

(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



PCT



(43) 国际公布日
2009 年 7 月 23 日 (23.07.2009)

(10) 国际公布号
WO 2009/089784 A1

(51) 国际专利分类号:
H04W 92/04 (2009.01)

(21) 国际申请号: PCT/CN2009/070067

(22) 国际申请日: 2009 年 1 月 7 日 (07.01.2009)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
200810001013.5
2008 年 1 月 10 日 (10.01.2008) CN

(71) 申请人 (对除美国外的所有指定国): 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基地总部办公楼, Guangdong 518129 (CN)。

(72) 发明人; 及

(75) 发明人/申请人 (仅对美国): 李鉴(LI, Jian) [CN/CN]; 中国广东省深圳市龙岗区坂田华为基地总部办公楼, Guangdong 518129 (CN)。 吕鸿(LV, Hong) [CN/CN]; 中国广东省深圳市龙岗区坂田华为基地总部办公楼, Guangdong 518129 (CN)。 姜玉萍(JIANG, Yuping) [CN/CN]; 中国广东省深圳市龙岗区坂田华为基地总部办公楼, Guangdong 518129 (CN)。

(74) 代理人: 北京挺立专利事务所(BEIJING TINGLI PATENT AGENCY); 中国北京市西城区宣武门西大街129号金隅大厦804、805、806室, Beijing 100031 (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, AO, AT, AU, AZ, BA, BB, BG,

[见续页]

(54) Title: METHOD?SYSTEM AND EQUIPMENT FOR ACCESSING NETWORK DEVICE TO PACKET EXCHANGE NETWORK

(54) 发明名称: 网络设备接入分组交换网络的方法、系统及装置

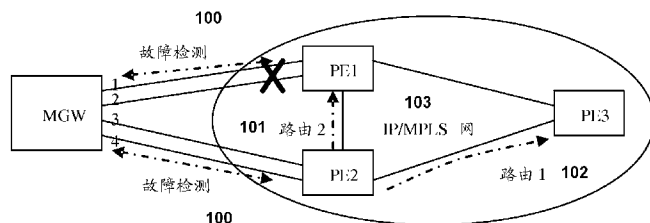


图 4 /Fig.4

100 FAULT DETECTION
101 ROUTE 2
102 ROUTE 1
103 IP/MPLS NETWORK

(57) Abstract: A method, system and equipment for accessing network device to packet exchange network are provided, the method applied to the network device which connects to the PE through the primary and secondary mode, in which the network device accesses to the system of packet exchange network, the method includes: the primary PE and the secondary PE sending a fault detection message to the network device through the interface connecting to the network device, respectively; the primary PE determining the received fault detection response from the interface during the preset time, then setting the interface "UP", and issuing route to a remote terminal PE; otherwise setting the interface "DOWN"; and cancelling

route to a remote terminal PE; the secondary PE receiving the fault detection response from the interface connecting to the network device itself, then setting the interface "UP", and issuing route to a remote terminal PE.

(57) 摘要:

本发明公开了一种网络设备接入到分组交换网络的方法、系统及装置, 所述方法应用于网络设备, 该设备通过主备方式连接提供商边界路由器接入到分组交换网络的系统中, 包括: 主 PE 和备用 PE 分别通过连接网络设备的接口, 向网络设备发送故障检测报文; 主 PE 判定在预设的时间段内, 从该接口接收到故障检测应答, 则将该接口置为 up, 向远端 PE 发布路由; 否则将该接口置为 down; 撤销向远端 PE 发布的路由; 备用 PE 从自身连接网络设备的接口接收到故障检测应答, 则将该接口置为 up, 并向远端 PE 发布路由。

WO 2009/089784 A1



BH, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW。

SD, SL, SZ, TZ, UG, ZM, ZW), 欧亚 (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), 欧洲 (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG)。

(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, NA,

本国际公布:

— 包括国际检索报告。

网络设备接入分组交换网络的方法、系统及装置

本申请要求于 2008 年 1 月 10 日提交中国专利局，申请号为 200810001013.5，发明名称为“网络设备接入分组交换网络的方法、
5 系统及装置”的中国专利申请的优先权，其全部内容通过引用结合在本申请中。

技术领域

本发明实施例涉及网络技术领域，具体涉及一种网络设备接入分
10 组交换网络，例如网际协议（IP，Internet Protocol）或多协议标签交换（MPLS，Multi-Protocol Label Switching）网络（下文中称作 IP/MPLS 网络）的方法、系统及装置。

