计算机科学2000Vol 27№.5

基于 LDAP 的目录服务器的研究与实现*)

Study and Implementation of LDAP Directory Server

赵 明 郭常杰 卢文龙 常桂然
(东北大学软件中心 沈阳 110006)

Abstract This paper describes the directory server implemented by ourselves, discusses its structure and some important problems, implementation of storage mechanism, replication mechanism and security mechanism of directory. At last, we compare it with other products in the world.

Keywords LDAP, Directory service, Object, Directory replication

1 前言

最初目录服务是作为查询网络上一些变化不频繁信息—网络地址、电子邮件地址等的技术手段出现的[1]。随着计算机网络的日益发展，特别是 Intérnet 的发展，网络上的资源日益增多，为了能统一管理这些资源，开始把目录服务技术用于对网络上各种资源—用户、设备、应用软件等的管理[2]。因此可以看到，目录服务已成为当今网络上的一种基础平台软件、并且逐渐成为网络的核心。

目录服务的国际标准是 ISO 和 ITU 定义的 X.500系统建议，此建议系列规定了信息模型、通讯协议、安全认证等各个方面[3]。但由于 X.500运行在 OSI 的七层模型上，并没有得到广泛的支持。此外，X.500只定义了协议而未定义应用程序接口 API，阻碍了 X.500应用软件的开发，为了解决这些问题，密执根大学的学者开发了轻量级目录访问协议 LDAP 来解决这些问题，LDAP 最初是作为一个目录访问协议出现的[4]，用于简化 X.500客户端软件的开发，而 LDAP 第三版协议的推出，则为开发独立的 LDAP 服务器提供了标准[5]。

本文所介绍的目录服务器是一个支持 LDAP 第三版协议标准、支持主从和主主复制模式以及具有良好用户界面的管理工具的完整系统。

2 服务器结构模型

整个服务器由支持 LDAPV3的目录服务模块、支持主从、主主复制的复制服务模块和一个基于 WWW 浏览器的管理维护工具模块三部分组成。

目录服务模块的主要作用是处理 LDAP 客户程序的请求，并将从数据库中查询的结果返回给 LDAP 客户程序，因此可以将目录服务模块分为两大部分：前端和后端。前端负责服务器和客户程序之间的通讯，主要功能是完成协议解释和分析；后端则具体完成数据库中数据的查询和修改。

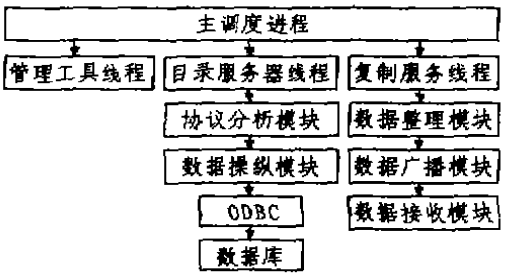

图1 目录服务器总体结构

复制服务模块用于目录服务器之间数据的复制，以维护系统的稳定性。

目录服务器的管理工具模块主要是完成对目录服务器的配置，同时也为管理员监控目录服务系统提供了一个方便的接口。

3 系统的实现的若干问题讨论

LDAP 协议只规定了目录服务的信息模型和通讯中的格式，并没有规定目录服务器如何实现，特别是 LDAP 协议目前并没有有关目录复制的标准。在具体的实现过程中，除了有关协议的分析部分外，各种产品在存贮机制、复制机制、安全机制等方面各有不同，这

*)本文得到国家863项目资助，课题编号863-306-ZT05-05-5。

几方面也是实现目录服务器的关键技术。

3.1 存贮机制

目录的信息模型是一个面向对象的层次模型，其信息存放在构成一个称为目录信息树 DIT 的结构中。

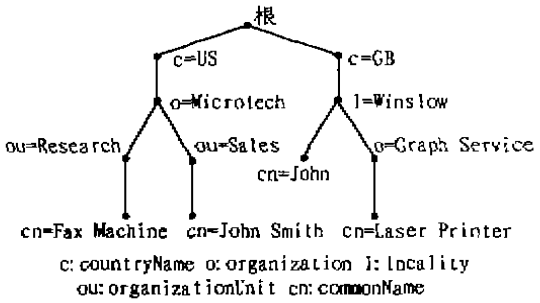

图2 目录信息树(DIT)

如图2所示，DIT 中的每个结点表示一个对象，称为目录项(entry)，根结点到每个目录项的路径称为 DN，如：cn=John Smith，ou=Sales，o=Microtech，c=US。由于每个目录项是一个对象，各目录项具有的属性不同。为了存储 DIT，我们采用对象的序列化技术，分别设计了以下表：dn2id，id2entry，id2children 和属性索引表。

dn2id	id2entry	id2children	attribute index
dn1->id1	id1->entry1	id->id1，…，idn	value1->id1，…，idn
……	……	……	……
dnN->idN	idN-entryN	idn->	valueN->id1，…，idn

