

基于 IP 的 SAN 互联技术的研究^{*}

傅湘林 韩德志 谢长生

(华中科技大学计算机学院信息外存储国家重点实验室 武汉430074)

Study on Interconnecting Technology of IP-based SAN

FU Xiang-Lin HAN De-Zhi XIE Chang-Sheng

(National Storage System Laboratory, School of Computer Science, Huazhong University of Science and Technology, Wuhan Hubei 430074, P. R. China)

Abstract Both the requirement of the application and the development of the technology invoke the research of the new networking technology in network storage, currently the IP -based SAN has become the new focus. The main networking technology used in the IP -based SAN is: iSCSI, FCIP, iFCP, mFCP. These protocols all implement the transmitting of SCSI block I/O over the TCP/IP. To understand the protocols deeply, this paper analyzes and compares the protocols intensively.

Keywords IP-based SAN, Networking technology, iSCSI, FCIP, iFCP, mFCP

1 前言

传统 SAN 的互联技术是 FC(Fibre Channel),但其自身的缺陷如实现成本高、设备的互操作性差、跨越的地理距离短(10km)等阻碍了它的进一步普及;远程应用的发展,如远程数据备份、容错、灾难恢复等,要求 SAN 跨越的距离越来越远,甚至是跨越 MAN/WAN,传统 FC-SAN 已不能满足目前应用的需要,因此新的 SAN 互联技术应运而生;技术的进步,如吉比特以太网等高速网络技术的出现也为 SAN 互联技术的研究提供了物质基础。目前,可以用于 SAN 互联技术除 FC 之外,还有 iSCSI、FCIP、iFCP、mFCP 等。

iSCSI 定义的是 SCSI 到 TCP/IP 的映射,即将主机的 SCSI 命令封装成 IP 数据包,在 IP 网络上传输,到达目的节点后,再恢复成封装前的 SCSI 命令,从而实现 SCSI 命令在 IP 网络上的直接、透明传输。由于 TCP/IP 网络技术及其应用的飞速发展,有理由相信未来的网络将建立在 TCP/IP 网络的基础之上,iSCSI 的出现正是应运这一发展趋势。它分别整合了现有的存储和网络方面的两种主流协议:SCSI 和 TCP/IP,实现了存储和网络的无缝融合。FCIP 是一种隧道技术,简单地说,它将 FC 帧封装到 IP 数据包中,再通过 IP 网络传输,到达目的 SAN 之后,再解封装,恢复封装前的 FC 帧。其作用是将两个或多个 SAN 孤岛用 IP 网络连接起来,使得多个 SAN 看起来象一个规模更大的 SAN,从而突破单个 SAN 地理范围的限制。iFCP 是一种网关到网关的协议,它是利用基于 IP 的互联设备(包括交换机和路由器等)来代替基于 FC 的互联设备以构建一个基于 IP 的存储网络。其应用的场合是用户有大量的 FC 设备,希望这些设备可以连接到 IP 网络。至于 mFCP,除了用 UDP 代替 TCP 外,mFCP 和 iFCP 是相同的。

这些标准分别由各大公司提交给 IETF 组织,目前仍是该组织的协议草案,还没有形成正式的标准。随着时间的推移,

协议草案可能会进一步的完善。由于涉及到不同厂商的商业利益,IETF 目前不参与对这些草案标准的优劣进行评价。事实上,由于应用的场合不同,也难以有一种统一的评判标准,目前它们是在平行发展。由于这些标准都是实现 SCSI 块 I/O 在 TCP/IP 网络上的传输,都涉及到 TCP 封装其他协议的 PDU 等,尤其是后三者,都是将 SCSI 命令、数据封装到 FC,形成 FC 帧,然后再封装到 TCP 数据报中,最终在 IP 网络中传输,很容易混淆。为此,本文从协议的层次结构、协议的实现模型、命名和寻址、帧的封装等网络协议的几个重要方面做深入的分析 and 对比,来阐述这几种协议的最新发展,从而更好地认识和掌握这几种协议。

