



(12) 发明专利

(10) 授权公告号 CN 101039239 B

(45) 授权公告日 2013. 03. 06

(21) 申请号 200710096845. 5

G06Q 20/32 (2012. 01)

(22) 申请日 1998. 08. 26

(56) 对比文件

(30) 优先权数据

08/917761 1997. 08. 27 US

09/081012 1998. 05. 19 US

WO 90/04837 A1, 1990. 05. 03, 说明书第 5 页第 1 行 - 第 6 页第 36 行, 第 8 页第 34 行 - 第 14 页第 24 行、附图 1-4.

US 5326959 A, 1994. 07. 05, 全文.

US 5602936 A, 1997. 02. 11, 全文.

(62) 分案原申请数据

98810515. 2 1998. 08. 26

EP 0593209 A2, 1994. 04. 20, 说明书第 2 栏第 30-44 行, 第 3 栏第 18 行 - 第 4 栏第 15 行, 第 4 栏第 20 行 - 第 6 栏第 8 行、附图 1-2.

(73) 专利权人 数据宝库公司

地址 美国德克萨斯州

审查员 陈晨

(72) 发明人 C·R·巴拉德

(74) 专利代理机构 中国专利代理(香港)有限公司

司 72001

代理人 张志醒

(51) Int. Cl.

G06Q 10/10 (2012. 01)

权利要求书 1 页 说明书 20 页 附图 11 页

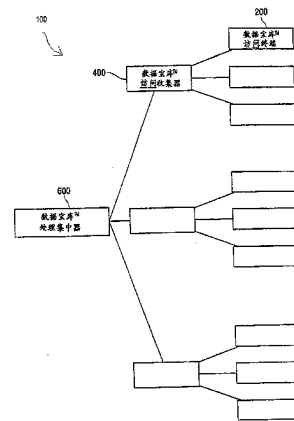
(54) 发明名称

具有集中处理和存储的远端图像捕获的系统和方法

(57) 摘要

公开了一种叫做数据宝库™(Data Treasury™) 系统的用于远端数据捕获与集中处理和存储的系统。该数据宝库™系统提供对处理与包含销售、商业、银行业和一般消费者交易的不同应用有关的文件和电子数据的全面支持。该系统检索在一个或多个远端位置的以电子或纸件形式的诸如信用卡收据支票的交易数据、加密数据、将加密数据发送到中央位置、将数据变换为可用形式、用签名数据和生物统计学数据进行识别验证、从该数据产生信息报告并将信息报告发送到远端位置。数据宝库™系统具有许多有益特征, 这些特征共同提供高的性能、安全性、可靠性、容错能力和低费用。首先, 网络结构使远端位置和中央处理设施之间能安全通信。为了更快的性能和更高的利用率, 动态地址分配算法在系统的服务器之间进行负载平衡。最后, 划分方案改进纠错过程。

CN 101039239 B



1. 用于在一个或多个远端数据访问子系统 (200)、至少一个中间数据搜集子系统 (400) 和至少一个中央数据处理子系统 (600) 中和之间传送数据的通信网络, 所述子系统形成分层结构, 所述至少一个中央数据处理子系统 (600) 的每一个与所述至少一个中间数据搜集子系统 (400) 的相应的一些通信, 并且所述至少一个中间数据搜集子系统 (400) 的每一个与所述一个或多个远端数据访问子系统 (200) 的相应的一些通信, 所述一个或多个远端数据访问子系统 (200) 包括远端数据访问子系统 DAT 扫描器, 所述通信网络包括:

至少一个用于在相应的所述一个或多个远端数据访问子系统 (200) 之一内传输数据的第一局域网;

至少一个用于在相应的所述至少一个中间数据搜集子系统 (400) 之一内传输数据的第二局域网 (406);

至少一个用于在相应的所述至少一个中央数据处理子系统 (600) 之一内传输数据的第三局域网 (606); 和

至少一个用于在所述一个或多个远端数据访问子系统 (200)、所述至少一个中间数据搜集子系统 (400) 和所述至少一个中央数据处理子系统 (600) 之间传输数据的广域网 (412)。

2. 根据权利要求 1 所述的通信网络, 进一步包括:

至少一个第一调制解调器 (204), 用于通过所述至少一个广域网 (412) 而将所述一个或多个远端数据访问子系统 (200) 的所述至少一个第一局域网连接到相应的至少一个第二局域网 (406) 之一;

至少一组调制解调器 (404), 用于通过所述至少一个广域网 (412) 而将所述至少一个中间数据搜集子系统 (400) 的所述至少一个第二局域网 (406) 连接到所述一个或多个远端数据访问子系统 (200) 的所述至少一个第一局域网的相应的一些;

至少一个第一广域网路由器 (408), 用于将所述至少一个中间数据搜集子系统 (400) 的相应的至少一个第二局域网 (406) 之一连接到所述至少一个广域网 (412); 和

至少一个第二广域网路由器 (612), 用于将所述至少一个中央数据处理子系统 (600) 的相应的所述至少一个第三局域网 (606) 之一连接到所述至少一个广域网 (412)。

3. 根据权利要求 2 所述的通信网络, 其特征在于所述至少一个广域网 (412) 包括利用传输交易数据的帧延迟方法的载波云。

4. 根据权利要求 3 所述的通信网络, 其特征在于所述至少一个第二局域网 (406) 和所述至少一个第三局域网 (606) 进一步包括用于在所述至少一个第二局域网 (406) 和所述至少一个第三局域网 (606) 内选择交易数据路由的相应的至少一个网络交换器之一; 且其中所述数据包括 (a) 来自信用卡, 智能卡和借贷卡的电子交易, 签名数据和生物统计学数据, 或 (b) 来自文件和收据的纸件交易。

具有集中处理和存储的远端图象捕获的系统和方法

技术领域

[0001] 本发明通常涉及自动处理来自包含销售、商业、银行业和一般消费者交易的不同应用的文件和电子数据。特别是,本发明涉及一种自动系统,该系统检索远端位置的交易数据,加密数据,将加密数据发送到中央位置,将数据变换为可用形式,从该数据产生信息报告并将信息报告发送到远端位置。

背景技术

[0002] 本发明涉及处理文件和电子数据,例如从包括信用卡交易、智能卡交易、自动柜员机(ATM)交易、消费者购买、商业形式、W2形式、出生证、契约和保险文件的销售、商业和银行业交易中产生这些文件和电子数据。

[0003] 从销售、商业和银行业交易的文件和电子数据中产生的大量纸件和电子记录包含有价值的信息。首先,这些纸件和电子记录包含可用于证实由消费者、商人和银行家持有的记录精确性的信息。例如,消费者使用销售和银行业交易的纸件收据以证实他们从其银行或信用卡机构收到的周期结算表的信息。商人使用纸件收据以记录用于管理消费者投诉的销售交易。纳税人使用纸件收据以记录用于其税单返回准备的可减免的部分。雇员使用纸件收据以记录用于商业消费形式的商业消费。

[0004] 纸件和电子记录还包含可用于市场分析的信息。例如,生产商和零售商可以从纸件和电子记录包含的信息中确定不同区域消费者的偏爱还有消费者偏爱的趋势。

[0005] 但是,纸件和电子记录的持有和处理表现出困难的挑战。首先,纸件收据和文件容易被丢失、放错位置、窃取、损坏或毁坏。另外,这些纸件和电子记录包含的信息不容易处理因为信息都分散于各个记录中。例如,在商人持有的一组销售记录中包含的市场趋势信息不容易确定因为该信息分散于各个记录中。同样,在消费者持有的一组销售交易纸件收据中包含的税款信息不容易处理。

[0006] 以前已经建议了一些方案以符合与持有和处理纸件和电子记录有关的挑战。例如,数据档案服务公司将从其消费者得到的纸件收据和文件信息存储在中央设施的缩微胶片或光盘只读存储器(CD-ROM)。顾客典型地把纸件收据和文件发送到中央设施。对于不能离开顾客位置的敏感文件,某些数据档案服务公司在顾客位置进行数据捕获并转换为磁带和把磁带传送到中央设施。

[0007] 这些数据档案服务公司提供的方案存在缺点。首先,该方案费用高性能差因为该方案需要把纸件收据或磁带从顾客位置昂贵而消耗时间地物理传送到中央设施。另外,该方案不可靠因为信息在传送期间可能丢失或损坏。该方案还具有有限的容量因为它不能在一个系统中处理电子记录还有纸件收据。

[0008] 其它的方案集中于消除纸件收据和文件。US专利5,590,038公开了一种作为替换纸件收据的通用电子交易卡(UET卡)或智能卡,这种卡把交易信息存储在嵌在卡上的存储器中。同样,US专利5,479,510公开了一种在购买时电子传送和存储买方信息的方法,在较晚的时间读取该信息以便确定购买的货物或服务发送给正确的人。

[0009] 虽然这些方案避开了与纸件收据有关的问题,但它们存在其它的缺点。首先,这些方案不能用交易数据的第三方接收者提供对由顾客、商人或银行家所持记录的独立正确性证实。例如,如果 UET 卡在记录销售或银行交易后被丢失、窃取、损坏、或由不道德的持有者故意改变,这些方案都不能证实它方持有的交易的剩余记录。

[0010] 其次,这些方案不能在一个单独的综合系统中同时处理交易的纸件和电子记录。相应地,这些方案不能从事处理销售和银行交易已经产生的大量纸件和电子记录的任务。当在可预见的将来将继续产生纸件收据和文件时,这些方案缺乏同时处理纸件和电子记录的能力是显著的限制,因为涉及电子的交易可靠性与顾客和商人熟悉纸件收据。

[0011] 这些方案还缺乏安全性因为这些方案不能提供典型用于信用卡购买以避免偷窃和错误的签名验证。例如,小偷在暴力得到 UET 卡后盗用 UET 卡持有者的钱,操纵或窃取用户的个人识别号码 (PIN)。同样,罪犯在得到受害人的社会安全号码之后以受害人的名义获取信用卡并进行非法交易也不奇怪。这主要关系到例如在互联网上可以得到这类的个人信息。另外,由商人对信用卡购买手工进行的签名验证经常无法察觉伪造的签名。

[0012] 尽管智能卡或 UET 卡能在卡中存储签名和其它的生物统计学数据,但该系统仍然存在缺点。首先,偷卡者可以改变卡上存储的生物统计学数据以击败安全性措施。同样,如果卡被毁坏该生物统计学数据也会被破坏。最后,安全性措施费用很高,因为它在每张卡上或每个购买点的设备上需要昂贵的生物统计学比较特征。

[0013] 已经建议了包括签名验证系统在内的其它生物统计学验证系统以便解决安全性问题。例如,美国专利 5,657,393 公开了一种验证手写签名的方法和装置,涉及提取和比较包含选择部件线 (Componentlines) 的长度和角度的签名特征。另外,美国专利 5,602,933 公开了一种使用存在中央设施的相应数据验证远端获得的信息的方法和装置。

[0014] 但是,这些验证系统都没有提供对交易初始化、远端纸件和电子数据捕获、数据加密、数据通信、数据档案、数据检索、数据开采 (mining)、操纵、和分析业务的普遍支持。相应地,需要一种单个系统,该系统提供对自动处理销售、商业、银行业和一般消费者交易的文件、生物统计学和电子数据所涉及任务的全面支持。另外,需要一种单个的综合系统,该系统具有满足零售业、商业、银行业和一般消费工业要求的可靠性、性能、容错能力、容量、费用和安全性。

发明内容

[0015] 本发明提供一种具有最大安全性和可用性以处理电子和纸件交易的自动的、可靠的、高性能的、容错的、和低费用的系统,该系统命名为数据宝库™ (DataTreasury™) 系统。

[0016] 本发明的一个目的是提供一种集中管理,存储和验证从涉及销售、商业、银行业和一般意义的消费者应用的信用卡、智能卡、借贷卡、文件和收据中远端捕获的电子和纸件交易的系统,该系统包括:

[0017] 至少一个用于捕获和发送电子和纸件交易数据的远端数据访问子系统;

[0018] 至少一个用于收集和发送电子和纸件交易数据的数据收集子系统,该子系统包括用于管理交易数据的收集和发送的第一数据管理子系统;

[0019] 至少一个用于处理、发送和存储电子和纸件交易数据的中央数据处理子系统,该子系统包括用于管理交易数据的管理、发送和存储的第二数据管理子系统;和

[0020] 至少一个用于在所述至少一个数据访问子系统和所述至少一个数据处理子系统之间传送交易数据的通信网络。

[0021] 该数据宝库™系统处理诸如信用卡收据、自动柜员机(ATM)收据、商业花费收据和销售收据的纸件和/或电子收据并自动生成诸如信用卡结算表、银行结算表、免税税款报告、市场分析、和类似的报告。

[0022] 数据宝库™系统的另一个目的是在远端位置检索纸件和电子交易。

[0023] 数据宝库™系统的另一个目的是在用户端使用扫描器和数据入口终端以便从纸件交易中检索数据并能对扫描的信息分别进行添加或修改。

[0024] 数据宝库™系统的另一个目的是为了独立验证由消费者、商人和银行家持有的记录,提供从智能卡的存储器中检索交易数据的输入设备以便防止由于丢失、偷窃、损坏或故意改变智能卡而导致数据丢失。

[0025] 数据宝库™系统的另一个目的是从由帐号和通行字识别的数据宝库™无名智能卡中检索和处理交易数据。因为数据宝库™无名智能卡交易可以不用消费者的姓名来识别,所以消费者可以把钱加到数据宝库™无名智能卡中并与现金买进和消费具有相同的保密程度地用卡进行消费。