DIT 中的每个目录项都可用一个 DN 表示，为了查询方便，我们为每个结点分配一个 id，这些信息通过 dn2id 表表示。DIT 中每个目录项的具体内容用目录项表示，存放在 id2entry 表中。Id2children 表中存放每个非叶结点的孩子结点的信息，每个非叶结点的 id 对应其孩子结点的 id。这样，通过这三个表就可以把整个 DIT 完整地表示出来了。

属性索引表是供用户通过属性查询时使用的，对哪些属性建立索引由用户通过目录服务器的管理工具指定，如对用户姓名等建立索引文件后，就可以通过用户的姓名，快速地查询到目录中对应该姓名用户的其它信息，如 email 等。

查询过程如下：目录服务器得到用户输入的属性信息后，通过属性索引表找到对应的 id，再通过 id2entry 表得到所需的目录项，返回给 LDAP 客户。

当用户需要把目录中的数据导出时，目录服务系统可以通过 id2childiren 表将目录中的信息按 LDIF[6] 格式输出。

3.2 复制机制

只有一个独立的目录服务器显然无法构造分布式环境，因为它无法保证目录信息的分布式传输，同时单目录服务器也降低了系统的可靠性。解决这些问题的方法是建立一套灵活的分布式复制方案。复制是这样的一种机制，通过它目录数据可以在多个分布在不同地理位置的服务器之间自动地拷贝，这些目录数据可以是整个目录树，或是一个子树，甚至是一个单独的节点。

目录复制技术，从更新的方式上，一般分为两种，即主从式目录复制和多主式目录复制。主从式目录复制又称为单主式复制，这种复制模型假设只有一个主服务器具有写或更新操作的权限，所有提交到从服务器的更新操作都会被推荐到主服务器，待执行完成后，再广播给每个从服务器。这种服务器的特点是结构简单，易于实现，并且速度较高，得到了广泛的应用。多主式目录复制又称为主主式复制，这种复制模型使得目录项能够在任何一个服务器上更新，而不用在更新之前和其他的服务器之间协商，待更新完成后，再通知其他的服务器。多主式又分为两类，一类是以 Microsoft 的 Active Directory 采用的弱一致性收敛的多主式目录复制(multi-master loose consistency with convergence)，这种模型能保证系统在更新完成后收敛于某一稳定点，即达到一致，但在更新过程中的某一点不一定每一个服务器都保持相同的视图；另一类称为强一致性收敛的多主式复制模型(multi-master strong consistency with convergence)，典型的有 Novell 公司的 NDS(Novell Directory Server)，这种模型使得任一时刻，所有的服务器都保持一致的视图，这虽然比较理想，但需要耗费大量的 CPU 和网络资源，而且大大增加了系统实现的复杂性。

可以看出，主从式复制是一种快速且易于实现的方案，但它过于简单，有时不能满足需求；而多主式复制较为强大，但无论是配置还是实现都比较复杂，并且耗费资源，尤其是强一致性收敛的方式。

在这样的一种情况下，我们在能够满足需求的前提下制定了一套快速而且易于管理的复制模型，即主从式和弱一致性收敛的多主式相结合的复制模型。使用这种模型，管理员根据实际的情况，灵活配置，即可将系统变为自己需要的方案。

复制服务器主要完成的工作是在目录服务器之间复制目录数据，在多个目录服务器已经建立的情况下，主要任务是复制数据的变化，因此，主要包括待复制数据的广播和复制数据的接受两大主要部分。另外，配置的管理，服务器的管理，日志的管理等部分也是十分重要的部分。

图3描述了整个复制服务模块的框架结构。图中，复制服务模块是做为目录服务器的一组线程存在，并且由服务器在启动时创建的。复制服务模块主要分为数据整理模块、广播模块和接收模块等。另外从管理的角度出发，复制服务模块还包括配置管理、日志管理等模块。下面描述各个模块的主要功能：

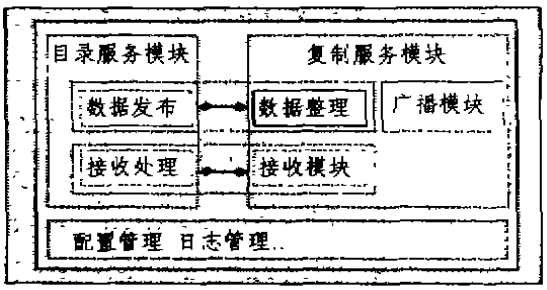

图3 目录复制模块结构图

数据整理模块：该模块包括在目录服务器中的数据发布部分和一个独立的数据整理线程，它们配合在一起，主要的功能是将需要发布的数据整理成一个队列，以便让广播模块广播出去。之所以分成两个部分（采取独立的数据整理线程），是因为需要减少目录服务器的负担，不降低目录服务器的性能。