2 协议的层次模型

如图1所示为几种协议的协议栈层次结构图。在 iSCSI 协议栈中,SCSI 命令集相当于 OSI 模型中的应用层,和 FC 不同,它是由 TCP/IP 来提供传输层和网络层的服务。SCSI 和 TCP/IP 是两种不同的协议:SCSI 作为存储协议,它的特点是高速、低延迟、传送距离短,在存储总线上,计算机和存储设备是主从关系,计算机主机是所有外部设备的控制者,它适合传送大块数据(block data);而 TCP/IP 是网络协议,它的特点是低速、高延迟、传送距离长,在 TCP/IP 网络中,所有的设备都是平等的,它适合传输大量的小块消息(message)。为了利用网络来传送大数据块,必须在 SCSI 和 TCP 之间增加一个 iSCSI 层以提供两种协议之间的一个接口,使得 SCSI 的命令和数据可以在 TCP 上顺利传输。另外也应该对 TCP 层进行改造和加强,如在 TCP 之上增加同步和导向层(Sync and Steering),在 TCP 之下增加较低功能层(Lower Function Layer)等。

FCIP、iFCP、mFCP 都是实现含 SCSI 块数据的 FC 帧在 IP 网络上的传输,它们的协议栈结构非常相似。从图1可以发现,在 FC-2层以上三种协议完全相同。和 iSCSI 不同,在这三

^{*} 本文受国家自然科学基金(编号:60173043)和国家“973”重大基础研究项目(编号:G1999033006)资助。傅湘林、韩德志 博士研究生。研究方向:网络存储系统。谢长生 教授,博士生导师,研究方向:网络存储系统,采用新原理的超高密度、超高速存储技术。

种协议中, TCP/IP 封装的对象不是 SCSI 命令集, 而是 SCSI 命令集经过 FC-2 层封装之后的 FC 帧。在 FC 协议栈中, FC-2 层除负责数据组帧外, 还负责顺序管理、链路管理、缓冲存储器管理、错误检测和纠正、流量控制、服务类型、登录/注销等传输控制服务。FCIP、iFCP、mFCP 三种协议中, 除利用 FC-2 层的组帧服务外, FC-2 层其他的传输控制服务由 TCP/IP 来

取代。因此准确的说, 这三种协议仅部分经历过 FC-2 层。图 1 还显示 FCIP、iFCP 的协议栈结构相同, 但实际上它们在对 FC 帧的封装, 封装后的 IP 包的寻址、路由等方面的机制迥然不同, 文中后续部分将做详细阐述。而 mFCP 和 iFCP 相比, 除用 UDP 代替 TCP 外, 其余是相同的。

应用程序层 (application layer)				
操作系统 (OS)				
SCSI 命令集 (SCSI command set)				
FC-4: FCP.....	iSCSI	FC-4	FC-4	FC-4
FC-3: 公共服务	Syc and Steering	FC-3	FC-3	FC-3
FC-2: 数据成帧等	TCP	FC-2	FC-2	FC-2
FC-1: 8/10 编码解码等	Lower Functional Layer (LFL)	TCP	TCP	UDP
FC-0: 物理介质	IP	IP	IP	IP
F C	Link	Link	Link	Link
	iSCSI	FCIP	iFCP	mFCP

图1 协议栈层次结构图

3 协议的实现模型

iSCSI 采用 Initiator/Target (启动设备/目标设备) 的实现模式: Initiator 发出 iSCSI 命令, 而 Target 接收到该命令后, 负责解释和执行。图 1 所示 iSCSI 协议栈中, 在 Initiator/Target 端实现的功能主要包括 iSCSI 层和 TCP/IP 协议栈的功能, 根据这两部分功能实现的层次不同可以分为以下三种实现方式: ①纯软件方式: 采用通用的以太网卡, 只需在发起端和目的端设备上安装相应的 iSCSI 软件模块, iSCSI 和 TCP/IP 协议栈功能层都由主机 CPU 完成。由于采用的是标准的网卡, 这种方式无需额外的硬件, 其硬件成本最低, 但由于 iSCSI 层和 TCP/IP 层的功能都由主机 CPU 完成, 通过该网卡的既有网络流量又有存储流量, 随着两种流量的增加, 主机的运行开销就会大大增加, 从而造成主机性能的下降, 严重时还可能使主机成为系统的瓶颈, 影响应用的进行。②智能 iSCSI 网卡实现方式: iSCSI 层的功能由主机来完成, 而 TCP/IP 协议栈功能由智能 iSCSI 网卡来完成。和方式①相比, 部分降低了主机的运行开销。③ iSCSI HBA 卡实现方式: 采用主机总线适配器的方式, iSCSI 层和 TCP/IP 协议栈功能均由该主机总线适配器来完成。这种方式对主机的 CPU 的需求最