[0026] 数据宝库™系统的另一个目的是从雇员时间文献中检索消费者帐单数据和从帐单数据中产生消费者帐单结算表。

[0027] 数据宝库™系统的另一个目的是开始包含互联网交易在内的电子交易和通过捕获和比较签名和生物统计学数据来提供身份验证。

[0028] 本发明数据宝库™系统的另一个目的是利用分层结构处理电子和纸件交易,该分层结构包括数据宝库™系统访问终端(DAT)、数据宝库™系统访问收集器(DAC)和数据宝库™系统处理集中器(DPC)。

附图说明

[0029] 从下面结合附图的详细描述中,可以更清楚地理解本发明的这些和其它的目的和特征,其中:

[0030] 图1是表示本发明的三个主要工作部分:数据宝库™访问终端(DAT)、数据宝库™访问收集器(DAC)和数据宝库™处理集中器(DPC)的方框图;

[0031] 图2是DAT结构的方框图;

[0032] 图3a是描述DAT捕获图象的流程图;

[0033] 图3b显示了由DAT处理的样品纸件收据;

[0034] 图4是DAC结构的方框图;

[0035] 图5是描述由DAC轮询DAT的流程图;

[0036] 图6是DPC结构的方框图;

[0037] 图7是描述由DPC轮询DAC的流程图;

[0038] 图8是描述由DPC进行数据处理的流程图;

[0039] 图9是描述由DPC进行数据检索的流程图;和

[0040] 图10是描述使用数据宝库™处理个人支票的流程图。

具体实施方式

[0041] 图 1 表示数据宝库™系统 100 的结构。数据宝库™系统 100 具有三个工作部分：数据宝库™系统访问终端 (DAT) 200 (远端数据访问子系统)、数据宝库™访问收集器 (DAC) 400 (中间数据收集子系统) 和数据宝库™处理集中器 (DPC) 600 (中央数据处理子系统)。

[0042] 数据宝库™系统 100 结构包括三层。在底层，DAT 200 从用户端检索数据。在下一层 DAC 400 轮询 DAT 200 以便接收在 DAT 200 积累的数据。在顶层，DPC 600 轮询 DAC 400 以便接收在 DAC 400 积累的数据。DPC 600 在中央位置存储消费者数据、从该数据产生信息性报告并把该信息性报告发送给在远端位置的消费者。

[0043] 在优选实施例中，数据宝库™系统 100 符合 Price WaterhouseSAS70 工业标准。特别地，数据宝库™系统 100 符合 Price WaterhouseSAS70 指定的软件发展标准、系统配置标准和可靠性标准。通过遵守 Price Waterhouse SAS70 标准，数据宝库™系统 100 提供特定于银行和股票经纪公司的财政应用的使命所需的安全性、可用性和可靠性。

[0044] 本领域普通技术人员都知道，数据宝库™系统 100 也可以使用其它的软件发展标准、其它的系统配置标准和其它可靠性标准，只要符合这些可替换的标准提供特定于财政应用的使用所需的安全性、可用性和可靠性。

[0045] 图 2 表示了 DAT 200 结构的方框图。DAT 200 位于用户端。数据宝库™系统 100 的用户端包括商人、消费者和银行家。DAT 200 充当与数据宝库™系统 100 提供的一套业务的用户连接点。在优选实施例中，根据运行 SUN 微系统的 JAVA/OS 操作系统的一般意义上的客户网络计算机 (NC) 传统设计 DAT 200。传统设计的 DAT 200 包括 DAT 扫描器 202、DAT 调制解调器 204、DAT 数字存储器 206、DAT 控制器 210 (工作站)、DAT 卡接口 212、可选的 DAT 打印机 208 和签名垫 214。

[0046] 本领域普通技术人员都知道，也可以根据对运行其它操作系统的一般意义上的网络计算机来传统设计 DAT 200，只要所选择的操作系统提供对数据宝库™系统 100 所需的多处理、存储器管理和动态链接的支持。

[0047] DAT 扫描器 202 扫描纸件收据并生成叫做收据的位映像图象 (BI) 的数字位映像图象表示。在优选实施例中，DAT 扫描器 202 能支持整个范围的通常以每英寸的点数 (DPI) 测量的图象分辨率值。其次，DAT 扫描器 202 能进行全双工成像。通过全双工成像，扫描器同时捕获纸件文件的正面和背面。DAT 扫描器 202 还可以支持在每个像素深度值的任何位的灰度等级和全彩色成像。DAT 扫描器 202 还支持用于身份验证的手写签名的捕获。

[0048] 除了扫描图象和电文，DAT 扫描器 202 还扫描可以从 Xerox 公司得到的数据图示符™ (DataGlyph™) 元素。本领域技术人员都知道，Xerox 数据图示符™ 技术用编码为许多、微小、独立的图示符元素的机器可读数据表示数字信息。根据扫描和打印设备的分辨率，每个图示符元素包括短到一英寸的 1/100 的 45 度对角线。分别根据其向下斜到左边或右边，每个图示符元素表示二进制数 0 或 1。相应地，数据图示符™ 元素可以表示如 ASCII 或 EBCIDIC 二进制表示法的字符串。另外，本领域技术人员都知道的加密方法加密数据用数据图示符™ 技术表示的数据。

[0049] 在数据宝库™系统 100 中使用图示符技术改进了系统的精确性、费用和性能。Xerox 数据图示符™ 技术包括纠错码，可以引用来纠正扫描错误或纠正由洒墨水或自然损

耗对文件造成的破坏。数据图示符™技术还使得系统费用减少,因为由于与数据图示符™元素有关的改进的精确性该系统需要较少的人工干预数据入口和纠正。因为数据图示符™元素在少量的空间中表示大量的信息,所以 DAT 扫描器 202 需要少量的时间以输入大量的信息。

[0050] DAT 卡接口 212 和 DAT 签名垫 214 还有通过 DAT 调制解调器 204 互联网和电话接入使得数据宝库™系统 100 用户使用包括借贷卡、智能卡和信用卡的各种卡经互联网或电话和 DAT 200 开始安全的销售和银行交易。在通过标准的互联网接口选择购买或银行交易之后,数据宝库™系统 100 用户把借贷卡、智能卡和信用卡插入或刷卡到 DAT 卡接口 212。

[0051] DAT 卡接口 212 从卡中检索识别信息以便随后发送到互联网交易的目的地。另外, DAT 扫描器 202 可以捕获文件中的手写签名或者 DAT 签名垫 214 可以捕获在其上用特殊的笔写的电子签名。同样,这些安全性特征允许智能卡接收器用位于商人端的 DAT 200 启动该卡。这些安全性特征可以检测借贷卡、智能卡和信用卡的未授权使用以免其非法窃用。相应地,数据宝库™系统 100 的安全性特征提供为互联网和电话交易比典型方法更高的安全性,典型方法只需要发送卡的帐号和到期日期。

[0052] 本领域普通技术人员都知道, DAT 200 还可以包含为了额外安全性用于捕获其它生物统计学数据的额外设备。这些设备包括面部扫描、指纹、语音印迹、虹膜扫描、视网膜扫描和手的几何形状。

[0053] 除了开始销售和银行交易, DAT 卡接口 212 还从智能卡的存储器中读取在其它地方进行的销售和银行交易。以便由数据宝库™系统按序储存和处理。如果在 DAT 卡接口 212 读取其交易数据之后,智能卡丢失、偷窃、损坏或被不道德的持有者故意改变,则数据宝库™系统 100 可以为用户再生交易数据。相应地, DAT 卡接口 212 提供支持对消费者、商人和银行家持有的记录的独立验证以防止由于智能卡的丢失、偷窃、损坏或故意改变而导致数据丢失。

[0054] DAT 卡接口 212 还支持启动和检索与数据宝库™系统无名智能卡的销售和银行交易。与标准的借贷卡和信用卡不同,数据宝库™系统无名智能卡不用名字识别持卡者。相反,数据宝库™系统无名智能卡只需要帐号和通行字。因为数据宝库™系统无名智能卡不需要用户的名字就能识别,所以数据宝库™系统 100 用户可以购买数据宝库™系统无名智能卡、在卡上加钱、用卡进行消费和监视卡的帐户,与现金买进、消费和管理的保密程度一样。

[0055] DAT 扫描器 202、互联网访问、签名垫 214 和它的生物统计学数据捕获设备还支持调查信息和购买订单的远端捕获。例如, DAT 扫描器 202 捕获在饭店和酒吧在帐单背面出现的调查。同样的, DAT 扫描仪 202 能够捕获来自住宅的购买订单,这使得顾客可以在他们的家里实现对通过电子邮件促进的货物的及时的购买。因此,国内市场批发商可以通过使用 DAT 扫描仪 202 以取代为住宅提供带有预付邮资的信封的方式以更加具有经济效率和可靠的方式进行销售的传输。

[0056] DAT 扫描器 202 也捕获在随后的纳税返回处理和税收审计所需要的收据。同样的, DAT 扫描器 202 捕获来自批发商的销售单据,提供一种非现场的保护,可靠的贮存方式以抵御由洪水、火灾或其它环境导致的丢失。这样的特点也使得批发商能够自动的以可靠和具有经济效率的方式完成商品清单编目。

[0057] DAT 控制器 210 完成通常由处理器完成的处理任务和输入/输出 (I/O) 任务。DAT

控制器 210 压缩, 加密并且标识 BI 以形成一个被标识的加密的压缩的位映射图像 (TECBI)。DAT 控制器 210 还管理输入 / 输出 (I/O)。特别地, DAT 控制器 210 管理诸如 DAT 扫描器 202, DAT 数字存储器 206, 任选的 DAT 打印机 208 和 DAT 调制解调器。

[0058] DAT 数字存储器 206 保持象 TECBI 一样的数据。DAT 调制解调器 204 将数据从 DAT200 依据 DAT 控制器 210 的指示传输至适当的 DAC400。特别地, DAT 调制解调器 204 将 TECBI 从 DAT 数字存储器 206 传输到适当的 DAC400。在优选的实施例中, DAT 调制解调器 204 是一个拨号接续性的高速调制解调器。DAT 数字存储器 206 的存储空间足够大以储存在传输到 DAC400 之前的输入数据。DAT 数字存储器 206 可以是随机读取存储器 (RAM) 或者是硬盘驱动器。

[0059] 图 3a 是详细描述 DAT 操作过程的流程图 300。在步骤 310 当中, DAT 扫描器 202 通过一个操作员的操作将纸件收据扫描入 DAT200。在步骤 312 中, DAT 控制器 210 决定操作是否被成功执行。如果扫描成功, DAT 扫描器 202 建立一个位映射图像 (BI)。如果扫描不成功, DAT 控制器 210 向操作员通报问题并提示操作员在步骤 370 修复。

[0060] 如果 BI 被建立, DAT 控制器 210 在步骤 314 中执行一个常规的例如标识的图像格式 (TIFF) 程序这样的图像压缩算法来压缩 BI。在步骤 316 中, DAT 控制器 210 决定压缩程序是否被成功执行。如果程序被成功执行, 就产生一个压缩的位映射图像 (CBI)。如果压缩不成功, DAT 控制器 210 向操作员通报问题并提示操作员在步骤 370 修复。

[0061] 如果 CBI 被建立, DAT 控制器 210 在步骤 318 中执行一个对于本领域的普通技术人员来说是熟知的加密算法来加密 CBI。加密保护了下面将详细说明的在随后的数据传递过程中的未经授权的访问。在步骤 320 中, DAT 控制器 210 决定加密程序是否被成功执行。如果程序被成功执行, 就产生一个加密的压缩的位映射图像 (ECBI)。如果加密不成功, DAT 控制器 210 向操作员通报问题并提示操作员在步骤 370 修复。

[0062] 如果 ECBI 被建立, DAT 控制器 210 在步骤 322 中使用包括扫描时间, 验证编码以证实批发商的原始扫描和任一附加的有用的信息这样的时间标记对 ECBI 进行标识。在步骤 324 中, DAT 控制器 210 决定标识程序是否被成功执行。如果程序被成功执行, 就产生一个被标识的加密的压缩的位映射图像 (TECBI)。如果标识不成功, DAT 控制器 210 向操作员通报问题并提示操作员在步骤 370 修复。

[0063] 如果 TECBI 被创建, DAT 控制器 210 在步骤 326 存储 TECBI 于 DAT 数字存储器 206 中。在步骤 328, DAT 控制器 210 决定存储操作是否成功执行。如果存储操作成功, DAT 数字存储器 206 将包含 TECBI。如果存储操作不成功, DAT 控制器 210 向操作员通报问题并提示操作员在步骤 370 修复。

[0064] 如果 TECBI 被适当的存储进 DAT 数字存储器 206 中, DAT 控制器 210 在步骤 330 中决定是否所有的纸件收据都已经被扫描。如果所有的纸件收据没有被扫描, 控制返回到下一份纸件收据将被按如上所述的过程进行处理的步骤 310。如果所有的纸件收据都已经被扫描, DAT 控制器 210 请求操作员核对在步骤 334 中的扫描的收据的数量。如果由 DAT 控制器 210 确定的扫描收据的数量与操作员所确定的扫描收据的数量不相等, DAT 控制器 210 将请求操作员是否决定重新按步骤 338 重新扫描所有的收据。

[0065] 如果操作员选择了按照步骤 338 重新扫描所有的收据, DAT 控制器 210 将删除所有的来自于步骤 342 中的 DAT 数字存储器 206 的批关联的 TECBI。当操作员准备好按步骤

346 重新扫描批收据之后,控制返回按上面的描述该批的首张收据被处理的步骤 346。

[0066] 如果操作员没有选择按照步骤 338 中的批重新扫描所有的收据,控制返回到如上所述的 DAT 控制器 210 请求操作员核对扫描的收据的数量的步骤 334 中。

[0067] 如果由 DAT 控制器确定的扫描收据的数量与操作员所确定的扫描收据的数量相等,DAT 控制器 210 就会在步骤 350 中在 DAT 打印机上打印出一张批记录单。操作员将把这张批记录单附加到已经被扫描的收据批当中。这张批记录单将包含相关的例如扫描时间,收据的数量以及数据操作员的验证编码这样的相关对话信息。如果在框图 300 中的图像捕获之后收据批处理困难产生的话,这张批记录单将使它们能被迅速地定位以使用 DAT200 重新扫描。