广播模块：需要复制到其他目录服务器的数据利用该模块复制出去，该模块实际上是一个线程组，对应于每一个需要复制的服务器有一个专门广播的线程。这些线程将要复制的数据解释成为 LDAP 标准的操作（如 modify，add，delete 等等），然后向所要复制的服务器进行操作。我们之所以采用标准的 LDAP 操作，是因为目前的目录复制尚未有一个统一的标准，只有采用标准的 LDAP 操作，才能够在异类服务器之间实现简单复制。

接收模块：接收模块主要任务是辨别到达的 LDAP 操作是正常的客户请求还是复制信息，同时做相应的处理。对于复制信息，此时需要检查"邮戳"，以保证复制系统的收敛性，而对于非复制信息，要在入口中添入"邮戳"属性，以便在复制过程中使用。

管理模块：包括配置管理和日志管理。其中，配置管理十分重要，它主要是定义了一套文法，用于让客户描述对整个分布式复制环境的配置，这些配置信息在系统初始化或 SYNC 的时候传递到系统的内核中去。

3.3 安全机制

目录服务器作为 Internet/Intranet 上的核心平台之一，存放许多重要信息，如：用户的权限、证书等。因此，一个目录服务系统是否"安全"得到越来越多的关注。我们的目录服务器的安全系统由以下四部分组成：口令验证、安全传输（SSL）、访问控制列（ACL）和日志系统。

我们的目录服务器除了支持非登录连接外，也支持口令验证登录连接，用户输入登录的 DN 和口令后，服务器可以验证用户是否是合法用户。如是，可根据用户的 ACL 为用户分配相应的权限。

通过输入口令进行身份认证并不可靠，因为别人可以在网络上监听到口令。我们的服务器支持 SSL3.0 协议，完成了基于公开密钥的身份认证。客户和服务器建立 SSL 连接时将协商一个密钥，以后通讯都用它加密，提高了信息传输中的安全性。我们的服务器允许目录管理员根据目录服务系统自身安全的需要，设定是否同时启动普通端口和 SSL 端口、只启动普通端口或只启动 SSL 端口。

LDAP 协议没有对访问控制作出规定，我们设计的 ACL 机制由两种 ACL：目录系统的 ACL 和目录项的 ACL 构成。系统的 ACL 供目录系统的管理员对整个 DIT 的访问权限进行设定，目录项的 ACL 供目录项的所有者设定自身属性的访问权限。

日志系统也是我们目录服务安全系统的一个重要组成部分，我们的目录服务器的日志由访问日志和出错日志两种日志组成，记录了访问者的信息、目录操作的信息、出错信息。当系统发生安全问题时，管理员可以通过查看日志分析原因。

3.4 其它特点

我们的目录服务器还具有以下特点：

·分别建立中英文索引，管理员可以对不同的属性分别建立中文索引或英文索引，提高了索引的效率。对中文索引还可以指定基于字还是双字索引。

·基于 WWW 的联机维护管理工具，作为目录服务系统的一个重要组成部分以一组目录服务系统的线程存在，可以在不关闭目录服务模块和复制服务模块的状态下，完成对目录服务器和复制服务器的设置和性能调整。

4 与国外同类产品的比较

当前国际上已有许多目录服务产品，比较流行的有：Novell 的 NDS，Netscape 的目录服务器以及 Microsoft 将要在 Windows2000中推出的 Active Directory。其它还有 Banyan 公司的 StreetTalk，IBM 以 DB2 为后台的目录服务器和 Oracle 公司以 Oracle8I 数据库为后台的目录服务器等等。下面分别把我们的服务器和 NDS，Netscape 的目录服务器简要作下对比。

Novell 的 NDS 以采用 X.500标准，据 Novell 自

（下转第60页）

$$\cap Y=A\}$$
$$=\frac{1}{1-k}\sum_{\cap Y=A}\Big(\sum_{e_1}P_{E_1}(e_1)\Big)\Big(\sum_{e_2}P_{E_2}(e_2)\Big)$$
$$=\frac{1}{1-k}\sum_{\cap Y=A}m_{E_1}(X)m_{E_2}(Y)$$
$$=\frac{\sum_{X\cap Y=A}m_{E_1}(X)m_{E_2}(Y)}{1-\sum_{X\cap Y=\varnothing}m_{E_1}(X)m_{E_2}(Y)}$$
$$=(m_{E_1}\oplus m_{E_2})(A)$$