少, 相对主机而言, 就是一个标准的 HBA, 可以很容易地移植到其他操作系统中。

FCIP 是一种隧道技术, 它通过 IP 网络使得本地的 FC-SAN 可以延伸到异地与之相连的 FC-SAN, 使得通过 IP 网络相连的两个或多个 FC-SAN 共同构成一个更大的 FC-SAN, 虽然存在 FC 和 TCP/IP 网络, 但在两种网络的接合处, 不发生协议的转换, 只进行协议帧的封装: 将 FC 帧封装到 TCP 数据报中, IP 网络的作用类似于网络互联中的网桥 (bridge), 所以 FCIP 一般实现为 FCIP 网桥 (FCIP bridge)。如图 2 所示, FCIP 网桥由两部分实体 (Entity) 构成: FC Entity 和 FCIP Entity。相应地可以将一个存储网络分为两个部分: 基于 FC 的 FC 域和基于 IP 的 IP 域, 分别负责提供和 FC Fabric、IP 网络的接口。在其 FC 域看来, FCIP 网桥相当于一个 E_Port, IP 域中的 IP 网络相对 FC Fabric 而言是透明的。因此整个一对 FCIP 网桥实现的是一对 E_Port 到 E_Port 的连接。在 FCIP 一侧的 FC 域中, 必须有 FC 的互联设备 (如 Hub、Switch、Router) 来构成 FC Fabric。即便没有 FCIP 的参与, FC 域本身也是一个独立的 FC-SAN, FCIP 的作用就在于将两个或多个相隔很远的 FC-SAN (FC Fabric) 连接起来形成一个统一的、规模更大的 FC Fabric。

