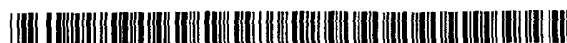


(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国际局



(43) 国际公布日:

2004年9月23日(23.09.2004)

(10) 国际公布号:

WO 2004/082222 A1

PCT

- (51) 国际分类号⁷: H04L 12/56
- (21) 国际申请号: PCT/CN2004/000192
- (22) 国际申请日: 2004年3月11日(11.03.2004)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:
03119758.2 2003年3月11日(11.03.2003) CN
- (71) 申请人(对除美国以外的所有指定国): 华为技术有限公司(HUAWEI TECHNOLOGIES CO., LTD.) [CN/CN]; 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。
- (72) 发明人;及
- (75) 发明人/申请人(仅对美国): 闫长江(YAN, Changjiang) [CN/CN]; 马达(MA, Da) [CN/CN]; 雷文阳(LEI, Wenyang) [CN/CN]; 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。
- (74) 代理人: 北京德琦知识产权代理有限公司(DEQI INTELLECTUAL PROPERTY LAW CORPORATION); 中国北京市海淀区花园东路10号高德大厦8层, Beijing 100083 (CN)。

- (81) 指定国(除另有指明, 要求每一种可提供的国家保护): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW
- (84) 指定国(除另有指明, 要求每一种可提供的地区保护): ARIPO(BW, GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), 欧亚专利(AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), 欧洲专利(AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI(BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG)

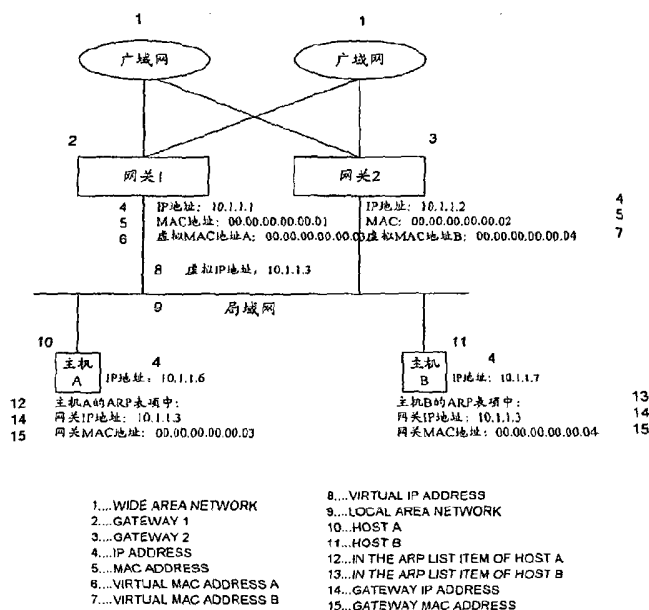
本国际公布:

- 包括国际检索报告。
- 在修改权利要求的期限届满之前进行, 在收到该修改后将重新公布。

所引用双字母代码和其它缩写符号, 请参考刊登在每期PCT公报期刊起始的“代码及缩写符号简要说明”。

(54) Title: COMMUNICATION METHOD HAVING THE FUNCTION OF PARTAKING THE NETWORK LOAD

(54) 发明名称: 具有网络负载分担功能的网络通信方法



(57) Abstract: Communication method having the function of partaking the network load The invention disclosed a network communication method that had the function of partaking the network load for the communication network which included the host of WAN (wide area network), LAN (local area network) and a plurality of gateways connecting the said WAN and LAN, the method include: configured all gateways which had the same virtual IP address and the different virtual MAC address, and configured the virtual IP address as the gateway address of each LAN host, when the LAN host transmitted message to WAN broadcast transmitting the ARP request for all gateways ; determined the gateway providing service based on the state of each current gateway, and made the response to the said ARP request using the virtual address of the gateway determined ; after the LAN host obtained the MAC address of the gateway determined, Ethernet network packaging the transmitted message using the virtual MAC address in the response, and transmitted it to the gateway which was determined providing service so as to forward . Adopting the method, not only implementing the function of partaking the network load and the redundance backup, but also resolving the problem that the host need configure a plurality of gateways.

[见续页]

WO 2004/082222 A1



(57) 摘要

本发明公开了一种具有网络负载分担功能的网络通信方法，应用于包括广域网、局域网主机和多个连接所述广域网与局域网主机的网关的通信网络，该方法包括：配置所有网关具有相同的虚拟 IP 地址和不同的虚拟 MAC 地址，并将虚拟 IP 地址配置为每个局域网主机的网关地址，当局域网主机向广域网发送报文时，向所有网关广播发送 ARP 请求；根据当前每个网关的状态确定提供服务的网关，并用被确定网关的虚拟 MAC 地址对所述 ARP 请求作出应答；局域网主机获得被确定网关的 MAC 地址后，用应答中的虚拟 MAC 地址对所发送报文进行以太网封装，并发送至被确定提供服务的网关以进行转发。采用该方法既能实现网关的负载分担和冗余备份功能，又能解决主机需要配置多个网关地址的问题。

具有网络负载分担功能的网络通信方法

技术领域

本发明涉及网络通信技术，特别涉及一种具有网络负载分担功能的网络通信方法。

5 发明背景

目前，随着基于传输控制协议/网际协议（TCP/IP）的网络通信技术的发展，人们对网络通信系统的要求也日益提高。图 1 是一个典型的网络通信系统的组成框图。如图 1 所示，该系统由包括多个主机的局域网、广域网以及用于连接局域网和广域网的网关。局域网中的主机可以通过多个网
10 关访问广域网。但是，当局域网中的主机所使用的网关发生故障时，就需要能够通过其它网关实现与广域网的通信，由此对通信系统提出了网关的负载分担和备份功能的要求。