[0068] 在步骤 354 中,DAT 控制器 210 决定扫描对话是否完成。如果扫描对话没有完成,控制将返回到如上所述的下一批扫描对话的第一张收据被处理的步骤 310。如果扫描对话已经结束,DAT 控制器 210 在步骤 358 中通过 DAT 打印机 208 有选择地打印一张对话记录。在步骤 362 中,DAT 控制器 210 通过 DAT 数字存储器 206 写下关于对话的统计信息。在步骤 366 中,DAT 控制器 210 终止对话。

[0069] 图 3b 描述了在图 3a 的流程图中中说明的由 DAT200 处理的纸件收据的样品。该纸件收据样品包含有四个参与者的信用卡交易:

[0070] A、THE ISSUER:一个例如 GE CARITAL、GM、AT&T 这样的为信用卡提供支持并向用户发行卡的银行或法人金融社团这样的实体。

[0071] B、THE PROCESSOR:通过进行包括用 ISSUER 数据库检验的基本交易核对以确保信用卡有足够的信用以允许交易确认执行内部信用卡交易处理。

[0072] C、THE ACQUIRER:专门用于营销,设备以及电子收款机 (POS) 信用卡终端设备的支持。象在数据宝库™系统中 DAC400 一样捕获器作为一个当信用卡被插入到 POS 终端时对初始的信用卡交易起到一个电子收集点的作用。在捕获之后,捕获器将处理信息传给 PROCESSOR。

[0073] D、THE MERCHANT:将信用卡插入一个 POS 终端并且记入交易数量以启动信用卡交易。

[0074] 在本发明的优选实施例中,DAT200 从下述如图 3b 所示的纸件收据样品中读取下述信息并按照下述的格式将信息存储起来。

[0075] CUSTOMER_ID 370:这个字段是一个 7 位置的十六进制的数值值。该字段唯一的确定使用终端的用户。在本样品中,这一字段可以识别信用卡批发商。

[0076] TERMINAL_ID 372:这个字段是一个 6 位置的十进制的数值值。该字段唯一的确定被用来打印信用卡收据的信用卡终端。

[0077] TRANSACTION_DATE 374:该字段含有信用卡交易的日期和时间。

[0078] TRANSACTION_LINE_ITEM 376:该字段是一个长度可变的字符串。头三个位置提供权力校验数字字段并且由引导的零表明该字段全长。该字段包含有包括项目价格的从属于购买项目的全部数据。DAT200 将为收据的每一个交易排列项存贮一个 TRANSACTION_LINE_ITEM 字段。因为并不是所有的信用卡交易都必须排列项,所以该字段是可以任选的。

[0079] TRANSACTION_SUBTOTAL 378:该字段是一个双精度浮点数。该字段指示了 TRANSACTION_LINE_ITEM 的小计。

[0080] TRANSACTION_SALES_TAX 380 :该字段是一个双精度浮点数。该字段包含了 TRANSACTION_SUBTOTAL 的销售税金。

[0081] TRANSACTION_AMOUNT 382 :该字段是一个双精度浮点数。该字段是 TRANSACTION_SUBTOTAL 与 TRANSACTION_SALES_TAX 的总和。

[0082] CREDIT_CARD_ACCT_NUM 384 :该字段是一个 12 个位置的十进制数。该字段识别被用来执行这一交易的信用卡。

[0083] CREDIT_CARD_EXP_DATE 386 :该字段识别信用卡的终止数据。

[0084] TRANSACTION_APPROVAL_CODE 388 :该字段是一个 6 位置的数值值。该字段指示了被给予特定交易的许可码。

[0085] DAT200 还储存了下述未在图 3b 中显示的附加的项目：

[0086] ISSUER_ID :该字段是一个 7 位置的数值值。该字段识别信用卡的发行者。

[0087] ACQUIRER_ID :该字段是一个 7 位置的数值值。该字段识别捕获者。

[0088] PROCESSOR_ID :该字段是一个 7 位置的数值值。该字段识别处理器。

[0089] TRANSACTION_LINE_ITEM_CNT :该字段是一个 3 位置的十进制数值值。该字段识别收据的交易排列项目的数量。零值显示没有任何收据的交易排列项目。

[0090] TRANSACTION_GRATUITY :该字段是一个双精度浮点数。因为只有有在宾馆和酒吧的收据上才会出现,所以是可供任选的。

[0091] FINAL_TRANSACTION_AMOUNT :该字段是一个双精度浮点数。因为只有有在宾馆和酒吧的收据上才会出现,所以是可供任选的。该字段是 TRANSACTION_AMOUNT 和 TRANSACTION_GRATUITY 的和。

[0092] 在如图 3a 流程图所示的步骤 322 中预先对 ECBI 的标识识别文件的初始化的时间和地点。特别的,该表示包括下列字段：

[0093] DAT_TERMINAL_ID :该字段是一个 7 位置的十六进制数值值。该字段唯一的识别被用户使用的 DAT200。

[0094] DAT_SESSION_DATE :该字段识别产生文件的格式的 Dat200 对话的数据和时间。

[0095] DAT_USER_ID :该字段是一个 4 位置的十进制数值值。该字段识别启动 DAT200 对话的顾客组织中的个人。

[0096] DATA_GLYPH_RESULT :该字段是长度可变的字符串。前 4 个位置保持一个权力校验数值位置并由引导的零表示字段的全长。第五个位置指示了数据图示符™元素状态。0 值表示在收据当中没有数据图示符。1 值表示数据图示符存在并且不包含任何错误。2 值表示数据图示符存在并且包含有一般错误。如果该字段的第五个位置的值为 2,字符串的剩余部分将识别错误字段的号。按照后面的描述,DPC600 将核对该字段的这部分,以使用交替的方法从收据中捕获错误数据。3 值指示了数据图示符存在但是有严重的错误。换句话说,就是数据图示符™元素被严重损坏并不能被读取。

[0097] 如图 3b 所示的收据还可以包含一个可被 DAT 扫描器 202 捕获的签名。数据图示符能够识别收据上的签名所在的位置。

[0098] 正像本领域的普通技术人员熟知的那样,数据宝库™系统 100 还可以象处理带有交替格式的收据,只要该收据还包括有诸如交易量,顾客, DAT200, 交易日期, 交易税金, 信用卡号, 信用卡终止日期等适当的识别信息。

[0099] 数据宝库™系统 100 以图 3b 所示的例子的方式将纸件收据划分成为图像小块。划分促进了在扫描操作过程中改正错误的处理的改进。如果在扫描过程中错误产生的话,数据宝库™系统 100 使用手工进入的方法来修复错误。通过划分,数据宝库™系统 100 将只需要将修复的努力聚集在由错误的图形小块上,而不是修复整个文件。后续讨论的数据宝库™系统 100 的方案详细描述了划分的概念的实施例。

[0100] DAC400 组成了如图 1 和图 4 所示的列结构的主干。如图 1 所示,每个 DAC400 支持一个包含一组 DAT200 的域。每一个 DAC400 轮询在其域上的 DAT200 并且接收由 DAT200 累积的 TECBI。DAC400 被定位于有最大批发商密度的密钥中心位置。

[0101] 在最优选的实施例中,DAC 服务器 402 包含被连接到通常在 WINDOWS NT 下运行的网络的独立的数字装置公司 (DEC) SMP Alpha 41002/566 服务器。DEC Alpha 服务器管理从 DAT200 上接受的图像和数据的收集和中间存贮。

[0102] 正像本领域的普通技术人员熟知的那样,数据宝库™系统 100 还可以使用任何一个其它计算机商可以提供的许多不同的服务器,只要它们可以满足系统所要求的负载,性能和可靠性。

[0103] 在最优选的实施例中, DAC 服务器 402 也包括用来存贮由 DEC Alpha 服务器所收集和管理的图像和数据的 EMC 300 SYMMETRIX CUBE 磁盘存贮系统。DAC400 结构还使用可从 EMC 得到的 SYMMETRIX 远端数据设备 (SRDF), 以使得容纳 EMC 存储系统的多个物理分离的数据中心能在通过广域网 (WAN) 时彼此保持冗余的备份。因为 SRDF 执行后台的备份操作,它并不影响数据宝库™系统 100 的运行性能。DAC 服务器 402 还具有第二存储器。在优选实施例中,第二存储器 410 是小规模 DLT 投币式自动电唱机。

[0104] DAC 服务器 402 的 DAC Alpha 服务器把从 DAT 200 接收的图象和数据插入到数据库中,该数据库存储在使用本领域普通技术人员所熟知的数据库操作语言的磁盘存储系统中在优选实施例中,该数据库是一个可从问答器得到的相关数据库。

[0105] 本领域普通技术人员都知道,数据宝库™系统 100 可以使用可从其它厂家得到的包括实体关系模型的任何一种大量不同的数据库模型,只要所选择的数据库符合系统的存储和访问效率要求。例如,参见 Korth 和 Silberschatz 的数据库概念第二章。

[0106] DAC 400 的结构使用具有增强域名服务 (DNS) 的基于 WEB 的范例、Microsoft 元件目标模型 (DCOM)、和 Windows NT 应用程序接口 (API), 以便平衡包括 DAC 服务器 402 的服务器之间的通信和负荷。本领域普通技术人员都知道,也叫做连接的 DNS 静态地把域名请求翻译成互联网协议 4 (IP4) 地址。在 DAC 400 结构中,增强的 DNS 动态地分配 IP4 地址以便平衡包括 DAC 服务器 402 的服务器之间的负荷。

[0107] 在优选实施例中,利用来自 Microsoft DCOM 的目标设计和实现增强的 DNS。通过使用 DCOM 目标,增强的 DNS 在设置间隔从 Windows NTAPI 得到对包括 DAC 服务器 402 的每个服务器上的实时服务器负荷性能统计。根据这些负荷性能统计,增强的 DNS 把域名请求的映射调整到 IP4 地址以便把数据导向更少负荷的服务器。

[0108] 一大组调制解调器 404 轮询在 DAC 400 区域的用户位置的 DAT 200。在优选实施例中,这组调制解调器,可以得到如 CISCO AS5200 是 48 个调制解调器设备的集合,具有允许 DAC 服务器拨号到 DAT 200 的局域网 (LAN) 406 连接,而不需要 48 个单独的调制解调器和串连接。

[0109] DAC 服务器 402 和这组调制解调器 404 连接到 LAN 406。在优选实施例中，LAN 使用 100BaseT/10BaseT 转换的通信硬件层协议。本领域普通技术人员都知道，100BaseT/10BaseT 协议根据以太网模型。另外，数目 100 和 10 是指每秒兆比特的通信链路速率。在优选实施例中，CISCO Catalyst 2900 网络转换支持在连接到包括 DAC 服务器 402 的 LAN 406 的设备和这组调制解调器 404 之间的 LAN 406 连接。

[0110] 本领域普通技术人员都知道，可以利用其它的 LAN 结构以便在 LAN406 的设备之间通信。例如，LAN 406 可以使用具有循环 robin 分配算法、时分多路复用算法或统计多路复用算法的集线 (hub) 结构。

[0111] 广域网 (WAN) 路由器 408 把 LAN 406 连接到 WAN 以便在 DAC 400 和 DPC 600 之间通信。在优选实施例中，WAN 路由器 408 是 CISCO 4700WAN 路由器。WAN 路由器 408 使用帧中继连接把 DAC LAN 406 连接到 WAN。本领域普通技术人员都知道，诸如 NORTEL Magellen Passport “50” 电信转换器的其它设备可用于在 DAC 400 和 DPC 600 之间通信，只要所选择的路由器符合系统的性能和质量通信要求。

[0112] 本领域普通技术人员都知道，帧中继是统计多路复用的分组交换数据通信的接口协议，在分组交换的数据通信中，可变大小的分组 (帧) 用于完全封装用户发送的用户分组。与保证特定数据速率的专用点对点链路不同，帧中继通信提供需要保证的最小数据速率的带宽。帧中继通信根据网络的可用性还允许偶尔的较短的高数据速率突发。

[0113] 每帧封装一个用户分组并加入地址和验证信息。帧中继数据通信典型地具有 56 千字节每秒 (kb/s) 和 1.544 兆字节每秒 (Mb/s) 之间的传输速率。帧在长度上可以变换直到大约 1 千字节的设计极限。

[0114] Telco 载波云 412 是接收从 DAC 400 由 WAN 路由器 408 发送到 DPC600 的帧的通信网。本领域普通技术人员都知道，载波在本地中心局提供通信服务。这些中心局包含网络设备和装置以便把电话和数据通信互联到其自身网络和其它载波网中的其它中心局。

[0115] 因为载波共享互连网络的组成链路，数据通信必须根据可用性动态地分配给网络的链路。因为数据选择路由的动态本性，互连网络称作通信带宽的载波云。

[0116] 所有的 DAC 400 装置都是完全冗余的在线 UPS 电源以确保最大的电源可用性。另外，为了最小化用于故障检测、故障分析和修理的时间，正如本领域普通技术人员所知道的，所有的 DAC 400 装置并入故障检测和远程报告诊断。

[0117] 图 5 是描述由 DAC 400 轮询 DAT 200 和从 DAT 200 发送 TECBI 到 DAC 400 的流程图 500。在步骤 502，DAC 服务器 402 读取其区域中的第一 DAT 200 的地址用于轮询。在步骤 502，DAC 服务器 402 在其轮询的区域中读第一 DAT200 的地址。在步骤 504，调制解调器组 404 中的一个调制解调器拨号到第一 DAT 200。在步骤 506，DAC 400 确定到 DAT200 的呼叫是否成功。如果到第一 DAT 200 的呼叫不成功，则在步骤 522，DAC 400 将在通话总结报告中记录错误状态并把错误报告给 DPC600。