到此就推出了 Dempster 规则。

讨论与结论 由引理1～3及上述定理可见，Dempster 规则的成立需要四个条件，即前面的假定1～3以及证据框 E_1 与 E_2 之间的先验独立性。Shafer[3]对 Dempster 规则的表述强调了被合成的两个信度函数必须是完全独立的(entirely distinct)，但他对这种独立性未给出严格的描述。这个概念的不精确性是导致指责 Dempster 规则的根源之一。例如，一些批评论据本质上是基于对信度函数独立性的不恰当理解，因此其实并不成立。本文在相容性观点的框架内给出了信度函数独立性的确切定义(定义7)。不难看到，要验证两个证据(或两个信度函数)的这种独立性是困难的。

在 Dempster 规则成立所需要的三个假定中，假定1和假定2在直觉上是比较合理的。假定3相对较难理解。在对先验概率 $P_{E_1\wedge E_2}$ 更新时，假定3仅考虑了证据"命题 $C(E_1,E_2)$ 为真"。文[5]的实例表明，一方面，这个证据往往比证据框 E_1 与 E_2 的合取所包涵的信息少，但另一方面，它又是唯一确定的证据。从这个角度看，假定3并不是限制性很强的条件。

概括起来，在上述四个条件下，Dempster 规则是精确成立的。另一方面，这些条件也构成了应用 Dempster 规则的限制，特别是信度函数的独立性不易验证。

参 考 文 献

1 Dempster A P. Upper and lower probabilities induced by a multivalued mapping. Annl. Math. Stat., 1967,38

2 Dempster A P. A generalization of Bayesian inference (with discussion). J R Stat. Soc., 1968,B30

3 Shafer G. A mathmatical theory of evidence. Princeton University Press, 1976

4 张文修，梁怡. 不确定性推理原理. 西安交通大学出版社，1996

5 杨春，李怀祖. 证据理论在一类侦察信息综合中的应用. 二炮工程学院学报，1997,3

6 张尧庭，杜劲松. 人工智能中的概率统计方法. 科学出版社，1998

7 Halpern J Y, Fagin R. Two views of belief: belief as generalized probability and belief as evidence. Artl Intell, 1992,54

8 Lingras P, Wong S K M. Two perspectives of the Dempster-Shafter theory of belief functions. Int. J. Man-Mach. Stud., 1990,33

9 Kyburg H E. Bayesian and non-Bayesian evidential updating. Artl. Intell., 1987,31

(上接第86页)

已的介绍，最新的 NDS8将可以支持10亿个对象。缺点是 X.500协议比较复杂，特别是为了支持 LDAP 客户端的访问，增加了一个 LDAP/DAP 的网关来实现对 LDAP 客户端的支持，致使 NDS 的性能不如纯 LDAP 服务器。

Netscape 公司的目录服务器采用 LDAPV3协议实现，降低了实现的复杂性，性能比 Novell 公司的服务器好，最新的 Netscape DS4.0可支持5000万个对象。

我们的目录服务器是纯 LDAP 服务器，由于采用了 ODBC 连接数据库，其容量和查询速度与后台数据库的性能有很大关系。

在复制方面的区别已在第3.2节中做了比较。

总之，三种目录服务器的区别主要在于其应用定位不同，Novell 的 NDS 用于网络管理，需要较大的容量，Netscape 的目录服务器用于其应用服务程序的集成，所以没有使用关系数据库。我们开发的目录服务器用于建立企业信息目录，所以使用了 ODBC 去连接不同的关系数据库，以保证企业原有的投资。由于目录服务的分布式特征，单个目录服务器容量大小意义不大。

参 考 文 献

1 Tang A, Scoggins S. 开放式网络和开放系统互连. 电子工业出版社，227～275

2 Judd S. Strassner J. Directory-enabled Networks Information Model and Base Schema. Available at: http. // murchiso. com/den

3 ITU-T X.500(93)建议系列. X.500, X.501, X.509, X.511, X.518, X.519, X.520, X.521, X.525

4 Yeong W, et al. Lightweight Directory Access Protocol, RFC1777. March 1995

5 LDAPV3 标准系列: RFC2251, RFC2252, RFC2253, RFC2254, RFC2255, RFC2256

6 Gordon Good. The LDAP Data Interchange Format (LDIF)-Technical Specification, draft-ietf-asid-ldif-02. txt, July 1997. Available at: ftp: // ietf. org/internet-drafts/draft-ietf-asid-ldif-02. txt