征，通过把应用程序的功能作为对象暴露给其他应用程序，这些对象拥有属性(数据)和方法(函数)，应用程序可以通过定义、调用OLE中的对象去执行相应的任务。Domino既可以作为一个OLE提供者被外部程序调用，也可以通过OLE去控制其他OLE对象。例如：[Visual Basic](http://www.csai.cn/incsearch/search.asp?key=Visual+Basic)可以利用OLE功能，通过CreatObject获得Notes.NotesUIWorkspace 和 Notes.NotesSession对象来访问Notes类。使用OLE有许多优点但也存在一些弊端，如Domino的OLE服务不支持早期绑定，应用程序中对OLE编程进行类型检查、快速启动时，需要强迫启动Notes的客户端，对系统压力大且不能同时启动多个NotesSession等。

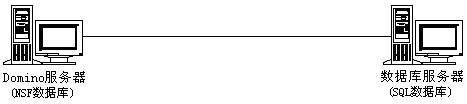
　　②．Lotus Notes/Domino在R5.03以后的版本中，提供了一个后台的COM类NotesSession。它是一个功能强大的对象模型，它提供了其他Notes对象的接口，即只有通过它才能访问Notes中的数据库对象NotesDatabase、NotesView等对象。 由于COM与OLE的兼容性，使应用程序操作Domino变得十分容易，应用程序不需要调用Lotus Notes客户端就可以访问Notes的函数；且可以在设计时使用早期绑定，其速度比晚绑定快；COM的运行速度比Lotus Script快；可以独立于Lotus Notes客户端进行应用程序发布。如在VB中支持后台的Domino对象，即Session，但对于UIWorkSpace 并不支持。

　　3．使用第三方厂商提供的开发工具开发Notes应用系统，例如：Powersoft公司的Notes库。PowerBuilder（以下简称为PB）作为客户／服务器计算环境下优秀的客户端开发工具，在访问Notes数据库时就采用了不同于以往的数据库连接方式，即利用PB提供的Lotus Notes数据库应用开发库(LLN，Power Builder Library for Lotus Notes)来实现对Lotus Notes数据库的存取。依托PB下的Notes库就可利用数据窗口访问Notes数据库，并将其结合到应用程序中，利用PB的特点开发出更完美的Notes应用程序。

**四．Notes数据库与其他关系型数据库的数据**[**交换**](http://www.csai.cn/incsearch/search.asp?key=%BD%BB%BB%BB)**工具：**

　　异构数据环境已经是一个普遍存在的问题，Windows常用开放数据库互连标准ODBC进行数据交换与操作，后来[微软](http://www.csai.cn/incsearch/search.asp?key=%CE%A2%C8%ED)又推出了ADO技术。在Notes应用开发过程中，经常需要实现Notes数据库与其他关系型数据库(如Oracle、Sybase、SQL Server、DB2等)的数据交换，为此Lotus公司推出了相关组件来完成Notes数据库与RDBMS(关系型数据库管理系统)之间高效、快速、可伸缩的批量数据交换，如Notes Pump和LEI等。

　　早期的Notes Pump组件，它作为Lotus公司Domino/Notes 系列产品之一，是一个企业级规模的数据分发服务器，它可实现RDBMS和Notes 数据库之间数据的周期性复制，主、辅Domino服务器之间数据的及时和自动的同步。数据的移动可以是调度安排的，也可以是轮询或手动的，还可以被某个特定的条件所激发。是一个基于服务器的32位多任务数据传输引擎（见图1）。

  
图１

　　Notes Pump包含三部分：服务器、管理员、开发客户端。Notes Pump服务器必须安装在已装有必要的Notes DLL的平台上，但其运行可独立于已安装的Notes服务器或客户端。"管理员"是一个Notes 的应用数据库，存储在Domino服务器上，且可运行于任何装有Notes Client4.11或更高版本的平台上。Notes Pump开发客户是一个可选组件，安装后可为后面创建"活动(Activity)"时提供方便，并支持Notes Pump LotusScript Extensions编程。装有Notes Pump 服务器的机器还需要有相应的数据库通讯软件，如Oracle SQL \*NET，DB/2CAE，ODBC等，以便和相应的数据库服务器连接。

　　LEI（Lotus Enterprise Integrator）是Notes Pump的后继产品，利用LEI工具的连接及行为来实现NOTES与SQL数据库的连接，实际上起到中间服务器的作用，它使不同数据库通过它的connection及activity文档来实行数据交换。示意图如下（见图2）：

  
图２

**五．分析Notes数据库的工具：**

　　在这里本人向大家推荐一款Notes数据库分析工具——《Domino数据库浏览器》，可以到南山软件屋网站（<http://www.nssoft.net/>）中的“原创软件”栏目中下载。《Domino数据库浏览器》是一个小工具，可以用来浏览NSF（Domino数据库）文件中的所有文档数据，并将其导出。既可以浏览Domino服务器上的所有文档，也可以打开浏览独立的NSF文件，作为软件开发的协助分析工具。

　　综上所述，我们必须清醒地认识到文档数据库与传统关系型数据库应用[程序开发](http://www.csai.cn/incsearch/search.asp?key=%B3%CC%D0%F2%BF%AA%B7%A2)系统有很大的差别，当然，这涉及到具体应用等很多其他的问题，如：应用是否适合用Notes来开发。试图将传统的应用程序开发技术映射到Notes环境中的开发人员，将很难获得Notes应用程序的优点。