[0118] 如果到第一 DAT 200 的呼叫成功了，则在步骤 508，DAC 400 将证实 DAT 200 准备好发送。如果 DAT 200 没有准备好发送，则在步骤 522，DAC 400 将在通话总结报告中记录错误状态并把错误报告给 DPC 600。

[0119] 如果在步骤 508 DAT 200 准备好发送，则 DAT 200 将在步骤 510 将 TECBI 分组首部发送到 DAC 400。DAC 400 将确定在步骤 512 发送 TECBI 分组首部是否成功。如果 TECBI

分组首部的发送不成功,则在步骤 522, DAC 400 将在通话摘要报告中记录错误状态并把错误报告给 DPC600。

[0120] 如果在步骤 512 TECBI 分组首部的发送成功,则 DAT 200 将在步骤 514 将 TECBI 分组发送到 DAC 400。DAC 400 将在步骤 516 确定 TECBI 分组的发送是否成功。如果 TECBI 分组首部的发送不成功,则在步骤 522, DAC 400 将在通话总结报告中记录错误状态并把错误报告给 DPC600。

[0121] 如果在步骤 516 TECBI 分组的发送成功,则 DAC 400 在步骤 518 将比较在步骤 510 发送的 TECBI 分组首部和在步骤 514 发送的 TECBI 分组。如果 TECBI 分组首部和 TECBI 分组不匹配,则在步骤 522, DAC 400 将在通话总结报告中记录错误状态并把错误报告给 DPC 600。

[0122] 如果在步骤 518 TECBI 分组首部和 TECBI 分组匹配,则 DAC 400 在步骤 520 将设置 TECBI 分组的状态以表示它准备发送到 DPC 600。DAC400 还在步骤 520 发送状态给 DAT 200 以表示轮询和发送通话的成功结束。然后,在步骤 524, DAC 400 将确定 TECBI 是否已经从其区域的所有 DAT 200 中发送出去。如果 DAC 400 区域中的所有 DAC 200 都已经将 TECBI 发送到 DAC 400,则 DAC 400 在终止该通话之前在步骤 528 编辑 DAT 200 状态报告。

[0123] 如果 DAC 400 区域中的一个或多个 DAT 200 还没有发送 TECBI 到 DAC 400,则 DAC 400 将在步骤 526 得到区域内下一个 DAT 200 的地址。然后,控制返回到步骤 504,在那里,将如前所述地轮询 DAC 400 区域中的下一个 DAT 200。

[0124] 在优选实施例中, DAC 服务器 402 开始以最佳长话速率次数轮询和数据传送以降低数据传送的费用。除了搜查驱动器和冗余服务器之外, DAC 400 还具有两个将周期备份整个数据组的磁带备份单元。如果 DAC400 出现灾难性故障,则磁带可以检索并直接发送到 DPC 600 用于处理。当 DAT 200 轮询和数据传送进行时, DAC 400 将随其状态周期地更新 DPC 600。如果 DAC 400 出现灾难性故障,则 DPC 600 将知道出现故障的 DAC 400 已经做了多少轮询和备份。相应地, DPC 600 可以很容易地分配另一个 DAC 400 以便在出现故障的 DAC 400 区域完成 DAT 200 的轮询和数据传送。

[0125] 图 6 是 DPC 600 结构的方框图。DPC 600 积累、处理和存储用于以后由数据宝库™系统检索用户检索的图象,这些用户被授权访问这些相关的信息。数据宝库™系统检索用户包括信用卡商人、信用卡公司、信誉信息公司和用户。如图 6 和图 1 所示, DPC 600 轮询 DAC 400 并检索在 DAC 400 中积累的 TECBI。

[0126] 在优选实施例中, DPC 服务器 602 包括连接到运行 Windows NT 的普通网络的独立的数字装置公司 (DEC) 的 SMP Alpha 4100 4/566 服务器。DEC Alpha 服务器管理从 DAC 400 接收的图象和数据的收集和立即存储。

[0127] 在优选实施例中, DPC 服务器 602 还包括 EMC 3700 SYMMETRIX CUBE 磁盘存储系统,该磁盘存储系统存储由 DEC Alpha 服务器收集和管理的图象和数据。与 DAC 400 结构相似, DPC 600 结构使用可以从 EMC 得到的 SYMMETRIX 远端数据设备 (SRDF), 以使容纳 EMC 存储系统的多个物理分离的数据中心能彼此保持通过广域网 (WAN) 的冗余备份。

[0128] 正如上面讨论 DAC 400 的结构所述的,与 DAC 400 结构相似, DPC600 结构使用具有增强的域名服务 (DNS) 的基于 WEB 的范例、Microsoft 元件目标模型 (DCOM)、和 Windows NT 应用程序接口 (API), 以便平衡包括 DAC 服务器 402 的服务器之间的通信和负荷。

[0129] 工作站 604 执行 DPC 600 网络的操作控制和系统监视和管理。在优选实施例中，从 Compaq 得到的工作站 604 是运行 Microsoft WindowsNT 4.x 的 Intel 平台工作站。当 Microsoft Windows NT 5.x 可以得到时工作站 604 应当能运行 Microsoft Windows NT 5.x。工作站 604 执行 CA Unicenter TNG 软件以执行网络系统监视和管理。工作站 604 执行 SnoBound 成像软件以便显示和处理 TECBI。

[0130] 工作站 604 还通过比较由 DAT 200 远端检索的签名数据和在 DPC600 存储的签名数据进行识别验证。在优选实施例中，在工作站 604 上执行的签名验证软件（从加利福尼亚 Redwood Shores 的通信智能公司可以得到）进行识别验证。本领域普通技术人员都知道，工作站 604 可以执行其它的软件以便通过比较包括面部扫描、指纹、语音印迹、虹膜扫描、视网膜扫描和手的几何形状的生物统计学数据进行识别验证。因此，DPC 600 可以通过比较远端捕获的生物统计学数据和在 DPC600 存储的生物统计学数据来验证用信用卡进行购买的人的身份。

[0131] 本领域普通技术人员都知道，数据宝库™系统 100 可以使用来自其它集成电路厂商的具有集中处理单元的工作站，只要选择的工作站具有执行诸如传送指令、传送数据、执行具有传送数据的传送指令和存储结果的标准操作的能力。同样，数据宝库™系统 100 可以使用另外的 windows 操作系统和网络监视软件，只要所选择的软件可以监视网络中的工作站和链路并把确定的状态显示给操作员。

[0132] 远端数据入口网关 614 和远端非现场数据入口设备 616 纠正在 DAT200 数据捕获期间出现的错误。因为如讨论图 3b 的样品收据时描述的数据宝库™系统 100 划分文件，在远端数据入口网关 614 或远端非现场数据入口设备 616 的操作员只需要纠正包含错误的这部分文件或图象片断。

[0133] 划分改进系统性能、减少系统费用并提高系统质量。通过划分，DPC 服务器 602 只发送包含错误的这部分文件给远端数据入口网关 614 或远端非现场数据入口设备 616。因为在这些数据入口位置的操作员只看到包含错误的这部分文件，她可以快速识别和纠正错误。没有划分的话，操作员必须在整个文件中搜索错误。由于这种没有效率的操作，操作员需要更多的时间并且更有可能由于遗漏错误或在错误的位置进行修改而出错。相应地，划分通过提高速度和纠错过程的精确性改进系统性能和质量。

[0134] 同样，划分减少 DPC LAN 606 和 Telco 载波云 412 的业务量，因为 DPC 服务器 602 只发送包含错误的图象片断给远端非现场数据入口设备 616 或远端数据入口网关 614。相应地，划分通过减少在互联网中的带宽要求而减少系统费用。

[0135] DPC LAN 606 使连接到包括 DPC 服务器 602 和网络工作站 604 的 LAN 606 的设备之间进行通信。在优选实施例中，DPC LAN 606 使用类似前面讨论的 DAC LAN 406 的转换的 100BaseT/10BaseT 通信硬件层协议。在优选实施例中，DPC LAN 606 是支持 TCP/IP 高速 OC2 网络拓扑骨干网。CISCO Catalyst 5500 网络转换支持在连接到 LAN 606 的设备之间的 DPC LAN 606。

[0136] 本领域普通技术人员都知道，可以利用其它的 LAN 结构以便在 LAN406 的设备之间通信。例如，LAN 406 可以使用具有循环 robin 分配算法、时分多路复用算法或统计多路复用算法的集线结构。

[0137] 广域网 (WAN) 路由器 612 把 LAN 606 连接到 WAN 以便在 DAC 400 和 DPC 600 之间

通信。在优选实施例中，WAN 路由器 612 是 CISCO 7507WAN 路由器。WAN 路由器 612 使用帧中继连接把 DPC LAN 612 连接到 WAN。本领域普通技术人员都知道，诸如 NORTEL Magellen Passport “50” 电信转换的其它设备可用于在 DAC 400 和 DPC 600 之间通信，只要所选择的路由器符合系统的性能和质量通信要求。

[0138] DPC 600 具有三层存储结构以支持数据宝库™ 系统 100 中的多种存储要求。在优选实施例中，存储结构包括根据 EMC Symmetrix 公司存储系统的光纤信道 RAID 技术，在那里每箱 (cabinet) 支持一兆兆存储能力。在 TECBI 图象已被处理和在线 30 天后，它们将移到基于 DVD 的投币式自动电唱机系统。在 TECBI 图象已经在线 90 天后，它们将根据用户的要求移到基于写一次读多次 (WORM) 的投币式自动电唱机系统 608 中用于直到 3 年的更长时间的存储。

[0139] 在另一种实施例中，DPC 600 还想把高密度只读存储器 (HD-ROM) (从 New Mexico 的 Los Alamos 的 NORSAM 技术中可以得到) 配置到光存储投币式自动电唱机系统 610 (诸如从 Hewlett Packard 可以得到) 以便代替 DVD 元件增加存储容量。HD-ROM 符合 CD-ROM 式要素金属 (factor metallic) WORM 盘。HD-ROM 目前在一个上底板具有超过 320 吉字节 (320GB) 的非常大的存储容量和在一个底板上具有超过若干兆兆字节 (TB) 的预期容量。DPC 600 使用 IBM 和 Philips 技术以便从 HD-ROM 读取和写到 HD-ROM。

[0140] DPC 服务器 602 的 DPC Aalpha 服务器把从 DAC 400 接收的图象和数据插入到一个数据库中，该数据库是使用本领域普通技术人员所熟知的数据库操作语言在数字存储工作系统上储存。在优选实施例中，数据库是 V8.0 Oracle 关系数据库，该数据库设计为在一个资源丰富区支持数据和图象存储。

[0141] 本领域普通技术人员都知道，关系数据库包括具有唯一名字的表集。例如参见 Korth 和 Silberschatz 编著的数据库系统概念的第三章。一个数据库方案是数据库的合理的设计。相互关联的数据库的每一张表都有属性。表中的每一行代表着表中的属性的一组值的相互关系。每一张表有一个或多个超密钥。每一个超密钥都是可以唯一的确定表中的一行的一组一个或多个属性。由于没有合适的子集也是一个密钥，所以一个候选的密钥也是一个超级密钥。一个原始的密钥也是一个被数据库的设计者选择的一个候选的密钥作为识别表中的一行的工具。

[0142] 正像本领域的普通技术人员熟知的那样，数据宝库™ 系统 100 还可以使用任何一种包括实体关系模型的其它计算机商可以提供的数据库模型，只要所选择的数据库可以满足系统的存贮和访问效率要求。参见：例如由 Korth 和 Silberschatz 编著的数据库系统概念的第二章。

[0143] 一个示范性的 DPC600 基本的方案有下列的表组成。由于属性的名称是说明性的，它们充分地定义了属性的内容。每张表格的原始的密钥被两个星号 (**) 定义。对于一个原始密钥的特定值唯一对应的数值属性都被使用 “NO” 后缀所指示。对于唯一对应内部的所有相互关联的数据库的数值属性被使用 “NUM” 后缀所指示。

[0144] I. CUSTOMER : 该表描述了数据宝库™ 系统的顾客。

[0145] A. **CUSTOMER_ID

[0146] B. COMPANY_NAME

[0147] C. CONTACT

- [0148] D. CONTACT_TITLE
- [0149] E. ADDR1
- [0150] F. ADDR2
- [0151] G. CITY
- [0152] H. STATE_CODE
- [0153] I. ZIP_CODE
- [0154] J. COUNTRY_CODE
- [0155] K. VOX_PHONE
- [0156] L. FAX_PHONE
- [0157] M. CREATE_DATE
- [0158] II. CUSTOMER_MAIL_TO :该表描述了数据宝库™系统的顾客的邮件地址。
- [0159] A. **MAIL_TO_NO
- [0160] B. **CUST_ID
- [0161] C. CUSTOMER_NAME
- [0162] D. CONTACT
- [0163] E. CONTACT_TITLE
- [0164] F. ADDR1
- [0165] G. ADDR2
- [0166] H. CITY
- [0167] I. STATE_CODE
- [0168] J. ZIP_CODE
- [0169] K. COUNTRY_CODE
- [0170] L. VOX_PHONE
- [0171] M. FAX_PHONE
- [0172] N. CREATE_DATE
- [0173] O. COMMENTS
- [0174] III. CUSTOMER_DATA_SITE :该表描述了数据宝库™系统的顾客的 DAT 地址。
- [0175] A. **DAT_SITE_NO
- [0176] B. **CUST_ID
- [0177] C. CUSTOMER_NAME
- [0178] D. CONTACT
- [0179] E. CONTACT_TITLE
- [0180] F. ADDR1
- [0181] G. ADDR2
- [0182] H. CITY
- [0183] I. STATE_CODE
- [0184] J. ZIP_CODE
- [0185] K. COUNTRY_CODE
- [0186] L. VOX_PHONE