背景技术

15 当前，网络发展日新月异，在任何场景下，都力求更好的降低设备复杂度，减少设备成本和其管理成本，同时提高故障切换时的业务收敛速度。

现有技术中下一代网络（NGN，Next Generation Network）通过主备方式直连接入 IP/MPLS 网络的网络提供商边界路由器（PE，
20 Provider Router）设备上。如下就以 NGN 设备中的多媒体网关（MGW，Media Gateway）为例，说明现有技术中 NGN 设备以主备方式接入 IP/MPLS 网络的具体处理过程。

图 1 为现有技术中 MGW 接入 IP/MPLS 网络的示意图。如图 1 所示，MGW 工作在主备方式，与 IP/MPLS 网络的两个 PE 设备：PE1
25 和 PE2 直连。MGW 上的主端口通过主链路 1 和 2 与 PE1 连接，备用端口通过备用链路 3 和 4 与 PE2 连接，链路 3 和 4 分别作为链路 1 和 2 的备用链路。每条主链路和其对应的备用链路配置相同的 IP 地址。正常情况下，MGW 的备用端口不工作，即不收发报文，因此备用链路不进行数据流的收发处理。MGW 主端口通过主链路，周期性

地给 PE1 发送地址解析协议 (ARP, Address Resolution Protocol) 请求报文, PE1 收到之后会回送 ARP 应答报文。如果 MGW 在预定的时间内没有接收到 PE1 回复的 ARP 应答报文时, 则确定主链路发生故障, 继而触发主备切换, 将备用端口切换为主端口, 则备用链路变为 5 主链路。

PE 设备上运行虚拟路由容错协议 (VRRP, Virtual Router Redundancy Protocol) 和虚拟专用局域网网段 (VPLS, Virtual Private LAN Segment)。VRRP 将 PE1 配置为主设备, PE2 配置为备用设备, 将 VRRP 的虚拟 IP 地址配置在主设备 PE1 上。两个 PE 设备需要内 10 置环回板, 即: 主设备 PE1 运行 VRRP 的物理端口, 周期性地发送 VRRP 组播报文, 该 VRRP 组播报文通过其间的 VPLS 承载扩散到备用设备 PE2 运行 VRRP 的物理端口。如果备用设备在三个 VRRP 组播报文发送周期内都没有接收到 VRRP 组播报文, 则确认为主设备发生故障, 触发 VRRP 的主备切换, 升级为主设备。PE1 和 PE2 处于 15 同一子网, 分别向远端的 PE3 发布路由, 如图 1 中虚线所示。

正常情况下, MGW 通过主端口及连接主端口的主链路, 将流量转发到 PE1, 再由 PE1 通过 IP/MPLS 网络转发至 PE3, 或者是其它 PE 设备。

对于由 PE3 经 PE1 发布的路由到达 PE1 的部分返程流量, 经主 20 链路直接转发至 MGW。如图 1 中双向箭头所示。

对于由 PE3 经 PE2 发布的路由到达 PE2 的部分返程流量, 因为 PE2 连接的 MGW 备用接口不能收发流量, 故此, 需要经 PE2 和 PE1 之间的 VPLS 网络透传至 PE1, 再通过主链路转发至 MGW。如图 1 中单向箭头所示。

25 当 MGW 检测到主链路出现故障时, 例如主链路 1 出现故障, MGW 触发主备切换, 连接主链路 1 的主端口切换为备用端口, 原主链路变为备用链路, 不再收发数据报文; 而连接备用链路 3 的备用端口切换为主端口, 原备用链路变成主链路, 开始收发报文。此时, 由于 PE1 和 PE2 设备仍然正常工作, 所以 VRRP 不会触发主备切换。

图 2 为现有技术中主链路 1 出现故障时 MGW 接入 IP/MPLS 网络的示意图。如图 2 所示, MGW 通过主端口及链路 3, 将流量发送到 PE2, 因为 VRRP 并没有由于 MGW 的主备切换而触发主备切换, PE2 仍然充当备用设备, 因此 PE2 将接收到的流量经其 L3VPN 端口, 通过 VPLS 网络透传至 PE1 的 L3VPN 端口, 再由 PE1 转发至 PE3 或其他 PE 设备。