在现有技术中，通常采用虚拟路由器冗余协议（VRRP）技术为网关提供备份功能。该技术由因特网工程任务组（IETF）的 RFC 2338 定义，它
15 针对局域网主机通过多网关访问广域网提出了冗余备份方案。该方案的主要内容是：在局域网中为主机设置网关地址为网关的虚拟 IP 地址，这些出口网关可以分为主用网关和备用网关，主用网关负责为局域网内主机提供报文转发服务。主机在通过主用网关发送数据报文的过程中，如果主用网关失效，主机可通过备用网关发送数据。主用网关与备用网关的切换与主
20 机的操作无关，主机可以不知道有多个网关存在、也不必知道是否发生了网关的切换。对主机而言，只要有网关为它提供正确的服务即可。

图 2 是现有技术中利用 VRRP 技术进行数据传输和网关切换的原理示意图。如图 2 所示，各个主机的网关地址都指向 VRRP 的虚拟 IP 地址，比

如：图 2 中各主机的网关地址都指向虚拟 IP 地址 10.1.1.3，主机通过地址解析协议（ARP）进行地址解析，即获得 IP 地址与媒体访问控制（MAC）地址之间的对应。当主机向广域网发送数据时，首先需要解析网关 IP 的 MAC 地址，这个网关 IP 就是 VRRP 备份组的虚拟 IP，于是主机发送 ARP 请求时，主用网关将应答以一个虚拟 MAC 地址。之后，主机就通过该虚拟 MAC 地址向主用网关发送数据。主用网关接收目的 MAC 为其虚拟 MAC 地址的数据并转发。当网关向主机发送数据时，如果还没有该主机的 MAC 地址，则主用网关会向该主机发送 ARP 请求，以获得该主机的 MAC 地址，然后再将数据发送至该主机。

另外，主用网关在正常工作的同时，还定期向备用网关发送 VRRP 协议报文，以使备用网关监控主用网关的状态。当备用网关连续丢失了指定个数的从主用网关发出的 VRRP 协议报文后，备用网关判断主用网关出现故障，并自动接替主用网关来提供服务，这样就完成主用、备用网关之间的切换。

在图 2 所示的上述方案中存在着一个问题，即：由于主用网关和备用网关都连接着广域网链路，但是局域网内各主机的网关 MAC 地址都是指向同一网关的 MAC 地址，这里就是指向主用网关的 MAC 地址，平时只利用了连接主用网关的广域网链路传送数据，而另外的连接备用网关的链路则是空闲的。仅当主用网关失效时，才通过备用网关传送报文数据。因此这种技术没有负载分担功能，容易导致负载过多集中于某一些链路的不均衡情况发生。为了解决这一问题，需要采用负载分担技术。VRRP 的办法是提供多个 VRRP 备份组。各 VRRP 备份组的主用网关不同，如图 3 所示，网络通信系统中配备了两个备份组，分别是 VRRP 备份组 1，其中网关 1 为主用、网关 2 为备用；以及 VRRP 备份组 2，其中网关 2 为主用、网关 1 为备用。各备份组的虚拟 IP 地址和虚拟 MAC 地址是不同的。将局域网中

的一部分主机的网关设定为 VRRP 备份组 1 的虚拟 IP 地址，另外一部分主机的网关设定为 VRRP 备份组 2 的虚拟 IP 地址。这样，上述两部分主机分别使用不同的 VRRP 备份组的虚拟 IP 地址作为网关 IP 地址，实现了对负载的统计平均分担功能。

- 5 虽然上述现有技术解决了负载分担问题，但是它仍存在着一个问题，即，在该技术中，若要实现负载的分担，就必须对同一个局域网内的主机分别设定不同网关，而这是用户通常不能接受的。另外，在上述现有技术中，当所有主机都指定网关后，网关的流量模型也就基本固定。如果需要对网关的流量模型重新进行调整，其困难相对较大。

10 发明内容

有鉴于此，本发明的目的在于提供一种具有网络负载分担功能的网络通信方法，其既能够实现网关的负载分担和冗余备份功能，又能解决主机需要配置多个网关地址的问题。

为达到上述目的，本发明的技术方案是这样实现的：

- 15 一种具有网络负载分担功能的网络通信方法，应用于包括广域网、一个以上局域网主机和至少一个用于连接所述广域网与局域网主机的网关的通信网络，所述网关被分为主用网关和备用网关，关键在于，配置所有网关具有相同的虚拟 IP 地址和不同的虚拟 MAC 地址，并将所述网关的虚拟 IP 地址配置为每个局域网主机的网关地址，该方法还包括：

- 20 1) 局域网主机向广域网发送报文时，向所有网关广播发送 ARP 请求；
2) 收到 ARP 请求后，根据当前每个网关的状态确定为发起请求的局域网主机提供服务的网关，并用被确定网关的虚拟 MAC 地址对所述 ARP 请求作出应答；
3) 所述局域网主机获得被确定为自身提供服务的网关的 MAC 地址后，

用所收到应答中的虚拟 MAC 地址对所发送报文进行以太网封装，并将封装后的报文发送至被确定提供服务的网关以进行转发。

该方法进一步包括：所述主用网关与所述备用网关按预定的时间间隔互相发送用于监控相互状态的 HELLO 报文，并实时判断是否未能连续收到从某个网关发送的指定个数的 HELLO 报文，如果是，则确定相应网关失效，否则，相应网关有效。

该方法进一步包括：判断当前失效的网关为主用网关还是备用网关，如果是备用网关，则由主用网关确定替代当前失效的备用网关的另一网关；如果是主用网关，则由当前发现主用网关失效的备用网关确定新的主用网关，然后由所确定的新主用网关确定接管当前失效的主用网关所处理的主机流量的另一网关。

该方法进一步包括：对所述失效的网关设置检测时间，该检测时间大于所述失效网关发送 HELLO 报文的时间间隔；判断是否在所设置的检测时间内连续收到指定个数的从该失效网关发出的 HELLO 报文，如果是，则确定该失效网关恢复正常，否则，确定该失效网关未恢复正常。如果确定所检测的失效网关未恢复正常，该方法进一步包括：延长所设置的检测时间；判断是否在所设置的检测时间内连续收到指定个数的从该失效网关发出的 HELLO 报文，如果是，则确定该失效网关恢复正常，并将所设置的检测时间恢复到初始设置的检测时间。