- [0187] M. FAX_PHONE
- [0188] N. CREATE_DATE
- [0189] O. COMMENTS
- [0190] IV。CUSTOMER_SITE_DATA :该表描述了数据宝库™系统的顾客的 DAT 位置。
- [0191] A. **DAT_TERMINAL_ID
- [0192] B. **DAT_SITE_NO
- [0193] C. **CUST_ID
- [0194] D. INSTALL_DATE
- [0195] E. LAST_SERVICE_DATE
- [0196] F. CREATE_DATE
- [0197] G. COMMENTS
- [0198] V。DATA_SPEC :该表提供了文件分割和提取的数据说明。
- [0199] A. **DATA_SPEC_ID
- [0200] B. **CUST_ID
- [0201] C. DESCR
- [0202] D. RECORD_LAYOUT_RULES
- [0203] E. CREATE_DATE
- [0204] F. COMMENTS
- [0205] VI。DATA_SPEC_FIELD :该表提供了文件分割和提取的域数据说明。
- [0206] A. **DATA_SPEC_NO
- [0207] B. **DATA_SPEC_ID
- [0208] C. FIELD_NAME
- [0209] D. DESCR
- [0210] E. DATA_TYPE
- [0211] F. VALUE_MAX
- [0212] G. VALUE_MIN
- [0213] H. START_POS
- [0214] I. END_POS
- [0215] J. FIELD-LENGTH
- [0216] K. RULES
- [0217] L. CREATE_DATE
- [0218] M. COMMENTS
- [0219] VII。TEMPL_DOC :该表指定了预定的文件的分割。
- [0220] A. **TEMPL_DOC_NUM
- [0221] B. DATA_SPEC_ID
- [0222] C. DESCR
- [0223] D. RULES
- [0224] E. CREATE_DATE
- [0225] F. COMMENTS

- [0226] VIII。TEMPL_FORM:该表定义了预定文件的格式位置。
- [0227] A. **TEMPL_FORM_NO
- [0228] B. **TEMPL_DOC_NUM
- [0229] C. SIDES_PER_FORM
- [0230] D. MASTER_IMAGE_SIDE_A
- [0231] E. MASTER_IMAGE_SIDE_B
- [0232] F. DISPLAY_ROTATION_A
- [0233] G. DISPLAY_ROTATION_B
- [0234] H. DESCR
- [0235] I. RULES
- [0236] J. CREATE_DATE
- [0237] IX。TEMPL_PANEL:该表指定了预定文件的格式内的板的位置。
- [0238] A. **TEMPL_PANEL_NO
- [0239] B. **TEMPL_SIDE_NO
- [0240] C. **TEMPL_FORM_NO
- [0241] D. **TEMPL_DOC_NUM
- [0242] E. DISPLAY_ROTATION
- [0243] F. PANEL_UL_X
- [0244] G. PANEL_UL_Y
- [0245] H. PANEL_LR_X
- [0246] I. PANEL_LR_Y
- [0247] J. DESCR
- [0248] K. RULES
- [0249] L. CREATE_DATE
- [0250] X。TEMPL_FIELD:该表定义了预定文件格式的板内的字段的位置。
- [0251] A. **TEMPL_FIELD_NO
- [0252] B. **TEMPL_PANEL_NO
- [0253] C. **TEMPL_SIDE_NO
- [0254] D. **TEMPL_FORM_NO
- [0255] E. **TEMPL_DOC_NUM
- [0256] F. DISPLAY_ROTATION
- [0257] G. FID_UL_X
- [0258] H. FID_UL_Y
- [0259] I. FLD_LR_X
- [0260] J. FID_LR_Y
- [0261] K. DESCR
- [0262] L. RULES
- [0263] M. CREATE_DATE
- [0264] XI。DAT_BATCH:该表定义了在一个 DAT 对话过程中被处理的文件的批。

- [0265] A. **DAT_BATCH_NO
- [0266] B. **DAT_SESSION_NO
- [0267] C. **DAT_SESSION_DATE
- [0268] D. **DAT_TERMINAL_ID
- [0269] E. DAT_UNIT_CNT
- [0270] F. CREATE_DATE
- [0271] XII. DAT_UINT :该表定义了在一个 DAT 对话过程中被处理的文件的批的单元。
- [0272] A. **DAT_UNIT_NUM
- [0273] B. **DAT_BATCH_NO
- [0274] C. **DAT_SESSION_NO
- [0275] D. **DAT_SESSION_DATE
- [0276] E. **DAT_SESSION_ID
- [0277] F. FORM_CNT
- [0278] G. DOC_CNT
- [0279] H. CREATE
- [0280] XIII. DAT_DOC :该表定义了在一个 DAT 对话过程中被处理的文件的单元中的文件。
- [0281] A. **DAT_DOC_NO
- [0282] B. **DAT_UNIT_NUM
- [0283] C. DOC_RECORD_DATA
- [0284] D. CREATE_DATE
- [0285] DATA_SPEC, DATA_SPEC_FIELD, TEMPL_DOC, TEMPL_FORM, TEMPL_PANEL, TEMPL_FIELD 表实现了如图 3b 所示的对收据样本的描述总的文件分割算法。当 TEMPL_DOC, TEMPL_FORM, TEMPL_PANEL, TEMPL_FIELD 的交叉产品分割数据宝库™ 系统 100 中的预定的文件时 DATA_SPEC, DATA_SPEC_FIELD 的交叉产品任意地分割文件。TEMPL_FORM 定义了预定文件的格式位置。TEMPL_PANEL 定义了预定文件的内部格式控制板位置。最后, TEMPL_FIELD 表预定文件的格式的内部控制板的字段位置。
- [0286] DPC600 通过从数据库返回信息来完成数据采集和报告用于应用程序的广泛的变化。例如, DPC600 通过分析由 DAT200 捕获的收据的数据为批发商产生市场趋势的分析报告和存货清单的报告。DPC600 也能为纳税人以报告形式向纳税人或通过来自数据库的检索纳税信息向纳税准备软件那样的软件应用程序提供重要的税收信息, 该数据库原始地驻留在由 DAT200 捕获的收据, 文件和电子交易上。同样的, DPC600 还可以为税收稽查的特定的时间段提供税收信息。
- [0287] 图 7 是描述由 DPC600 轮询 DAC400 的流程图和从 DAC400 到 DPC600 的 TECBI 的传送。在步骤 702 中, DPC600 读取在其轮询域中的第一个 DAC400 的地址。在步骤 704 中, DPC600 为了传输而与 DAC400 相连。在步骤 706 中, DPC600 决定与 DAC400 的连接是否成功。如果向 DAC400 的呼叫不成功的话, DPC600 将在步骤 722 中的对话摘要报告中记录错误的条件并且向 DPC600 的管理者汇报错误。
- [0288] 如果与 DAC400 的连接成功的话, 在步骤 708 中, DPC600 将校验 DAC400 已经准备

好传输。如果 DAC400 没有准备好传输, DPC600 将在步骤 722 中的对话摘要报告中记录错误的条件并且向 DPC600 的管理者汇报错误。

[0289] 如果在步骤 708 中 DAC400 准备好传输的话, DAC400 将在步骤 710 中传输给 DPC600 一个 TECBI 包标题。DPC600 将在步骤 712 中决定 TECBI 包标题的传输是否成功。如果包标题的传输不成功, DPC600 将在步骤 722 中的对话摘要报告中记录错误的条件并且向 DPC600 的管理者汇报错误。

[0290] 如果在步骤 712 中 TECBI 包标题的传输是成功的话, DAC400 将在步骤 714 中传输给 DPC600 一个 TECBI 包。DPC600 将在步骤 716 中决定 TECBI 包的传输是否成功。如果包的传输不成功, DPC600 将在步骤 722 中的对话摘要中记录错误的条件并且向 DPC600 的管理者汇报错误。

[0291] 在步骤 716 中 TECBI 包的传输是成功的话, 在步骤 718 中, DPC600 将比较在步骤 710 中传输的 TECBI 包标题和在步骤 714 中传输的 TECBI 包, 如果包标题不能与包匹配, DPC600 将在步骤 722 中的对话摘要中记录错误的条件并且向 DPC600 的管理者汇报错误。

[0292] 如果在步骤 718 中的 TECBI 包标题可以与 TECBI 包匹配的话, DPC600 将在步骤 720 中设置一个 TECBI 包的状态用来指示它已被 DPC600 接收。并在步骤 724 中 DPC600 也将传给 DAC400 这个状态以指示轮询和传递对话的成功完成。接下来, DPC600 将决定是否在步骤 724 中来自于在它们的区域中的所有的 DAC400 的 TECBI 已全部被传输。如果所有的在 DPC600 的域中的 DAC400 都已经将 TECBI 传输给了 DPC600, 在终止传输对话前的步骤 728 中 DPC600 将编辑一份 DAC400 的状态报告。

[0293] 如果一个或更多的在 DPC600 的域中的 DAC400 还没有将 TECBI 传输给 DPC600, DPC600 将在步骤 726 中获取下一个在它的域中的 DAC400 的地址, 接下来, 控制返回到如前所述的在 DPC600 的域中的下一个 DAC400 被轮询的步骤 740。

[0294] 图 8 描述了 DPC600 进行数据处理的流程图。在步骤 802 中, DPC600 取出第一个 TECBI 包。接下来, 在步骤 804 中, DPC600 从 TECBI 包中提取 TECBI。在步骤 806 中, DPC600 将 TECBI 插入到数据库中。在步骤 808 中, DPC600 将提取包括客户标识符, 加密密钥, 来自 TECBI 以获取 ECBI 的样板标识符的标识标题。

[0295] 在步骤 810 中, DPC600 解密 ECBI 图象以获取 CBI。在步骤 812 中, DPC600 解压缩 CBI 以获取 BI。在步骤 814 中, DPC600 读取和应用针对 BI 的 BI 样板。进一步的 DPC600 将 BI 分解成图形小片并且使用在步骤 814 中使用的数据捕获规则标识 BI 样板以形成标识的位映射格式小片 (TBIS)。在步骤 816 中, DPC600 将 TBIS 提交给数据捕获操作以形成 IS 衍生数据记录 (ISDATA)。DPC600 在步骤 816 中完成数据捕获的基础上将删除 TBIS。在步骤 818 中, DPC600 在数据库中使用 IS 衍生数据记录将 TECBI 记录升级。

[0296] 在步骤 820 中, DPC600 将决定是否 TECBI 包中的最后一个 TECBI 已经被处理。如果 TECBI 包中的最后一个 TECBI 没有被处理, DPC600 将在步骤 822 中将 TECBI 包中的下一个 TECBI 提取。接下来, 控制返回到如上所述的下一个 TECBI 被处理的步骤 806。

[0297] 如果 TECBI 包中的最后一个 TECBI 已经被处理, DPC600 将决定是否最后一个 TECBI 包在步骤 824 中已经被处理。如果最后一个 TECBI 包没有被处理, DPC600 将在步骤 826 中从下一个 TECBI 包取数。接下来, 控制返回到如上所述的下一个 TECBI 包被处理的步骤 804。如果最后一个 TECBI 包已经在步骤 824 中被处理, DPC600 将终止数据处理。

[0298] 正像本领域的普通技术人员熟知的那样,一个使用者可以通过查询语言向相互关联的数据库请求信息。参见例如:由 Korth 和 Silberschatz 编著的数据库系统概念的第三章。举例说明,一个用户通过指定在选择的操作上的所要求的原始密钥和表名能检索具有特定的值的原始密钥的数据库表的所有的列。类似地,通过指定要求的原始密钥的值和具有选择操作的表,一个用户能根据具有特定值的原始密钥的多数据表检索所有的叙列。

[0299] 数据宝库™系统为其检索用户提供了简单的操作界面以使得如图 9 所述的从其相互关联的数据库提取数据交易成为可能。举例来说,数据宝库™系统的用户可以从大量的特定的交易中检索时间,数据,位置。

[0300] DPC600 通过从数据库返回信息为广泛改变应用程序执行数据采集和报告的发生。举例来说,DPC600 通过分析以 DAT 200 捕获的收据上的数据为批发商产生市场趋向的分析报告和存货清单的报告。通过从原来位于由 DAT 200 捕获的收据、文件和电子交易数据库中检索税款信息,DPC 600 还可以以报告形式向纳税人或税款准备软件提供报告重要的税款信息。同样,DPC 600 还可以提供特定时间周期的税款信息用于税款审计。

[0301] 图 9 是描述由 DPC 执行的数据检索的流程图 900。在步骤 902,DPC600 接收 TECBI 检索请求。在步骤 904,DC 600 得到用户标识符。在步骤 906,DPC 600 确定用户标识符是否有效。如果用户标识符无效,控制返回到 DPC 600 将得到另一个用户标识符的步骤 904。

[0302] 如果在步骤 906 用户标识符有效,则 DPC 600 将在步骤 908 得到用户安全性状。在步骤 910,DPC 600 接收用户的检索请求。在步骤 912,DPC 600 确定用户检索请求是否和用户安全性状一致。如果用户检索请求和用户安全性状不一致,则控制返回到 DPC 600 将得到另一个用户检索请求的步骤 910。如果用户检索请求和用户安全性状一致,则 DPC 将在步骤 914 把结果发送到由用户安全性状指示的用户。

[0303] 图 10 是描述数据宝库™系统处理支票的流程图。在步骤 1004,在付款人向收款人提供支票之前,在优选实施例中数据宝库™系统捕获在付款人远端位置的支票。或者,付款人只简单地表示或邮寄支票给收款人。在步骤 1004 在付款人远端位置的支票捕获使得付款人写的支票和收款人接收的支票的依次比较。换句话说,该步骤使得支票变更检测免于欺诈支票方案,在那里在收款人接收之前截取支票和希望用化学方法允许罪犯处理空白支票。