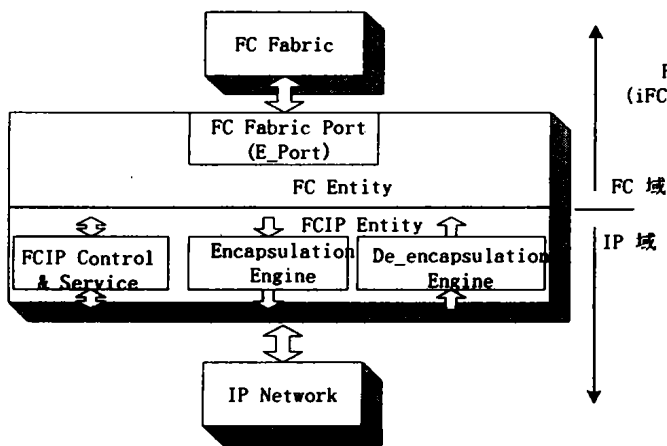


图2 FCIP网桥实现模型

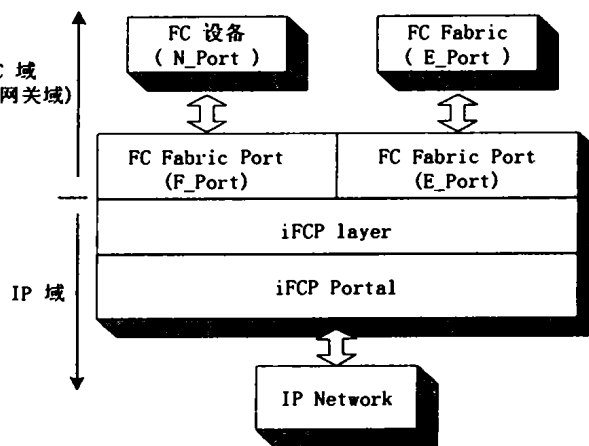


图3 iFCP网关实现模型

iFCP 则不同,在 FC 和 TCP/IP 两种网络的交接处发生的是 FC 和 TCP/IP 两种协议的转换,通过协议转换,使得 FC 帧可以在 TCP/IP 网络中独立自由地进行寻址和路由,因此 iFCP 一般实现为 iFCP 网关。如图3所示,iFCP 网关将整个网络也分为 FC 域和 IP 域。在 FC 域的一端,iFCP 网关实现为 FC Fabric Port (F_Port/E_Port),它向 FC 域中的设备提供 FC Fabric 的服务,所有的 FC 设备不用通过 FC 的互联设备就可以直接连到 iFCP 网关。在 FCIP 中,FC Fabric 的服务是采用 FC 来实现的,而 iFCP 中 FC Fabric 的服务是由 IP 网络通过 IP 协议来实现的。IP 网络相对 FC 域中的 FC 设备而言也是透明的,但它实现的是一对 F_Port 到 F_Port(或 E_Port 到 E_Port)的连接。在 FC 域中,iFCP 网关本身可以提供 FC Fabric 服务,不需要 FC 的互联设备来提供。另外,由于 iFCP 网关也可以对 FC 域中的设备提供 E_Port 的服务,它也支持 FC Fabric 联入 iFCP 网关,因此 iFCP 也可以实现 FCIP 的功能。

4 CDB/帧的封装和路由

在这几种协议中,TCP/IP 并不是直接传送 SCSI 的 CDB (Command Description Block)或 FC 帧,而必须先对 SCSI CDB/FC 帧进行封装。在 iSCSI 协议栈中,SCSI CDB 经过 iSCSI 层封装成 iSCSI PDU(Protocol Data Unit),再经 TCP/IP 传送,由于整个 iSCSI 协议栈中只有一种网络协议 TCP/

IP,不存在不同网络地址的转换以及网络协议间的映射等,因此,iSCSI 的封装操作比较简单:只需对 SCSI CDB 加上各种相应的头部信息(如 BHS: Basic Header Segment, AHS: Additional Header Segment 等)。

FCIP 采用了 FC 帧的公共封装机制,直接在 FC 帧上加入 FCIP 头部等相关信息构成 FCIP 帧,除了用于区别每个 FC 帧,在帧头和帧尾插入开始字节和结束字节外,FC 帧的结构和内容都不发生改变,也不对 FC 帧内部进行校验计算,如果 FC 帧中出现一个错误,那么这个错误的 FC 帧也会被原封不动地传送到远程的 FC-SAN 中。FCIP 帧仅其被封装的 FC 帧部分保存相应的原本在 FC-SAN 中的 FC 地址信息,而 FCIP 头部不含任何地址信息,因而 FCIP 帧不能在 IP 网络中独立寻址或路由。在 FCIP 协议中,由 FCIP Special Frame (SF)来实现 FCIP 帧在 IP 网络中集中的寻址和路由。在构造 FCIP SF 帧时,通过向相应的名字服务器发出查询请求,得到目的节点的 IP 地址,并将其封装到 SF 帧的头部。每次建立 TCP 连接的时候,首先传送一个 SF 帧以建立 TCP 连接,然后才是 FCIP 帧的传送。FCIP 帧到达目的 SAN 之后,FC 帧被解封装,再根据 FC 帧中的地址信息,利用 FC 的相关机制进行寻址和路由。在 FCIP 中,虽然使用了 FC 和 IP 两种寻址和路由机制,但它们分别作用于 FCIP 帧和 FCIP SF 帧,这两种机制分别单独发生作用,封装的时候不发生两种地址的映射和相互替代。