该方法进一步包括：所述局域网主机下线后，为下线的局域网主机提供服务的网关在 HELLO 报文中携带当前处理的主机个数的信息，通知其他网关；主用网关收到该 HELLO 报文信息后，根据局域网主机下线情况调整每个网关主机的分配。

新加入网关时，该方法进一步包括：

a. 所述新加入的网关向局域网内其它网关发送 HELLO 报文；

b. 所述其它网关正确检测到该新加入网关后，所述其它网关相应地更新自身的相关状态；

c. 所述主用网关在从该新加入网关接收到指定个数的 HELLO 报文后，开始对该新加入网关进行业务调度。

5 上述方案中，步骤 2) 中所述为发起请求的局域网主机确定提供服务的网关是由所述主用网关、或由所述局域网内设置的控制单元完成。步骤 2) 中所述应答由所述主用网关、所述被确定提供服务的网关、其它在所述局域网内的网关、或在所述局域网内的所述控制单元执行。

其中，所述提供服务的网关根据网关的出口带宽、或根据具有相同特
10 征的主机归相同的网关服务、或根据出口带宽负载或局域网入口的带宽负载、或根据各个网关平均分配主机的方式、或根据通过静态配置的指定的主机所指定的网关、或根据上述条件的任意组合来确定。

上述方案中，所述主用网关与备用网关是根据网关的设备优先级和设备 ID 决定的。

15 该方法进一步包括：当所述网关的优先级发生变化时，所述主用网关与备用网关采用强占方式、或采用非强占方式进行切换。其中，强占方式是指：当发生满足主用网关重新选定条件的情况时，即重新选定主用网关；非强占方式是指：在发生满足主用网关重新选定条件的情况后，一直等到主用网关失效、或人工强制要求重新选定主用网关时，才重新选定主用网
20 关。

综上所述，本发明所提供的具有网络负载分担功能的网络通信方法，能够简单地为局域网的主机通过多个网关与广域网之间的数据传送提供负载分担和冗余备份功能，这些功能的实现对局域网内的主机是透明的，局域网内主机不需要做任何特殊配置就可以充分享受负载分担和冗余备份带
25 来的好处。本发明充分利用了广域网链路的带宽和出口网关的处理能力。

附图简要说明

图 1 是一个典型的网络通信系统的组成框图；

图 2 是现有技术中利用 VRRP 技术进行数据传输和网关切换的原理图；

图 3 是现有技术中利用 VRRP 技术进行网络负载分担的原理示意图；

5 图 4 是本发明网络通信方法进行网络负载分担的实现原理示意图；

图 5 是本发明实施方案中各网关发送 HELLO 报文的示意图。

实施本发明的方式

下面结合附图及具体实施例对本发明再作进一步详细的说明。

参见图 4 详细描述本发明的具体实施方式，首先要说明的是：在本发明中，主用网关与备用网关的概念实质上与现有技术的概念是不同的。这是因为，在现有技术中主用网关是为局域网内主机提供服务的，而备用网关则处于闲置状态。而在本发明中，主用网关与备用网关的差别仅在于主用网关具有控制功能而备用网关没有，但是主用网关和备用网关都在为主机提供服务。这一点将在下面对本发明实施方案的描述中得到体现。

15 本实施方案为局域网的主机配置了相同的虚拟网关地址，这一点显然与上述现有技术不同。如图 4 所示，在本实施方案中主机的网关地址是网关的虚拟 IP 地址。设网关 1 为主用网关，其 IP 地址为 10.1.1.1，虚拟 IP 地址 10.1.1.3，MAC 地址为 00.00.00.00.00.01，虚拟 MAC 地址 A 为 00.00.00.00.00.03；网关 2 为备用网关，其 IP 地址为 10.1.1.2，虚拟 IP 地址 10.1.1.3，20 MAC 地址为 00.00.00.00.00.02，虚拟 MAC 地址 B 为 00.00.00.00.00.04；那么，主机的网关 IP 地址就设置为 10.1.1.3。在本例中，主机通过网关 1 和网关 2 两个网关完成数据报文的转发，网关 1 接收所有目的 MAC 地址为自身虚拟 MAC 地址 A，即：00.00.00.00.00.03 的报文，网关 2 负责接收处理所有目的 MAC 地址为自身虚拟 MAC 地址 B，即：00.00.00.00.00.04 的

报文。也就是说，在图 4 所示的实施例中，主机 A 的报文由网关 1 处理，主机 B 的报文由网关 2 处理。

当主机 B 要向广域网发送数据报文时，首先要向网关发送 ARP 请求。该请求为广播报文，所以网关 1 和网关 2 都能够接收到该请求。由于主机
5 B 的网关地址在本例中为网关的虚拟 IP 地址，因此，主用网关 1 接收到该请求后，根据目前各个网关的状态，决定由哪个网关为主机 B 服务，方法是使用该指定网关对应的虚拟 MAC 地址对主机 B 的 ARP 请求作出应答。作出应答的可以是主用网关也可以是其他网关，优选为使用被指定为主机 B 服务的网关作出应答。这样就可以控制主机 B 访问广域网的出口网关，
10 也就能够达到负载统计均衡的目的。

主用网关作出上述决定所依据的条件可以非常灵活，例如可以依据以下至少一种条件：根据出口带宽比来分配主机、或者把一组具有相同特征的主机，如服务器，分配到相同的网关服务、或者平均分配主机到各个网关上、也可以根据外出接口带宽负载或者局域网接口的带宽负载等策略进
15 行灵活分配。这样就可以解决局域网主机的访问广域网的数据分担问题。另一方面，这种技术对主机是透明的，而且能够根据条件灵活调度哪个网关为哪个主机服务。

假设主用网关 1 根据条件判断主机 B 应该由网关 2 来提供服务，于是就使用虚拟 MAC 地址 B: 00.00.00.00.00.04 来应答，应答网关可以是主用
20 网关或者主用网关通知网关 2 进行应答，甚至也可以由主用网关通知其他网关应答，这样主机 B 就会把报文的的目的 MAC 都封装为虚拟 MAC 地址 B: 00.00.00.00.00.04 发送。网关 2 接收到这个报文并转发。另一方面，如果是主机 A 想发送数据到广域网，同样要先发送 ARP 请求。这时，如果主用网关 1 根据条件决定应该由自己为主机 A 提供服务，则将以虚拟 MAC 地址
25 A: 00.00.00.00.00.03 对这个 ARP 应答，这样主机 1 就会把报文的的目的 MAC