[0304] 在步骤 1006,数据宝库™系统在收款人远端位置捕获支票和付款人的生物统计学数据。在另一种实施例,数据宝库™系统从付款人远端位置向收款人远端位置发送表示支票的电子交易数据。在步骤 1008,数据宝库™系统通过比较远端捕获的数据和存储在中心位置的数据来执行支票和生物统计学数据的验证。有效性进而包括校验优待量和付款人的签名。

[0305] 在步骤 1010,数据宝库™系统确定验证是否成功。如果步骤 1010 的验证不成功,则系统在步骤 1012 把错误信息发送到远端位置并返回到步骤 1004 用于重新提交。如果在步骤 1010 的验证成功,则系统在步骤 1014 在中心位置产生表示支票的电子交易。表示支票的电子交易包括付款人银行的识别号码、路由选择信息、付款人帐号、付款人支票、付款人银行付款通知单、支票或付款通知单的数目、收款人银行的识别号码、收款人银行路由信息、和付款人帐号。在步骤 1016,表示支票的电子交易发送给收款人银行。在步骤 1018,收款人银行发送表示支票的电子交易到付款人银行。

[0306] 在步骤 1020, 付款人银行验证表示支票的电子交易和确定是否批准资金转帐。如果收款人银行在步骤 1020 同意批准, 则付款人在步骤 1022 将资金从付款人银行转帐到收款人银行。在步骤 1024, 数据宝库™ 系统把转帐的状态通知给付款人银行和远端位置。

[0307] 虽然已经参照特定的优选实施例描述了上述本发明, 但是本发明的范围不局限于这些实施例。本领域技术人员可以找到这些优选实施例变换方式, 但是都落入本发明的精神内, 本发明的范围由下面所附的权利要求书定义。

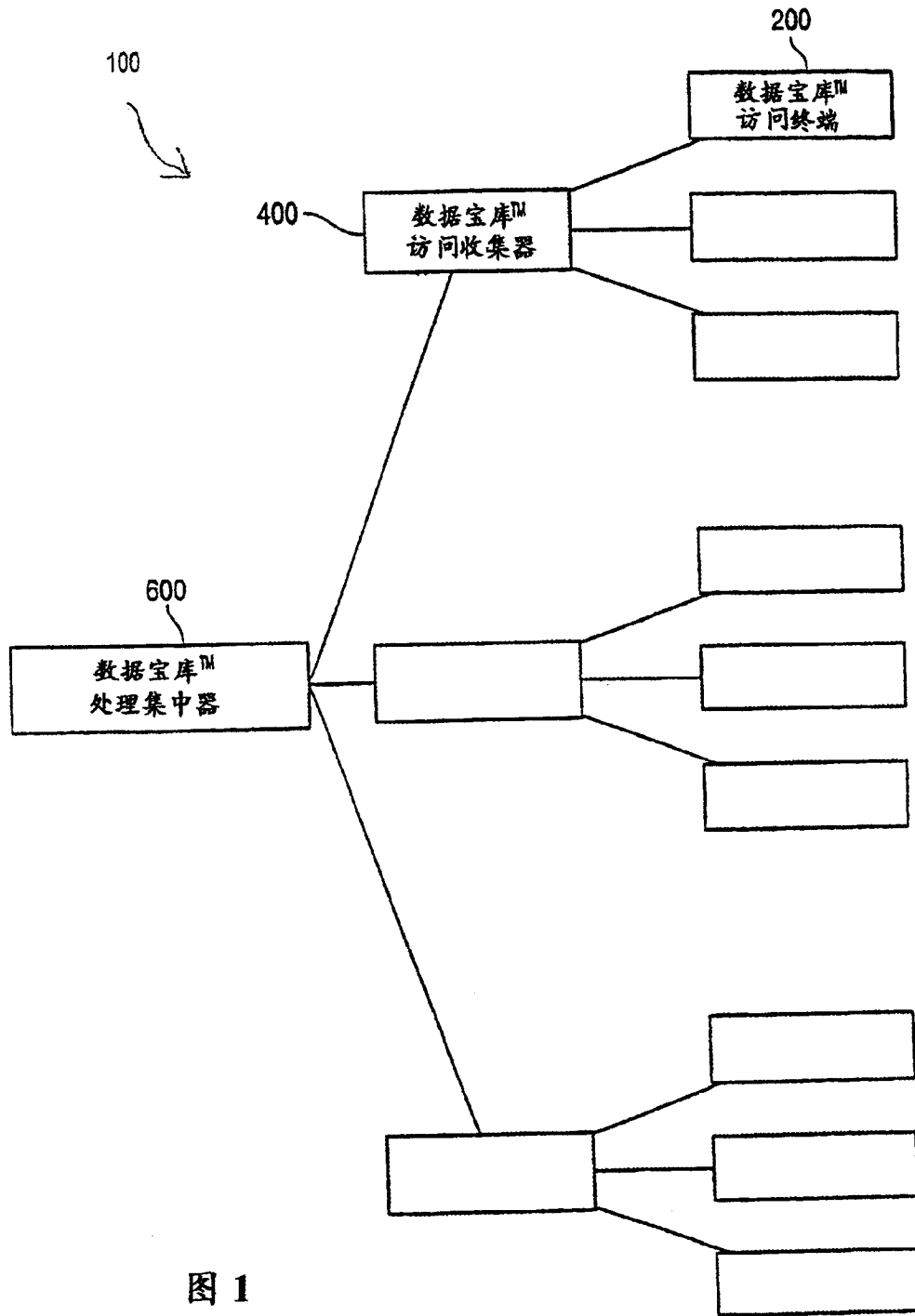


图 1

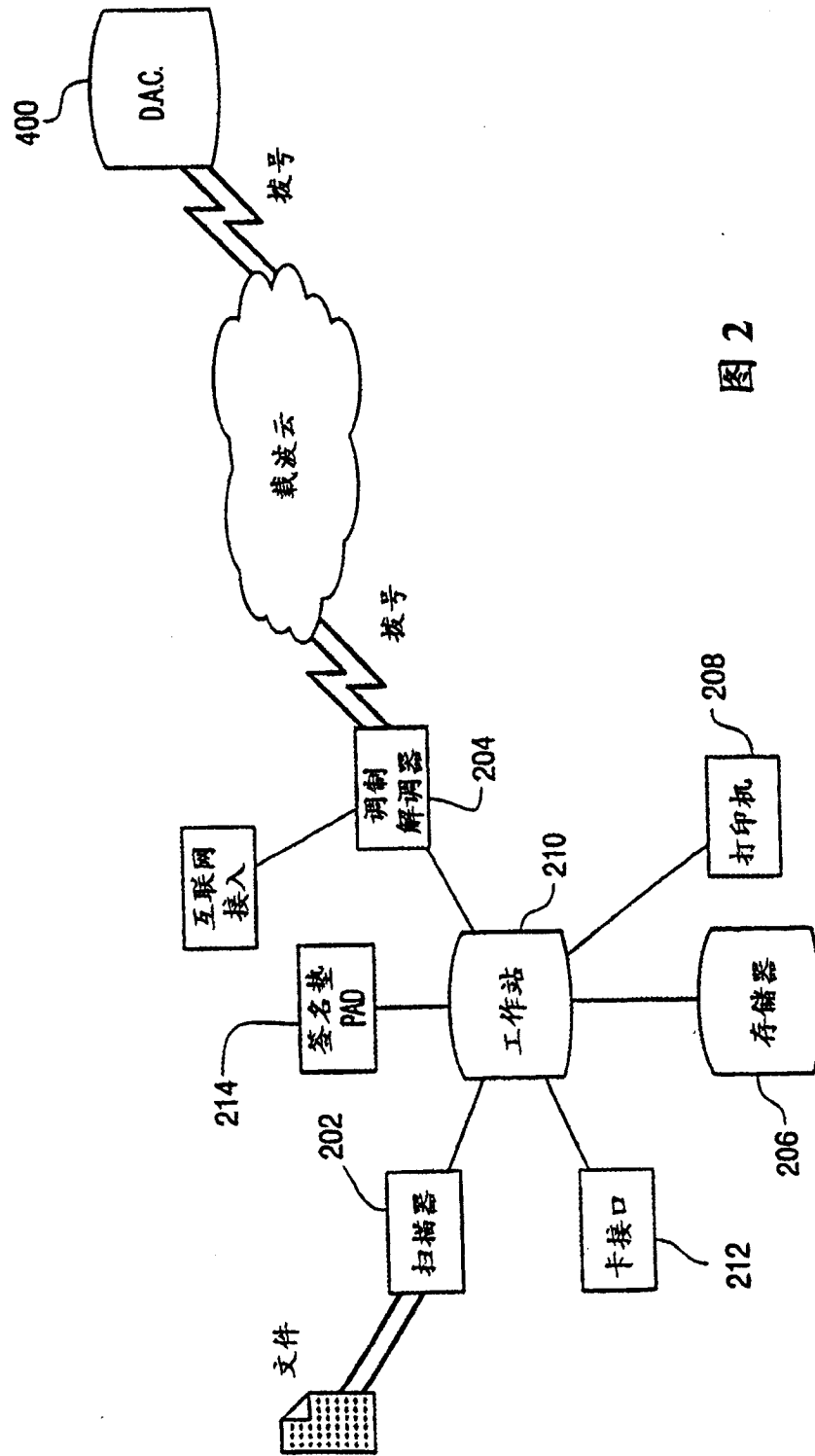
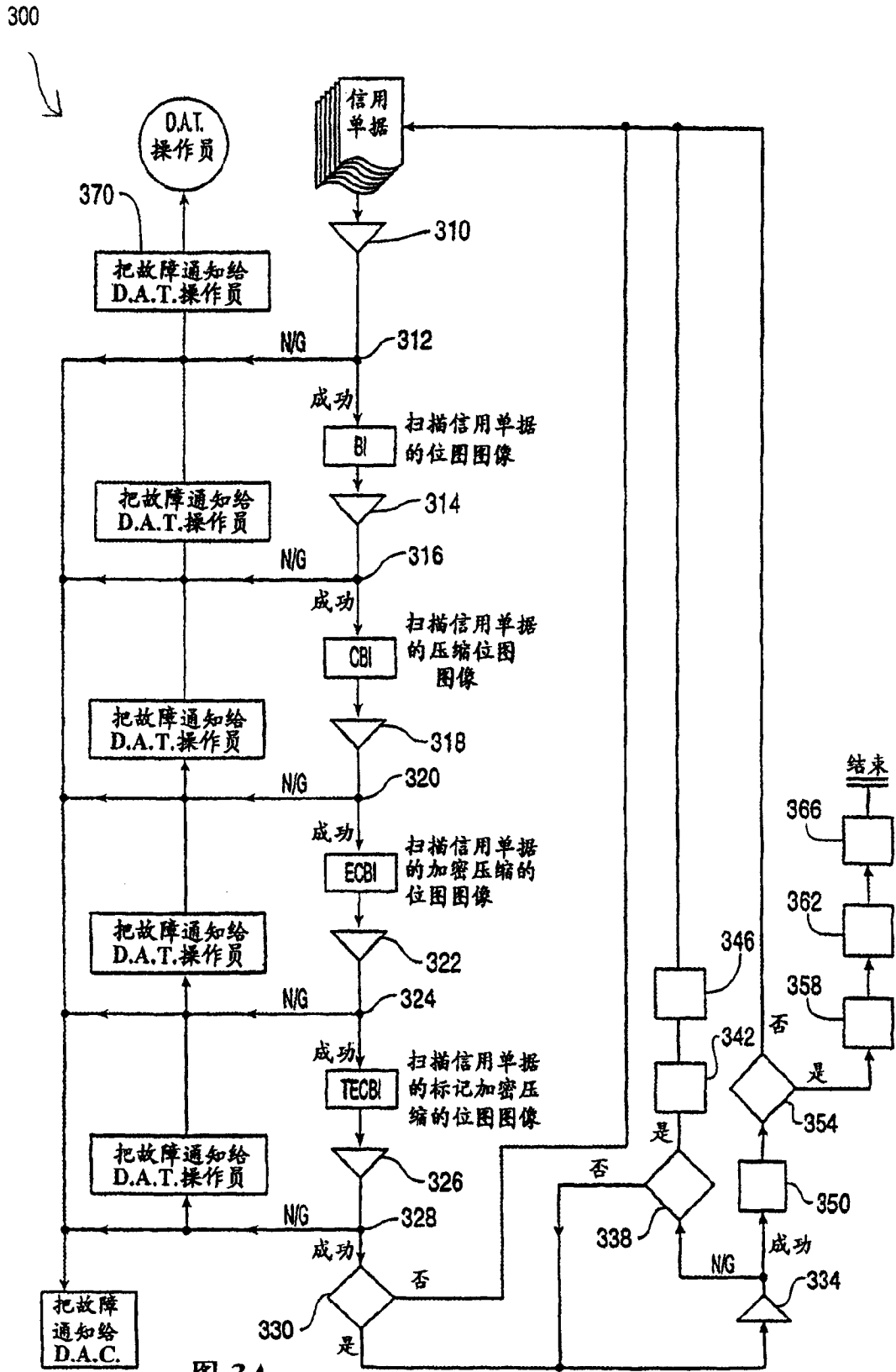