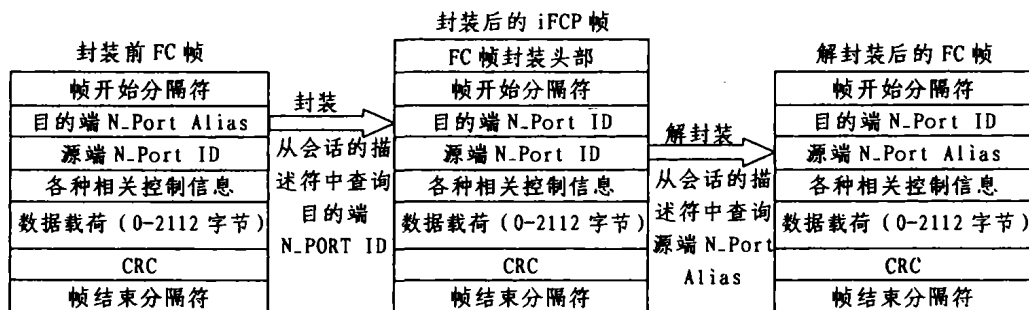


图4 iFCP 帧的封装

在 iFCP 的封装机制中,尽管采用的也是公共封装机制,但 FC 帧的结构和内容都发生改变。如图4所示,它将 FC 帧中目的端的 FC 地址映射为 IP 地址(目的端 N_Port Alias → 目的端 N_Port ID),在 iFCP 中,不存在类似 FCIP SF 的特殊帧。由于含有目的端的 IP 地址信息,封装之后的 iFCP 帧可以独立自由地在 IP 网络中路由。到达目的端后,源端的 IP 地址又被映射为 FC 地址(源端 N_Port ID → 源端 N_Port Alias)。因此在 iFCP 帧中同时存在 IP 和 FC 两种寻址和路由机制,它们同时在 iFCP 帧中发生作用。另外,iFCP 还对 FC 帧的内容进行校验计算等,对校验错误的帧或者丢失或者提示上层协议重发,防止错误的 FC 帧通过 iFCP 传送到目的地,因此一个 SAN 中的故障不会蔓延(而 FCIP 则会)。iFCP 中,每个 SAN 处于自治状态,它不像 FCIP 一样将两个 SAN 整合为一个单一的 SAN。通过 iFCP 协议,两个地理位置相距很远的 FC 设备可以通过 N_Port to N_Port 的方式建立单独的 TCP 连接,从而实现 FC 设备的远程通信。iFCP 实现的是设备到设备之间的会话,因此不同设备之间的通信可以实现不同的 QoS 策略。

5 协议应用

通过将支持 iSCSI 功能的服务器和存储设备直接联入 IP

网络,可以构建跨越 MAN/WAN 的广域存储区域网。如图5所示,FCIP 一般是通过 FCIP 网桥将两个交换式 FC-SAN 连接起来。FCIP 不支持将 FC 存储设备直接联入 IP 网络,也不支持环形拓扑结构 SAN(FCAL-SAN)。iFCP 则同时支持上述两种方式联入 IP 网络构成一个跨越 MAN/WAN 的广域存储网络,如图6所示,FC 设备(FC 服务器和 FC 存储设备)可以通过 iFCP 网关直接联入 IP 网络,对 FC 设备而言,iFCP 网关和 IP 网络共同提供 FC Fabric 服务,它在两个参与通信的异地 FC 设备之间建立一个 N_Port 到 N_Port 的连接,实

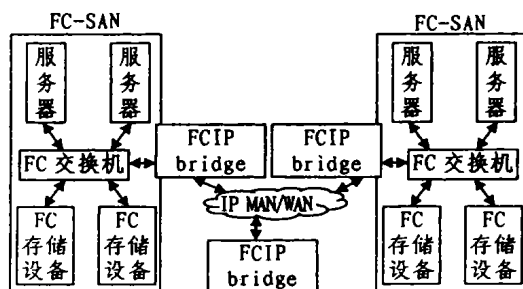


图5 基于 FCIP 的广域 SAN 的应用

现设备与设备之间的会话(session);iFCP 也可将一个或多个 FC-SAN 联入到 IP 网络,在这些 SAN 中,可以是交换结构(FC-SAN)也可以是仲裁环结构(FCAL-SAN)。因此 iFCP 也实现了 FCIP 的功能,即将地理位置上相隔很远的两个 FC-SAN 连接起来。