都封装为虚拟 MAC 地址 A: 00.00.00.00.00.03 并发送, 这样主用网关 1 就接收到这个报文并转发。

通过上述方式, 使得局域网的主机具有相同的网关 IP 地址, 但是实际访问广域网的数据可能通过不同的网关。其根本原因在于, 根据本发明的方案, 对于局域网主机所配置的相同的虚拟网关 IP 地址, 可以提供不同的以太网硬件地址, 也就是 MAC 地址, 而这些 MAC 地址是对应于不同网关的, 以此为基础, 就形成了本发明的对局域网主机透明的负载分担方案。

当广域网向局域网的主机 A 发送的数据到达网关时, 无论网关是主用网关还是备用网关, 如果该网关没有主机 A 的 MAC 地址, 那么该网关会向局域网内发送 ARP 请求。其中 ARP 请求报文的源 IP 使用该网关自己的 IP 地址例如网关 1, 就在 ARP 请求报文内的源 IP 字段内携带: 10.1.1.1。当主机 A 收到 ARP 请求报文, 就会使用自己的 MAC 地址对这个请求作出应答, 这样该网关就可以把去往主机 A 的上述报文用主机 A 所应答的 MAC 地址进行以太网封装, 然后发送给主机 A。

在本发明方法的上述实施例中, 主用网关和备用网关同时承担转发工作。主用网关通过为主机分配不同的网关来完成流量调度。为了保证网关失效后, 由失效网关服务的主机仍然能够正常与外部网络进行数据通信, 就需要对各网关进行监控。在本发明中, 包括主用网关和备用网关在内的各网关的监控是相互的。这种监控通过相互发送 HELLO 报文来实现, 这里 HELLO 报文就作为监控报文。HELLO 报文是一种设备间的通信报文, 该报文通常使用组播发送, 以确保各个网关能够接收处理, 其主要功能在于能够使其他网关监控 HELLO 报文发送者的工作状态, 即监控其是否正常工作; 能够使其他网关了解当前 HELLO 报文发送者的相关状态信息, 比如链路占用情况等。

在本发明的一个实施例中, HELLO 报文中可包括: 设备优先级、虚拟

IP 地址、网关设备 ID、网关设备 ID 和该网关处理虚拟 MAC 地址、出口链路可用带宽、链路占用情况、认证类型、认证字、HELLO 报文时间间隔、报文校验和等。这种 HELLO 报文通常使用组播发送，以确保各个网关能够接收处理。

5 其中，优先级用于决定哪个网关作为主用网关。当主用网关失效时，各个备用网关会通过这个优先级决定谁是主用网关。虚拟 IP 地址是为局域网内提供网关的地址。虚拟 MAC 地址和设备 ID 共同说明哪个网关应该处理哪个或者哪几个虚拟 MAC 地址。链路占用情况和出口链路可用带宽为主用网关提供主机调度的依据。认证类型和认证字是出于安全考虑，用于
10 确保各个网关是合法的网关。

网关的这种 HELLO 报文会根据自己在该 HELLO 报文中携带的时间间隔信息，以该时间间隔为周期向外发送 HELLO 报文，其他网关根据是否能够收到其它网关的这种 HELLO 报文来判断其它网关是否正常工作。当
15 其他网关连续地无法从某个网关接收到指定个数的 HELLO 报文时，则判断网关故障。图 5 示意性地显示了网关之间发送这种 HELLO 报文的情况。如图 5 所示，局域网内的各网关分别向该局域网内的其它网关发送 HELLO 报文。例如网关 1 向网关 2 和 3 发送 HELLO 报文，而网关 2 向网关 1 和 3 发送 HELLO 报文，等等。

20 这样，当备用网关失效，主用网关会根据一定的条件决定失效的备用网关的虚拟 MAC 地址该由哪个网关来接管，也就是原来通过这个失效的网关提供服务的主机需要调整到其他能够正常工作的网关上提供服务。该条件可以包括：目前的各个网关的链路使用情况、各个网关的服务的主机数、用户的指定配置等。

具体地说，如果检测到备用网关失效，则其他备用网关不做任何动作，
25 主用网关则根据一定策略决定哪个网关接替失效的备用网关工作。其方法

是在 HELLO 报文中指定一种报文类型，这种类型的报文包含有如下字段：

HELLO 报文版本号、HELLO 报文类型、发送网关设备 ID、设备优先级、虚拟 IP、网关设备 ID 和该网关处理虚拟 MAC 表、出口链路可用带宽、链路占用情况、认证类型、认证字、HELLO 报文时间间隔、报文校验和等
5 字段。

主用网关通过指定的网关设备 ID 和该网关设备处理的虚拟 MAC 地址表来通知接管失效网关的网关，以接替失效网关的服务。也就是说当网关发现自己需要处理某个失效网关的报文时，那么它通过发送这样一个 ARP 请求报文的方式完成主机业务切换，该 ARP 请求报文中的发送者协议地址
10 可以是该网关的 IP 地址，但不能是虚拟 IP 地址，该 ARP 请求报文中的目标协议地址也是该网关的 IP 地址，该 ARP 请求报文的以太网封装中的源 MAC 地址是失效网关的虚拟 MAC 地址，该被指定的网关同时打开对失效网关的虚拟 MAC 地址的报文接收处理。

当主用网关失效时，首先检测到主用网关失效的备用网关将发送重新
15 选举主用网关的信号，这可以通过发送一种类型的 HELLO 报文实现。当新的主用网关确定后，新主用网关将发布失效的网关该由哪个网关接替工作的指令，其后的操作可以与上述的备用网关失效处理方法相同。以上操作就完成了故障网关的切换工作。