图 2



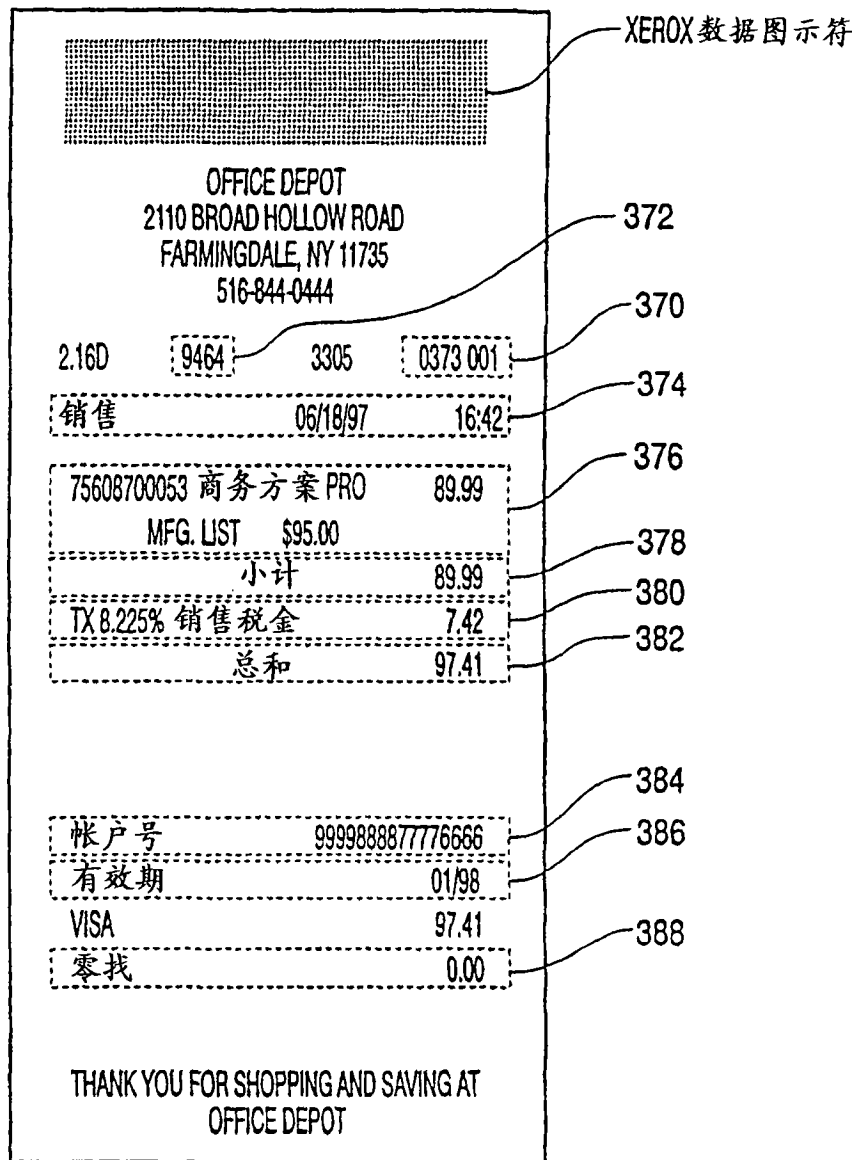


图 3B

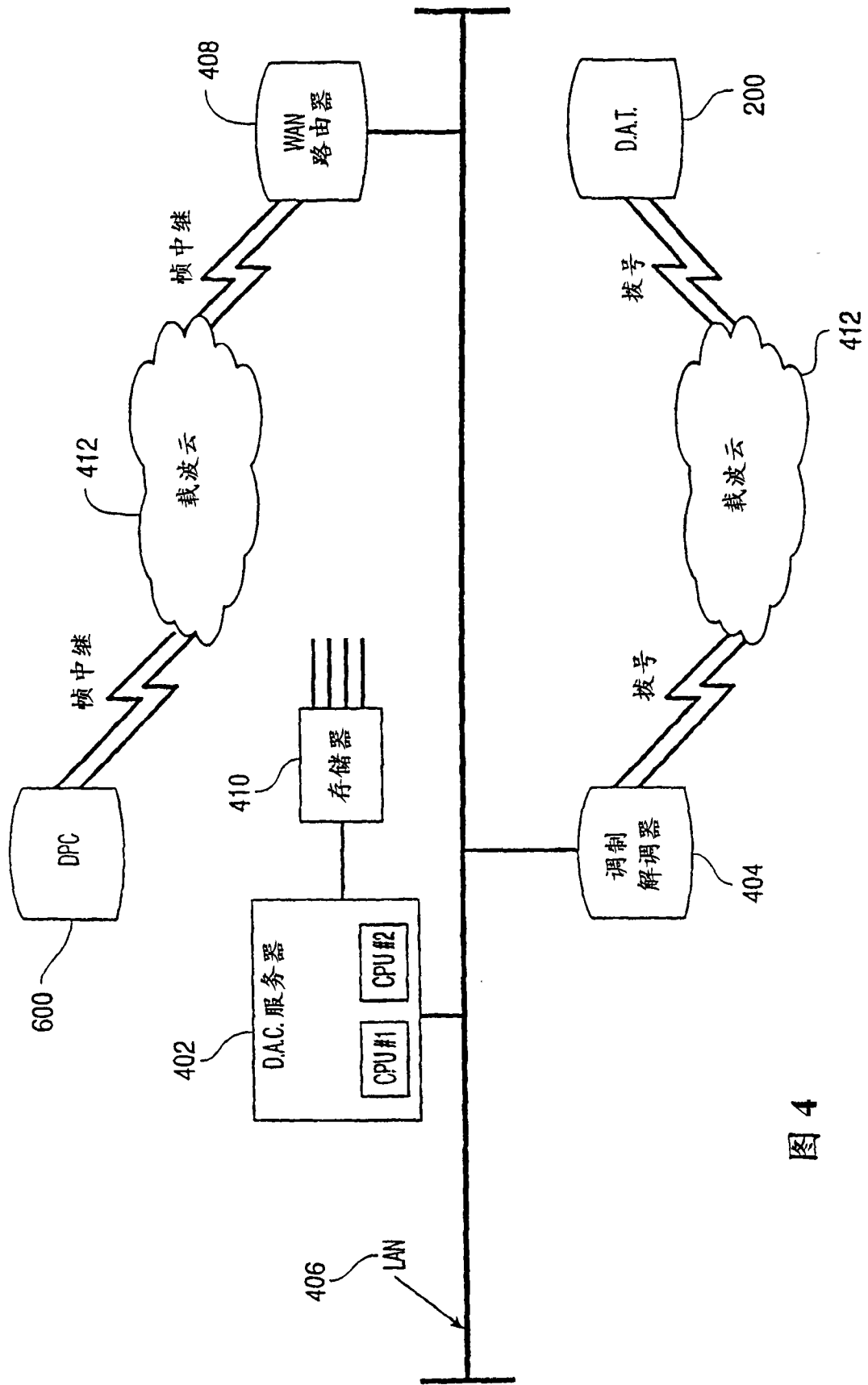


图 4

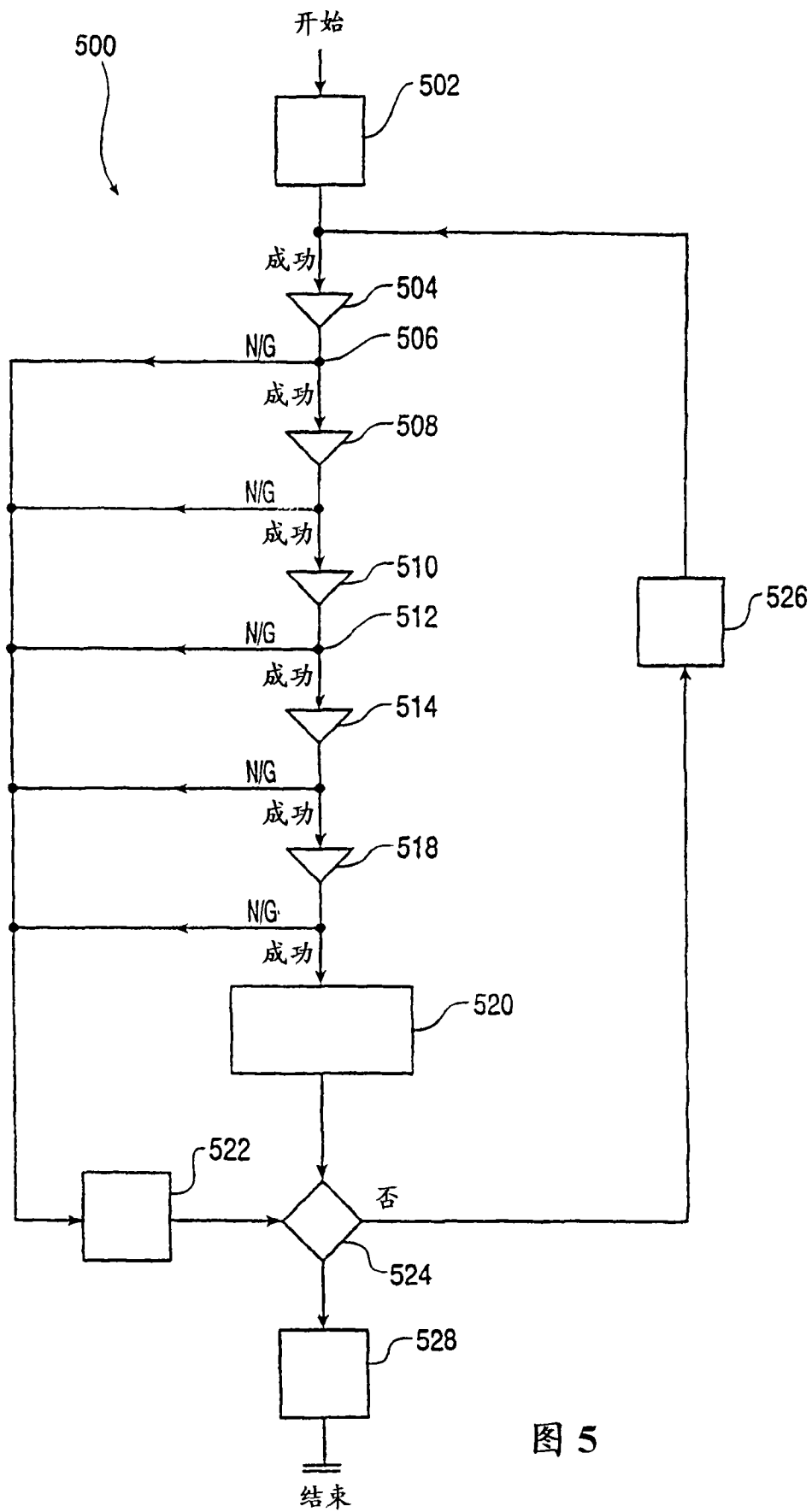


图 5

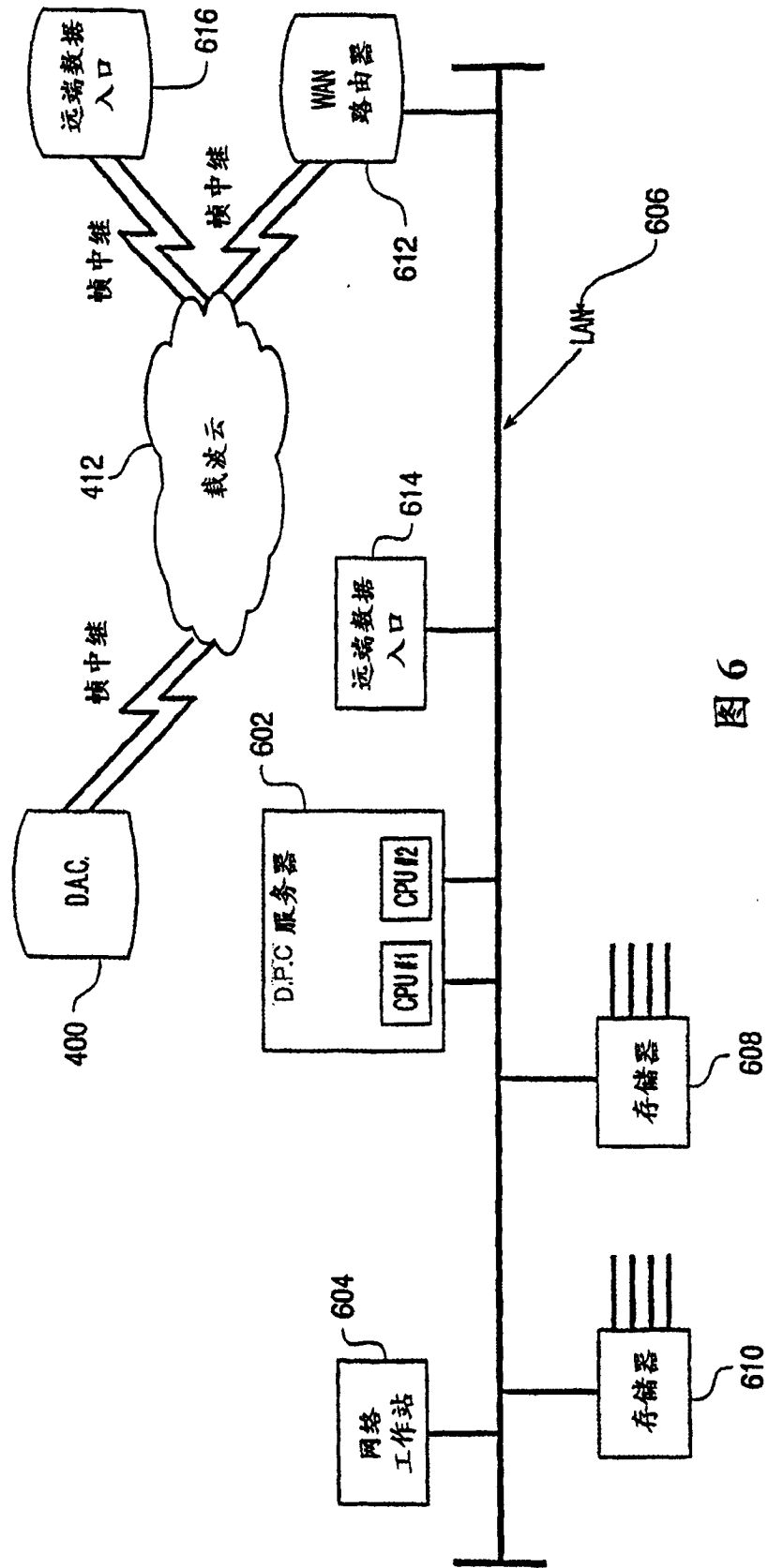