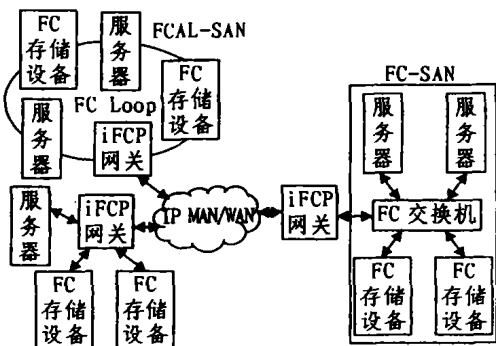


图6 基于 iFCP 的广域 SAN 的应用

总结 几种协议中,IP 网络的作用不同,因而它们应用的场合也不同。iSCSI 中,没有任何 FC 的存在,所有的设备都是基于 IP 的,TCP/IP 是 SAN 的构建技术,所得到的 SAN 也是基于 IP 的 SAN,它应用的场合是将分散于 Internet 上的存储设备、服务器(这些设备不是基于 FC)等连接起来,以构建一个跨越 MAN/WAN 的无处不在的存储网络。FCIP 中,

必须仍由 FC 互联设备来构建基于 FC 的 SAN,而 IP 网络的作用是将分散的 FC-SAN 互联成一个整体,它应用的场合是原来已经存在基于 FC 的 SAN,然后将分散的 SAN 互联起来。iFCP 则是利用 IP 网络来构建一个突破距离限制的基于 FC 的 SAN,它应用在目前大量的 FC 终端设备存在的场合下,将这些 FC 设备联入 IP 网络以便构建一个跨越 MAN/WAN 的 FC-SAN;iFCP 也可以象 FCIP 一样,用在将分散的 FC-SAN 互联起来的场合。不同的应用场合使得它们不能简单地以优劣来区分,在各自的应用场合下都具有其他技术难以比拟的优点,很难相互取代。尤其在目前大量 FC 设备存在的情况下,它们相互补充,平行地向前发展。

另外,FCIP 立足于 FC 仍是构建 SAN 的理想选择,只是利用 IP 网络来扩展 FC-SAN 跨越的距离,因此,FCIP 中,占主导地位的仍是 FC。而 iFCP 则不然,它认为 FC 不是构建 SAN 的最理想的选择,无论是在数据中心还是一些远程应用,必须采用新的构建技术,即 IP 来构建;但它正视目前大量的 FC 终端设备存在的现实,支持 FC 终端设备直接联入 SAN 或互联 FC-SAN,iFCP 中占主导地位的却是 IP。至于 iSCSI,则根本不存在 FC,而是由 TCP/IP 完全取代 FC,它构建的是一个基于 TCP/IP 的遍布 INTERNET/internet 的无处不在的存储网络。考虑到未来网络将统一在 TCP/IP 的基础之上,那么 iSCSI 应该更适合技术的长远发展。

(下转第100页)

《计算机科学》投稿须知

《计算机科学》由国家科技部主管,西南信息中心主办,系“中文科技核心期刊”、“中国科技论文统计与分析用期刊”、“中国科学引文数据库来源期刊”、“中国期刊方阵双效期刊”。《计算机科学》以其新颖、准确、及时为特色,突出动态性、综述性、学术性;主要报导国内外计算机科学与技术的发展动态,内容涉及程序理论、计算机软件、网络与信息、数据库、人工智能、人机界面、国际会议、应用等。

投稿者须知如下:

1. 内容充实、立论正确、有创新、重点突出、顺理成章。
2. 文字精炼、言简意赅,一般在 8000 字左右;并附 200-300 字中、英文摘要及题目、作者和工作单位英译名;编译稿应附原文;文末请列出主要参考文献。
3. 来稿须激光打印,字迹要清楚,上角或下角、英文大、小写字母须分明;数学符号应准确;附图须用 WORD 绘制,必须正确无误。
4. 请勿一稿两投。

论文书写格式如下(具体实例请参考我刊 2003 年第 1 期):

第 1 页:(通栏)中文标题→中文作者姓名→中文作者单位(单位地址,邮编)→中文摘要、关键词→英文标题→英文作者姓名→英文作者单位(单位地址,邮编)→英文摘要、关键词→(脚注)基金资助说明作者信息。

第 2 页起书写论文正文(双栏)。最后书写参考文献表(双栏)。

参考文献表的著录格式如下:

- a. 专著: 标引项顺序号 著者. 书名. 版本. 其他责任者. 出版地: 出版者, 出版年. 文献数量(选择项)
- b. 专著中析出的文献: 标引项顺序号 作者. 题名. 见(In): 原文献责任者. 书名. 版本. 出版地: 出版者, 出版年. 在原文献中的位置
- c. 论文集中析出的文献: 标引项顺序号 作者. 题名. 见(In): 编者. 文集名. 出版地: 出版者, 出版年. 在原文献中的位置
- d. 期刊中析出的文献: 标引项顺序号 作者. 题名. 其他责任者. 刊名, 年, 卷(期): 在原文献中的位置
- e. 报纸中析出的文献: 标引项顺序号 作者. 题名. 报纸名. 年-月-日(版次)
- f. 专利文献: 标引项顺序号 专利申请者. 专利题名. 专利国别. 专利文献种类, 专利号. 出版日期
- g. 学位论文或技术报告: 标引项顺序号 作者. 题名: [学位论文或技术报告]. 保存地: 保存者, 年份
- h. 会议论文: 标引项顺序号 作者. 题名. 会议名称, 会址, 会议年份

为了易于表示,我们先定义三个辅助函数 EXT 、 $RUP_{G_{ext}}^{e_{ext}}$ 、 $RUP_{G_{ext}}^{e_{ext}}$:

$e \in \{ \langle e_1, L, e_1, L, e_n \rangle \mid ML(m) \wedge m \in M \}$, 对 e 进行扩展增加 k 个 e_i :

$e_{ext} = \langle e_1, L, e_1, L, e_n, e_i, L, e_i \rangle, |e_{ext}| = n+k$, 定义双射函数 $EXT(e) = e_{ext}$.

以同样的方式对 G 进行扩展: $G_{ext} = \langle L_1, L, L_i, L, L_n, L_i, L, L_i \rangle$ 是 $n+k$ 元向量。

令 $G_{ext} = \langle L_1, L, L_i, L, L_n, L_{n+1}, L, L_{n+k} \rangle = \langle L_1, L, L_i, L, L_n, L_{R1}, L, L_{Rk} \rangle$

定义函数 $RUP_{G_{ext}}^{e_{ext}}(e_{ext}) = e'_{ext} = \langle e'_1, L, e'_1, L, e'_n, e'_{n+1}, L, e'_{n+k} \rangle \in dom(G'_{ext})$

if $L_i = L'_i, e'_1 = e_1$ else $e'_i = Rup_{L'_i}^{e_i}(e_i)$

定义函数 $RUP_{G_{ext}}^{e_{ext}}(e_{ext}) = e' = \langle e'_1, L, e'_1, L, e'_n, proj(e'_{n+1}), L, proj(e'_{n+k}) \rangle \in dom(G')$

$ML' = \{ \langle m', e' \rangle \mid \exists e' \in G' (m' = Agg(\{m \mid m \in M \wedge RUP_{G_{ext}}^{e_{ext}} \circ RUP_{G_{ext}}^{e_{ext}} \circ EXT \circ ML(m) = e'\}) \}$.

例5 我们将例2中定义的虚拟维 SupplierDim 加入多维数据集 MD, 其模式 MDS 同例4, $AddVirtualDim[Product \rightarrow Family, \{SupplierDim\}](MD) = MD'$, $MDS' = \{Dset' = \{ProductDim, StoreDim, CustomDim, TimeDim, SupplierDim\}, G' = \langle Family, Store, Custome, Day, Vsupplier \rangle, MS' = \langle Sales', (idT, number) \rangle\}$.

定义6.3.4(增加动态维) $AddDynaDim[L_n \rightarrow L'_i, \{A_1, L, A_k\}](MD) = MD'$

这一操作与增加虚拟维类似,不同的是:1. 这些维不是静态声明的,而是在查询中动态定义的,每个动态维只有两个等级 $L_j < ALL, dom(L_j) = dom(A_j)$; 2. 对于每一个 $A_j \in \{A_1, L, A_k\}$, P_{A_j} 是使 A_j 有定义的谓词,在这些维被加入多维数据集 MD 之前,先要对 MD 施加限定操作,限定谓词为: $P = P_{A_1} \wedge P_{A_2} \wedge L \wedge P_{A_1} \wedge L \wedge P_{A_k}$.