具体地说，如果某一备用网关最先检测到主用网关失效，该备用网关
20 应立即发出 HELLO 报文。该报文的类型为选举主用网关的报文类型，可包含如下字段：HELLO 报文版本号、HELLO 报文类型、发送网关设备 ID、设备优先级、虚拟 IP、网关设备 ID 和该网关处理虚拟 MAC 表、出口链路可用带宽、链路占用情况、认证类型、认证字、HELLO 报文时间间隔、报文校验和等字段。

25 其他备用网关收到这种选举主用网关类型的 HELLO 报文，即发送选

举主用网关类型的报文以重新选举主用网关。例如，图 4 中的网关 2 此时会成为主用网关。然后，主用网关根据策略决定谁来接管原有的主用网关处理的那些主机的流量。

下面说明主用网关的选定过程：

5 在本发明的一个实施方案中，主用网关是根据网关设备的优先级来确定的。例如，假定在某个局域网中有三个网关，其设备优先级分别是 100、80 和 60，并且设备优先级数值越大，该设备优先级越高，当然也可以采用相反设定。同时在该例子中假定选举主用网关的条件是：以设备优先级最高的网关为主用网关，如果设备优先级相同则比较网关设备 ID，ID 大的就
10 成为主用网关。这样当网关通过组播形式发送各自的 HELLO 报文，这个报文即标记为选举 HELLO 报文。各个网关设备都根据自己接收到的这种 HELLO 报文，进行比较，以确定主用网关。在本例中，网关 1 将成为主用网关。设备间的这种选举类型 HELLO 报文可包括以下字段：

HELLO 报文版本号、HELLO 报文类型、发送网关设备 ID、设备优先
15 级、虚拟 IP、网关设备 ID 和该网关处理虚拟 MAC 表、出口链路可用带宽、链路占用情况、认证类型、认证字、HELLO 报文时间间隔、报文校验和等
 字段。

下面更详细地说明主用网关对主机的服务网关的选择。

在本发明的实施方案中，主用网关是根据一定的策略决定使用哪个虚
20 拟 MAC 给各个主机的 ARP 请求进行应答。这种策略可以包括：1) 静态指定特定主机经过特定网关服务；2) 进行平均轮流分配，每个网关为 N/M
 个主机提供服务，其中 N 为网络内主机数，M 为网关数；3) 根据静态出口总可用带宽比例进行调度；4) 根据出口带宽占用情况进行动态调整；5)
 根据以太网入口带宽占用情况进行调度。

25 另外，当主机下线后，网关可以通过 ARP 的老化机制检测到这一情况。

网关会通过 HELLO 报文中携带的关于处理的主机个数的信息，使得其他网关能够了解各个网关目前负责处理的主机个数。以便主用网关根据这个情况调整主机的分配。

可以选择的主用、备用网关切换方式有两种，即强占式和非强占式的。

- 5 使用强占式时，当通信系统中网关设备的优先级发生变化时，可能会导致主用网关的强制切换。强制切换发生的条件是只要网关中的有设备的优先级比当前主用网关的优先级高，就要对主用网关进行重新选举。发生网关优先级变化的条件通常有在网关的设备优先级和出口可用带宽相关的情况下、在手工重新指配网关设备的优先级的情况下、或者失效的主用网关恢
- 10 复的等情况下，就可能会导致主用网关的切换。此时，需要作为新的主用网关的网关就会发送一次重新选举主用网关类型的 HELLO 报文，以便各网关开始重新选定主用网关。

- 对于非强占式的切换，虽然网关中有设备优先级比主用网关的优先级高，也不进行主用网关的强占，直到下一次主用网关选举发生。此时，各
- 15 个网关的备用状态保持稳定，不因为优先级的变化而发出重新选举主用网关类型的 HELLO 报文来重新选举主用网关，而是等到下一次重新选举主用网关。

- 例如，在失效网关恢复后，可以使用强占或者非强占的切换处理。无论采用上述哪种方式，主用网关都会把先前由该失效网关负责处理的虚拟
- 20 MAC 交还给恢复后的该失效网关处理。在本发明的一种实施方案中，对失效网关设置一检测时间。该检测时间是指在正常检测到连续个指定个数的监控报文后，才能认为该失效网关恢复的时间。该检测时间大于失效网关的网络 HELLO 报文发送的时间间隔。在网关发生失效后，将对该失效网关的检测时间加长。通过加长对失效网关的检测时间，可以避免曾经失效
- 25 网关的不稳定情况对系统通信产生影响。每次当不稳定的情况发生时，就

相应对检测时间进行延长，使得失效网关的不稳定对网络影响最小。

在本发明的实施方案中，还考虑到新加入网关的情况。在此情况下，新加入的网关首先发送 HELLO 报文，向网络中的其他网关声明该新网关的加入。其它网关相应地更新其相关的状态。同时，该新加入的网关也接收其他网关的 HELLO 报文，更新其相关状态。主用网关在检测到这个新加入的网关后，在从该新加入网关连续接收到指定个数的 HELLO 报文后，开始对它进行业务调度。主用网关检测新加入网关是否可以调度业务的过程和检测失效网关恢复的过程基本一致。

在上述这些实施方案中，也可以在局域网内单独配置控制单元，比如在服务器上配置。由该控制单元接替主用网关来选择为主机提供服务的网关。此外，显然也可以由控制单元来控制局域网内主用网关与备用网关的监控、切换等操作。在这种情况下，可以由控制单元指定局域网内的具有 ARP 功能的设备用被指定的服务网关的虚拟 MAC 地址来应答主机的 ARP 请求。而不仅限于由被指定的服务网关来应答。

总之，以上所述仅为本发明的较佳实施例而已，并不用以限制本发明，凡在本发明的精神和原则之内，所作的任何修改、等同替换、改进等，均应包含在本发明的保护范围之内。

权利要求书

1、一种具有网络负载分担功能的网络通信方法，应用于包括广域网、一个以上局域网主机和至少一个用于连接所述广域网与局域网主机的网关的通信网络，所述网关被分为主用网关和备用网关，其特征在于，配置所有网关具有相同的虚拟 IP 地址和不同的虚拟 MAC 地址，并将所述网关的虚拟 IP 地址配置为每个局域网主机的网关地址，该方法还包括：