图6

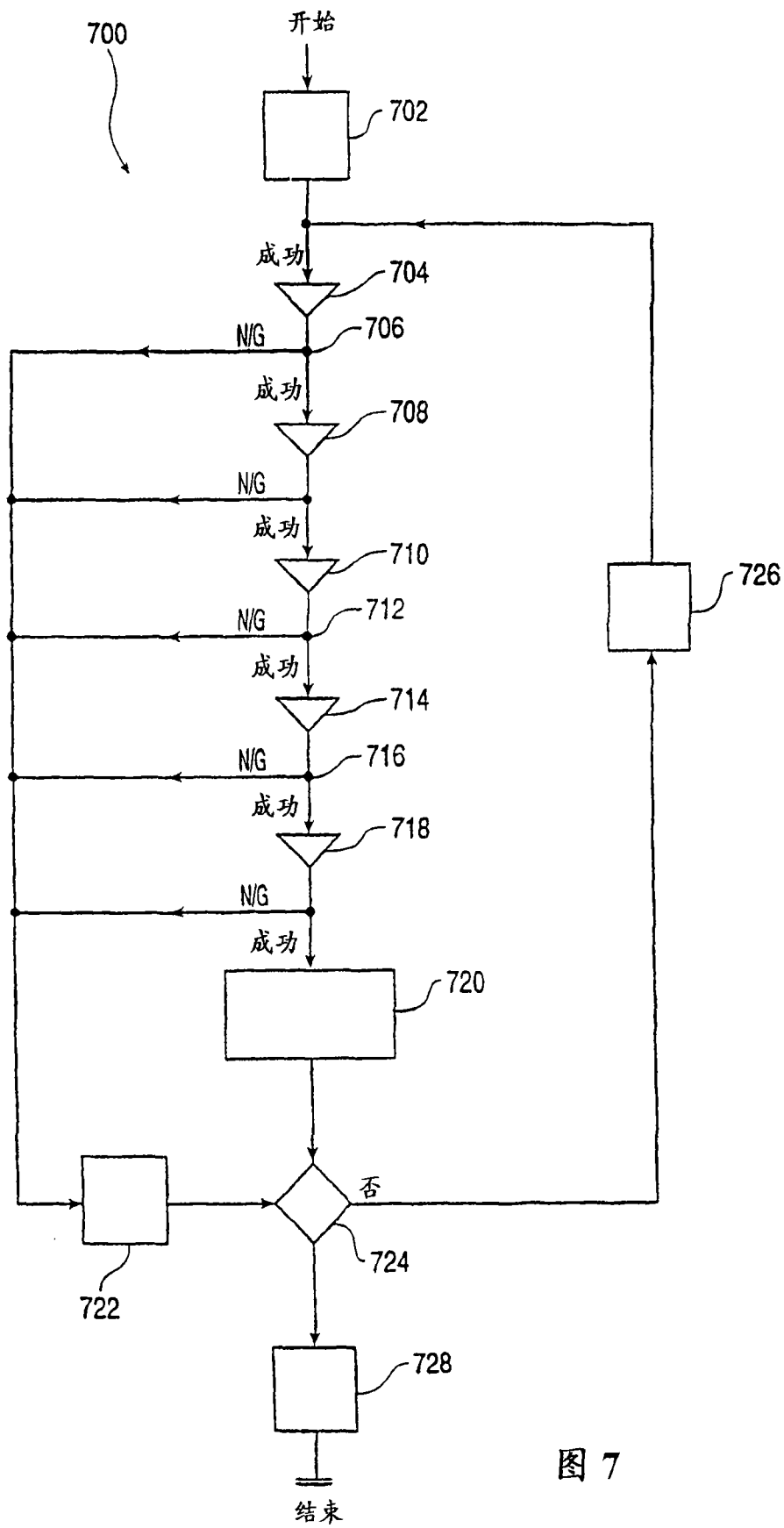


图 7

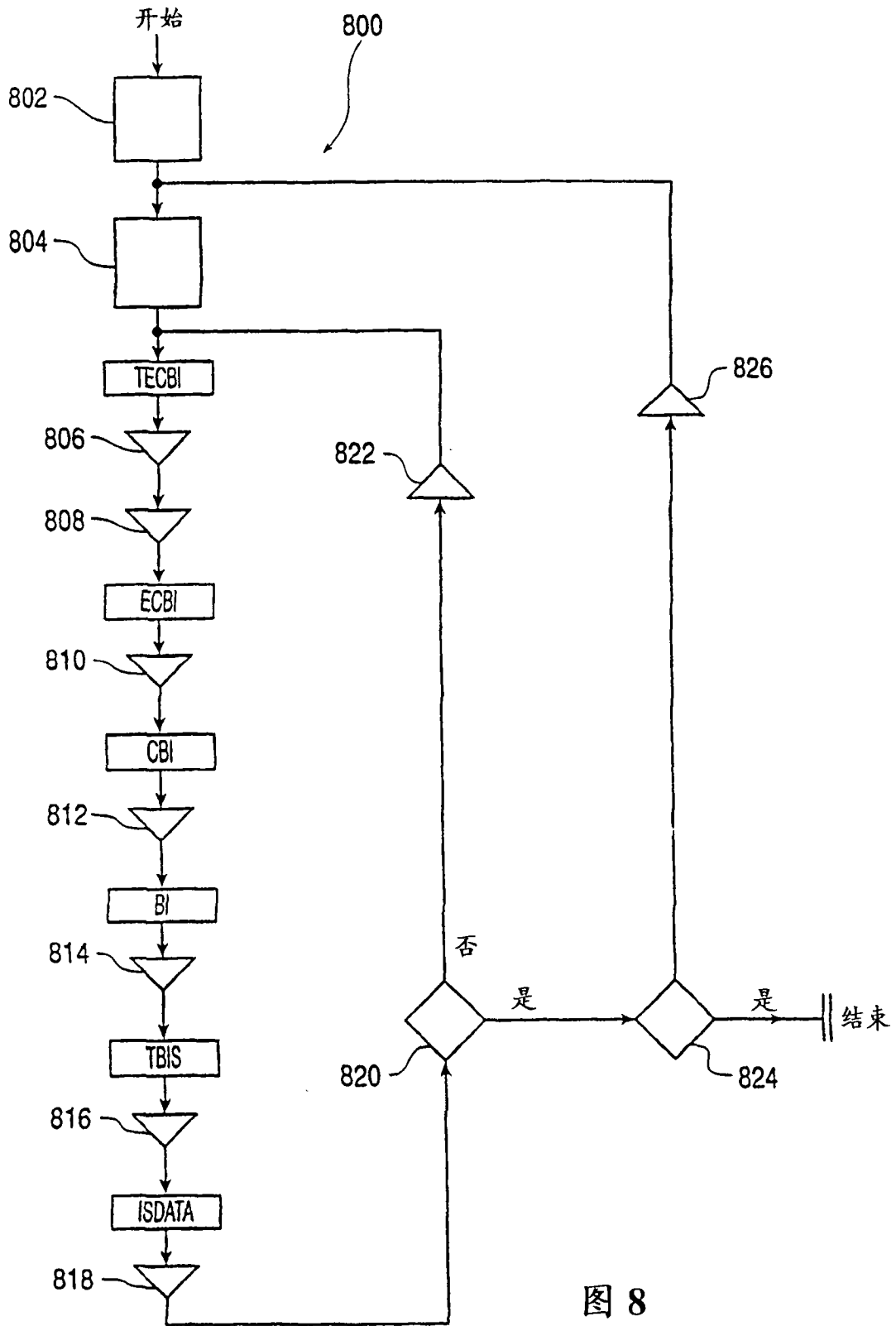


图 8

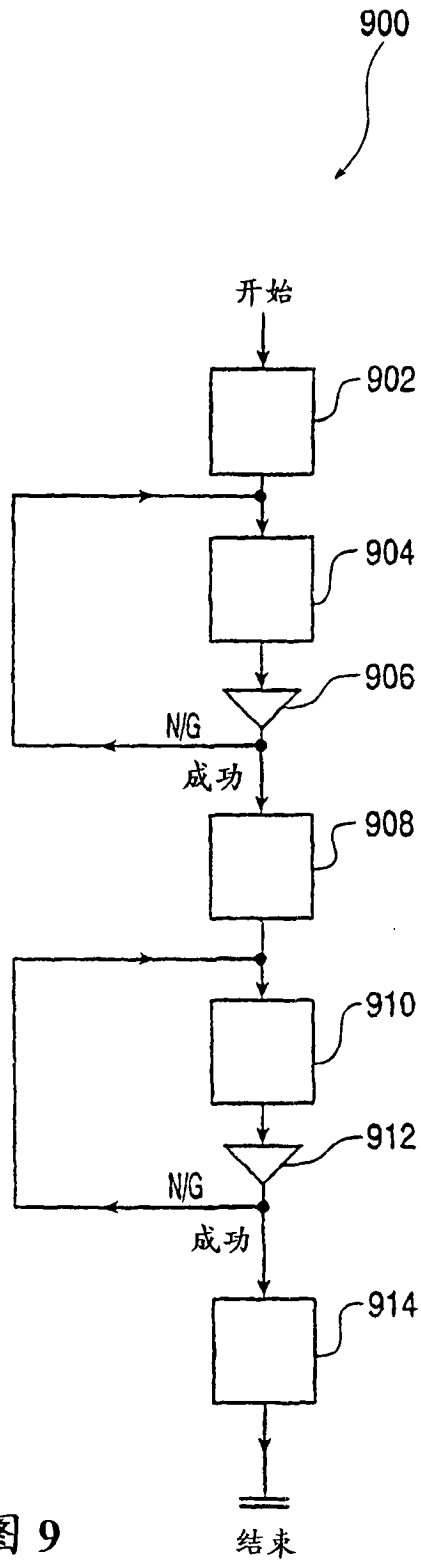


图 9

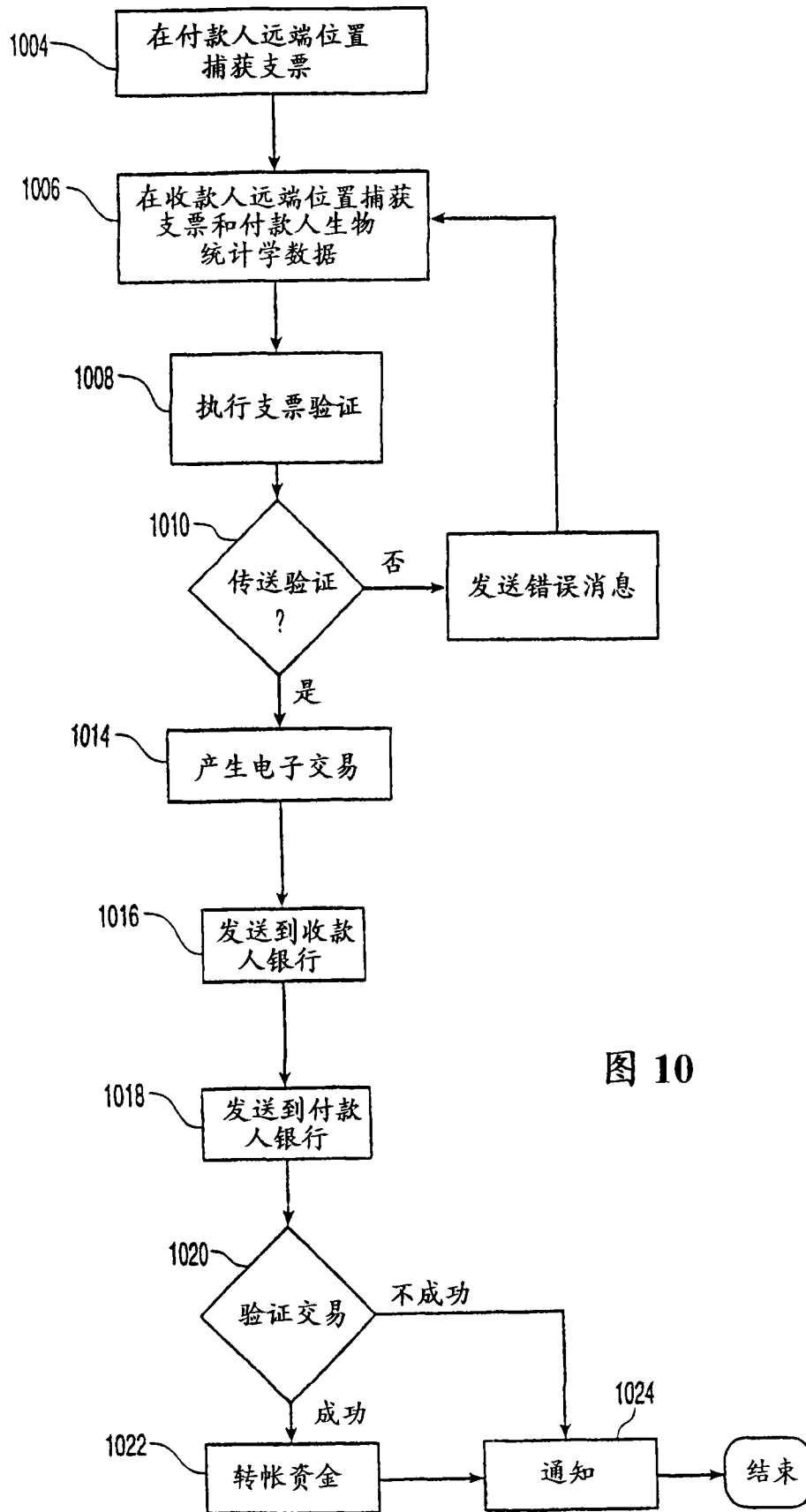


图 10