对于由 PE3 经 PE2 发布的路由到达 PE2 的部分返程流量, 经切换后的主链路及切换后的主端口, 转发给 MGW, 如图 2 中双向箭头所示。

对于由 PE3 经 PE1 发布的路由到达 PE1 的部分返程流量, 因为原主端口切换为备用端口, 不再收发报文, 因此需要经 PE2 和 PE1 之间的 VPLS 网络透传至 PE1, 再通过切换后的主链路及切换后的主端口, 转发给 MGW, 如图 2 中单向箭头所示。

以上所述可知, 现有技术中 MGW 通过向 PE1 发送 ARP 检测报文, 并接收 PE1 回复的 ARP 应答报文检测主链路的故障, 通过 VRRP 检测 PE 设备的故障。

如果 PE1 和 PE2 之间的物理链路出现故障, 则正常情况下, PE2 无法将从远端 PE 接收到的流量转发到 PE1; 而 MGW 主备切换之后, PE1 也无法将从远端 PE 接收到的流量转发到 PE2, 从而会造成严重的业务丢包。

在主链路出现故障时, MGW 进行主备切换, 但是由于 PE1 没有故障, PE1 和 PE2 不会进行主备切换, 所以发送的流量和返程的部分流量均需要由主设备 PE1 进行转发, 因此增加了流量的转发时间, 使得业务收敛速度慢。

而且, 在现有技术中, 网络设备通过 PE 设备接入 IP/MPLS 网络时, 需要在 PE 设备上配置环回板以运行 VRRP 和 VPLS, 因而 PE 设备复杂度高。同时, 为了提高 VPLS 的可靠性, 通常需要在 PE1 和 PE2 之间配置两条物理链路, 以确保 VPLS 网络的流量透传。因此增加了设备成本, 继而增大了整个系统的故障概率, 也就必然会增加

设备的管理成本。

对于 MGW 之外的其它网络设备,在以主备方式接入分组交换网络时,也存在以上所述的问题。

5 发明内容

本发明实施例提供一种网络设备接入分组交换网络的方法,减少故障切换时的业务丢包。

本发明实施例还提供一种网络设备接入分组交换网络的系统,减少故障切换时的业务丢包。

10 本发明实施例还提供一种 PE,减小故障切换时的业务丢包。

本发明实施例还提供一种网络设备,减少故障切换时的业务丢包。

为达到上述目的,本发明实施例的技术方案是这样实现的:

15 一种网络设备接入分组交换网络的方法,应用于所述网络设备通过以主备方式连接提供商边界路由器 PE 接入分组交换网络的系统中,该方法包括:

主 PE 和备用 PE 分别通过连接所述网络设备的接口,向所述网络设备发送故障检测报文;

20 主 PE 判定在预设的时间段内,从自身连接所述网络设备的接口接收到故障检测应答,则将自身所述接口的状态置为 up,向远端 PE 发布路由;否则将自身所述接口的状态置为 down,撤销发布给远端 PE 的路由;

备用 PE 从自身连接所述网络设备的接口接收到故障检测应答,则将自身所述接口的状态置为 up,并向远端 PE 发布路由。

25 一种网络设备接入分组交换网络的方法,应用于所述网络设备通过以主备方式连接提供商边界路由器 PE 接入分组交换网络的系统中,该方法包括:

网络设备确定在预定的时间段内从主端口接收到主 PE 发送的故

障检测报文, 则向所述主 PE 回复故障检测应答;

否则将与所述主端口对应的备用端口切换为主端口; 从切换后主端口接收备用 PE 发送的故障检测报文, 通过切换后的主端口向所述备用 PE 回复故障检测应答。

- 5 一种网络设备接入分组交换网络的系统, 所述网络设备以主备方式连接提供商边界路由器 PE 接入所述分组交换网络, 该系统包括网络设备、主 PE 和备用 PE,

所述网络设备, 用于确定在预定的时间段内从主端口接收到所述主 PE 发送的故障检测报文, 则向所述主 PE 回复故障检测应答; 否
10 则将与所述主端口对应的备用端口切换为主端口; 从切换后主端口接收所述备用 PE 发送的故障检测报文, 通过切换后的主端口向所述备用 PE 回复故障检测应答;