1) 局域网主机向广域网发送报文时，向所有网关广播发送 ARP 请求；

2) 收到 ARP 请求后，根据当前每个网关的状态确定为发起请求的局域网主机提供服务的网关，并用被确定网关的虚拟 MAC 地址对所述 ARP 请求作出应答；

3) 所述局域网主机获得被确定为自身提供服务的网关的 MAC 地址后，用所收到应答中的虚拟 MAC 地址对所发送报文进行以太网封装，并将封装后的报文发送至被确定提供服务的网关以进行转发。

2、根据权利要求 1 所述的方法，其特征在于，该方法进一步包括：所述主用网关与所述备用网关按预定的时间间隔互相发送用于监控相互状态的 HELLO 报文，并实时判断是否未能连续收到从某个网关发送的指定个数的 HELLO 报文，如果是，则确定相应网关失效，否则，相应网关有效。

3、根据权利要求 2 所述的方法，其特征在于，该方法进一步包括：判断当前失效的网关为主用网关还是备用网关，如果是备用网关，则由主用网关确定替代当前失效的备用网关的另一网关；如果是主用网关，则由当前发现主用网关失效的备用网关确定新的主用网关，然后由所确定的新主用网关确定接管当前失效的主用网关所处理的主机流量的另一网关。

4、根据权利要求 2 所述的方法，其特征在于，该方法进一步包括：对所述失效的网关设置检测时间，该检测时间大于所述失效网关发送 HELLO

报文的时间间隔；判断是否在所设置的检测时间内连续收到指定个数的从该失效网关发出的 HELLO 报文，如果是，则确定该失效网关恢复正常，否则，确定该失效网关未恢复正常。

5 5、根据权利要求 4 所述的方法，其特征在于，如果确定所检测的失效网关未恢复正常，该方法进一步包括：延长所设置的检测时间；判断是否在所设置的检测时间内连续收到指定个数的从该失效网关发出的 HELLO 报文，如果是，则确定该失效网关恢复正常，并将所设置的检测时间恢复到初始设置的检测时间，否则，确定该失效网关未恢复正常。

10 6、根据权利要求 1 所述的方法，其特征在于，该方法进一步包括：所述局域网主机下线后，为下线的局域网主机提供服务的网关在 HELLO 报文中携带当前处理的主机个数的信息，通知其他网关；主用网关收到该 HELLO 报文信息后，根据局域网主机下线情况调整每个网关主机的分配。

7、根据权利要求 1 所述的方法，其特征在于，新加入网关时，该方法进一步包括：

- 15 a. 所述新加入的网关向局域网内其它网关发送 HELLO 报文；
- b. 所述其它网关正确检测到该新加入网关后，所述其它网关相应地更新自身的相关状态；
- c. 所述主用网关在从该新加入网关接收到指定个数的 HELLO 报文后，开始对该新加入网关进行业务调度。

20 8、根据权利要求 1 至 7 任一项所述的方法，其特征在于，步骤 2) 中所述为发起请求的局域网主机确定提供服务的网关是由所述主用网关、或由所述局域网内设置的控制单元完成。

25 9、根据权利要求 8 所述的方法，其特征在于，步骤 2) 中所述应答由所述主用网关、所述被确定提供服务的网关、其它在所述局域网内的网关、或在所述局域网内的所述控制单元执行。

10、根据权利要求 8 所述的方法，其特征在于，所述提供服务的网关根据网关的出口带宽、或根据具有相同特征的主机归相同的网关服务、或根据出口带宽负载或局域网入口的带宽负载、或根据各个网关平均分配主机的方式、或根据通过静态配置的指定的主机所指定的网关、或根据上述
5 条件的任意组合来确定。

11、根据权利要求 1 或 3 所述的方法，其特征在于，所述主用网关与备用网关是根据网关的设备优先级和设备 ID 决定的。

12、根据权利要求 11 所述的方法，其特征在于，该方法进一步包括：
当所述网关的优先级发生变化时，所述主用网关与备用网关采用强占方式、
10 或采用非强占方式进行切换。