例6 我们将例5中输出多维数据集 MD' 作为本例中的输入, $AddVirtualDim[Custom \rightarrow Province, \{Gender, Married\}](MD') = MD''$, $MDS'' = \{Dset'' = \{ProductDim, StoreDim, CustomDim, TimeDim, SupplierDim, GenderDim, MarriedDim\}, G'' = \langle Product, Store, Province, Day, Supplier, Gender, Married \rangle, MS'' = \langle Sales'', (idT, number) \rangle\}$.

在我们的操作符集合中没有包含下钻操作,因为下钻操作总是要访问更细粒度的数据,最终仍要转换为上钻操作。同样我们也没有定义删除维的操作,在一个多维数据集中,一个维的粒度等级位于 ALL 时,就相当于维的隐含删除。

结论 在 OLAP 的研究领域,“成员属性”是一个长久以来没有得到广泛重视,更没有得到很好认识和定义的概念。本文的研究重点是一类特殊的成员属性(强/弱成员属性),它们在某些特定的条件下可以作为一个独立的维参与多维数据的表示,这种需求广泛存在于大量的 OLAP 的实际应用之中。为了给这类应用提供一个坚实的理论基础,本文提出了一个扩展的多维数据模型,用于合理表示“成员属性”及基于成员属性的虚拟维和动态维,并对相应的多维操作进行了扩展。与现有的 OLAP 产品相比,成员属性的完整性约束,以及对于动态维的支持是这一扩展模型的独特之处。

参考文献

- Hurtado C, Mendelzon A. Reasoning about summarizability in heterogeneous multidimensional schemas. In: Proc. of the 8th Intl. Conf. on Database Theory, London, UK, 2001
- <http://www.olapcouncil.org>
- Lehner W. Modeling Large Scale OLAP Scenarios. In: 6th Intl. Conf. on Extending Database Technology (EDBT'98, Valencia, Spain, 23-27 March)
- Cabibbo L, Torlone R. A Logical Approach to Multidimensional Databases. In: Proc. of the EDBT 1998
- Nguyen T B, TJOA A M, Wagner R R. An Object Oriented Multidimensional Data Model for OLAP. In: Proc. of the First Intl. Conf. on Web-Age Information Management (WAIM'00), Shanghai, China, June 2000. Lecture Notes in Computer Science (LNCS), Springer, 2000
- Agrawal R, Gupta A, Sarawagi S. Modelling Mul-tidimensional Databases. In: Proc. of the ICDE 1997
- Li C, Wang X S. A data model for supporting on-line analytical processing. In: Proc. Conf. on Infor-mation and Knowledge Management, Nov. 1996
- Gyssens M, Lakshmanan L V S. A Foundation for Multi-Dimensional Databases. In: Proc. of the VLDB 1997, Athens, Greece
- Vassiliadis P. Modeling multidimensional databases, cubes and cube operations. In: Proc. of 10 th SSDBM 1998, Capri
- Pedersen T B, Jensen C S. Multidimensional Data Modeling for Complex Data. In: Proc. ICDE' 99

(上接第49页)

参考文献

- intel iSCSI protect. <https://sourceforge.net/projects/intel-iscsi>, 2002
- Satran J, et al. iSCSI. <http://www.ietf.org/internet-drafts/draft-ietf-ips-iscsi-11.pdf>, 2002
- Rajagopal M, et al. Fibre Channel Over TCP/IP (FCIP). <http://www.ietf.org/internet-drafts/draft-ietf-ips-fcovertcpip-09.pdf>, 2002

- Monia C, et al. iFCP-A Protocol for Internet Fibre Channel Storage Networking. <http://www.ietf.org/internet-drafts/draft-ietf-ips-ifcp-10.pdf>, 2002
- Fibre Channel Industry Association. Fibre Channel Standard. <http://www.fibrechannel.com>, 2002
- Weber R, et al. FC Frame Encapsulation. <http://www.ietf.org/internet-drafts/draft-ietf-ips-fcencapsulation-06.pdf>, 2002
- Tseng J, et al. Internet Storage Name Service (iSNS). <http://www.ietf.org/internet-drafts/draft-ietf-ips-isns-09.txt>, 2002