1/3

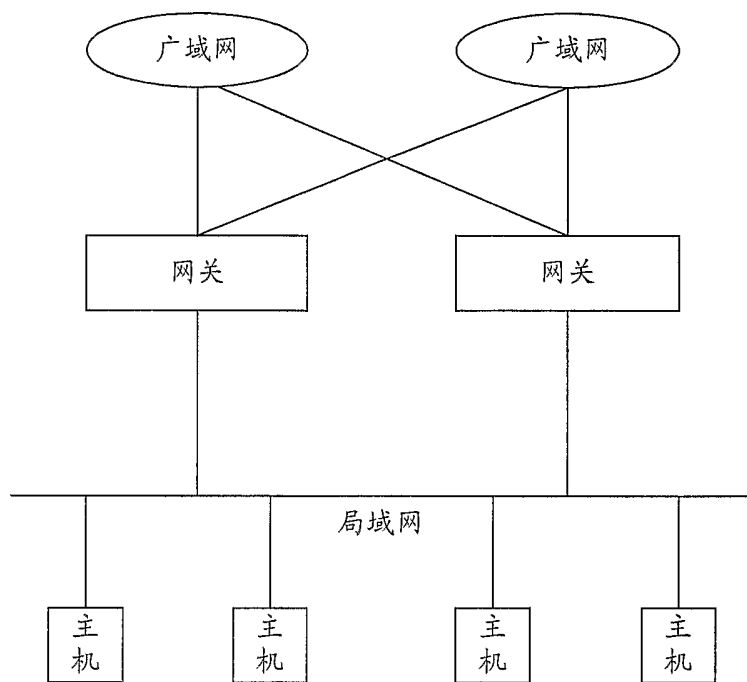


图 1

2/3

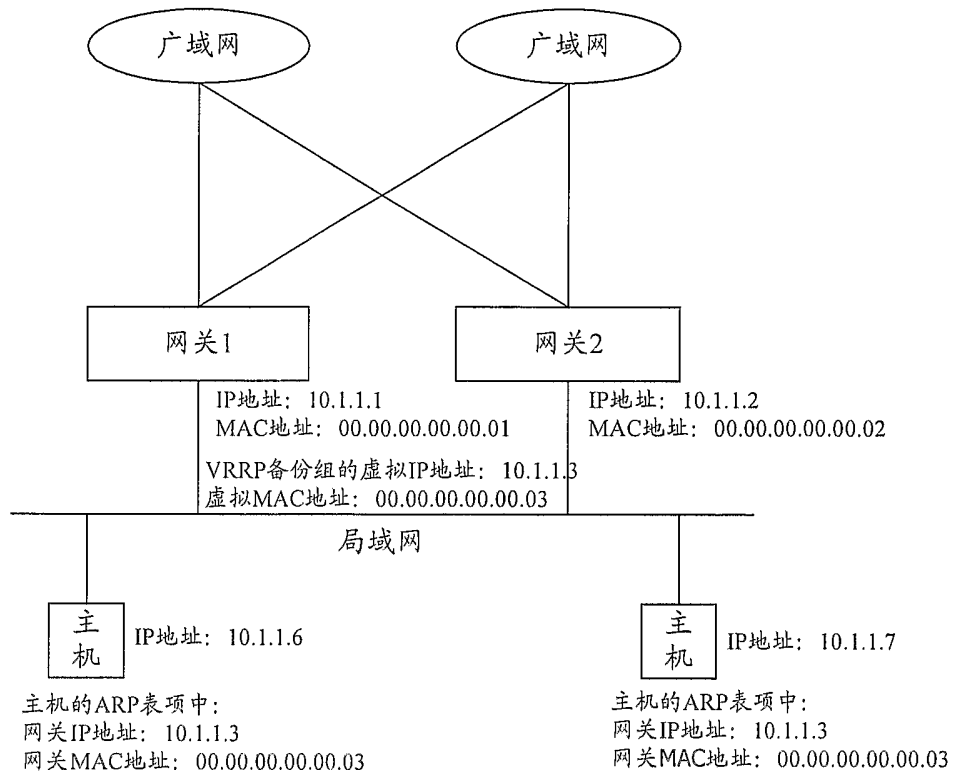


图 2

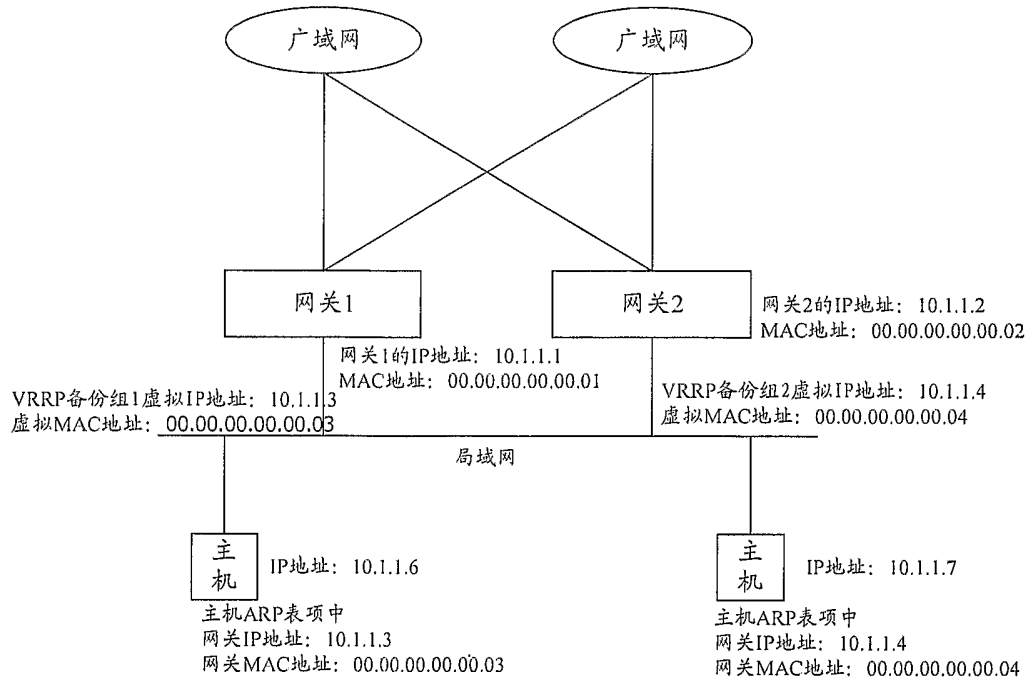


图 3

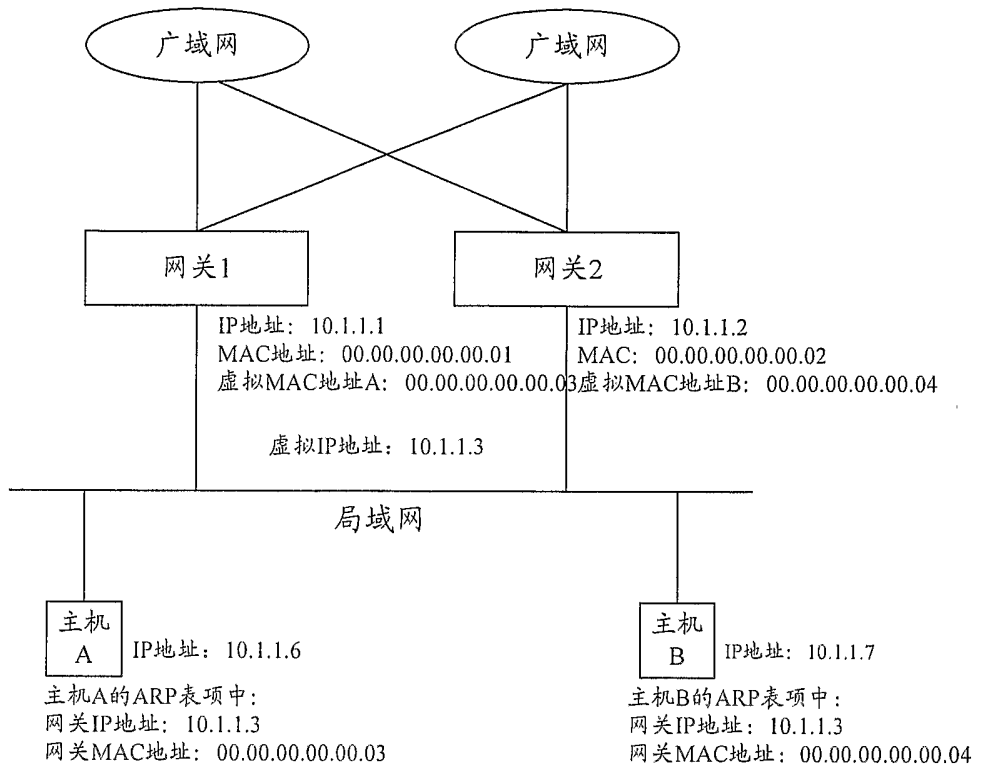


图 4

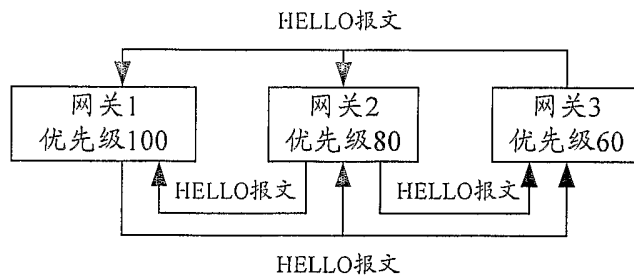
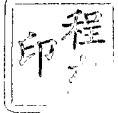


图 5

INTERNATIONAL SEARCH REPORT

International application
No.PCT/CN2004/000192

A. CLASSIFICATION OF SUBJECT MATTER <p style="text-align: center;">IPC7 H04L12/56</p> <p>According to International Patent Classification (IPC) or to both national classification and IPC</p>		
B. FIELDS SEARCHED <p>Minimum documentation searched (classification system followed by classification symbols)</p> <p style="text-align: center;">IPC⁷ H04L12/00 H04L25/00 H04Q7/00</p> <p>Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched</p> <p>Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)</p> <p style="text-align: center;">local wide area network gateway address virtual request</p>		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US5949753A, 07.Sep.1999,whole document	1-12
A	CN1357997A, 10.July.2002, whole document	1-12
A	JP2002368776A, 20.Dec.2002, whole document	1-12
A	US6282201B, 08.Aug.2001, whole document	1-12
<input type="checkbox"/> Further documents are listed in the continuation of Box C. <input checked="" type="checkbox"/> See patent family annex.		
* Special categories of cited documents: "A" document defining the general state of the art which is not considered to be of particular relevance "E" earlier application or patent but published on or after the international filing date "L" document which may throw doubts on priority claim (S) or which is cited to establish the publication date of another citation or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed	"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art "&" document member of the same patent family	
Date of the actual completion of the international search 11.Jun.2004	Date of mailing of the international search report 15 · JUL 2004 (15 · 07 · 2004)	
Name and mailing address of the ISA/CN 6 Xitucheng Rd., Jimen Bridge, Haidian District, 100088 Beijing, China Facsimile No. 86-10-62019451	Authorized officer <p style="text-align: right;">Chong Dong</p>  Telephone No. 86-010-62084524	

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2004/000192

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US5949753A	07/09/1999	JP10303968A JP2908785B2	13/11/1998 21/06/1999
CN1357997A	10/07/2002		
JP2002368776A	20/12/2002		
US6282201B	28/08/2001	US6091732A	18/07/2000

国际检索报告

国际申请号
PCT/CN2004/000192

A. 主题的分类

IPC7 H04L12/56

按照国际专利分类表(IPC)或者同时按照国家分类和 IPC 两种分类

B. 检索领域

检索的最低限度文献(标明分类体系和分类号)

IPC⁷ H04L12/00 H04L25/00 H04Q7/00

包含在检索领域中的除最低限度文献以外的检索文献

在国际检索时查阅的电子数据库(数据库的名称和, 如果实际可行的, 使用的检索词)

local wide area network gateway address virtual request IP 局域 广域 网络 网关 地址 虚拟
请求

C. 相关文件

类型*	引用文件, 必要时, 指明相关段落	相关的权利要求编号
A	US5949753A, 1999.9.07,全文	1-12
A	CN1357997A, 2002.07.10, 全文	1-12
A	JP2002368776A, 2002.12.20, 全文	1-12
A	US6282201B, 2001.8.28, 全文	1-12

其余文件在 C 栏的续页中列出。

见同族专利附件。

* 引用文件的专用类型:

“A” 明确叙述了被认为不是特别相关的一般现有技术的文件
 “E” 在国际申请日的当天或之后公布的在先的申请或专利
 “L” 可能引起对优先权要求的怀疑的文件, 为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件
 “O” 涉及口头公开、使用、展览或其他方式公开的文件
 “P” 公布日先于国际申请日但迟于所要求的优先权日的文件

“T” 在申请日或优先权日之后公布的在后文件, 它与申请不相抵触, 但是引用它是为了理解构成发明基础的理论或原理
 “X” 特别相关的文件, 仅仅考虑该文件, 权利要求所记载的发明就不能认为是新颖的或不能认为是有创造性
 “Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 权利要求记载的发明不具有创造性
 “&” 同族专利成员的文件

国际检索实际完成的日期
11.6 月 2004 (11.06.2004)

国际检索报告邮寄日期
15·7月2004 (15·07·2004)

国际检索单位名称和邮寄地址
ISA/CN
中国北京市海淀区西土城路 6 号(100088)
传真号: 86-10-62019451

授权官员
程东
电话号码: 86-10-62084524



国际检索报告
关于同族专利成员的情报

国际申请号
PCT/CN2004/000192

检索报告中引用的 专利文件	公布日期	同族专利成员	公布日期
US5949753A	1999.9.07	JP10303968A	1998.11.13
		JP2908785B2	1999.06.21
CN1357997A	2002.7.10		
JP2002368776A	2002.12.20		
US6282201B	2001.08.28	US6091732A	2000.07.18