所述主 PE, 用于通过自身连接所述网络设备的接口, 向所述网络设备发送故障检测报文, 确定在预设的时间段内, 从自身所述接口
15 接收到故障检测应答, 则将自身所述接口的状态置为 up, 并向远端 PE 发布路由; 否则将自身所述接口的状态置为 down, 并撤销发布给远端 PE 的路由;

所述备用 PE, 用于通过连接所述网络设备的接口, 向所述网络设备发送故障检测报文; 从自身连接所述网络设备的接口接收到故障
20 检测应答, 则将自身所述接口的状态置为 up, 并向远端 PE 发布路由。

一种提供商边界路由器 PE, 包括: 故障检测模块、接口状态设置模块和路由发布模块,

所述故障检测模块, 用于通过连接网络设备的接口, 向网络设备发送故障检测报文, 接收到网络设备发送的故障检测应答, 则通知所
25 述接口状态设置模块将所述接口的状态置为 up; 并通知所述路由发布模块向远端 PE 发布路由;

所述接口状态设置模块, 用于在所述故障检测模块的控制下, 将所述接口的状态置为 up;

所述路由发布模块, 用于在所述故障检测模块的控制下, 向远端

PE 发布路由。

一种网络设备，包括：主备端口切换模块和故障检测模块，

所述故障检测模块，用于确定在预定的时间段内从主端口接收到主 PE 发送的故障检测报文时，向所述主 PE 回复故障检测应答；否则，通知所述主备切换模块进行主备切换；并通过切换后的主端口接收备用 PE 发送的故障检测报文，向所述备用 PE 回复故障检测应答；

所述主备切换模块，用于在所述故障检测模块的控制下，将与所述主端口对应的备用端口切换为主端口。

与现有技术相比，本发明实施例所提供的技术方案中，分别连接网络设备主端口和备用端口的主 PE 和备用 PE 分别通过连接网络设备的接口，向网络设备发送故障检测报文，如果在预设的时间段内，从接口接收到网络设备回送的故障检测应答，则将其接口状态置为 up，同时向远端 PE 发布路由；否则将接口的状态置为 down，同时撤销向远端 PE 发布的路由。因此，本发明实施例的技术方案具有如下优点：

第一，PE 设备连接网络设备的出接口为 up 时，向远端 PE 发布路由，连接网络设备的出接口由 up 变化为 down 时，撤销向远端 PE 发布的路由，从而使得在正常情况下或在故障切换完成后，远端 PE 都能将返回的流量直接发送到连接网络设备主端口的 PE 设备，因此流量不需要在 PE 设备之间进行转发，从而减少了由于 PE 设备之间链路出现问题时造成的业务丢包。

第二，PE 设备向网络设备发送故障检测报文，由网络设备来判断主链路及 PE 设备是否出现故障，这样的检测机制代替了现有技术中的 PE 上运行的 VRRP 检测。使得在主链路发生故障，MGW 进行了主备切换时，实现了 PE 设备出接口与 MGW 主备切换的联动，并控制路由的发布。从而在发生故障时，向分组交换网络发送的流量和从远端 PE 设备返回的流量，只会通过与 MGW 主端口相连的 PE 设备转发，由此缩减了流量转发时间，提高了业务收敛速度。

第三，本发明实施例中，由 PE 设备与网络设备之间的故障检测

代替 PE 设备之间的 VRRP 检测,使得 PE 设备不需要同时运行 VRRP 和 VPLS, 因此不需要使用环回板, 从而降低了 PE 设备成本、实现复杂度以及管理成本。

5 附图说明

图 1 为现有技术中 MGW 接入 IP/MPLS 网络的示意图;

图 2 为现有技术中主链路 1 出现故障时 MGW 接入 IP/MPLS 网络的示意图;

图 3 为本发明实施例中 MGW 接入 IP/MPLS 网络的示意图;

10 图 4 为本发明实施例中主链路出现故障时 MGW 接入 IP/MPLS 网络的示意图;

图 5 为本发明实施例中网络设备接入 IP/MPLS 网络的系统结构图。

15 具体实施方式

下面结合附图及具体实施例对本发明进行详细说明。

本发明实施例中的网络设备接入分组交换网络的方法, 分别连接网络设备主端口和备用端口的主 PE 和备用 PE 分别通过连接网络设备的接口, 向网络设备发送故障检测报文。如果主 PE 在预设的时间段内, 从接口接收到网络设备回送的故障检测应答, 则将接口的状态置为 up, 向远端 PE 发布路由; 否则将接口的状态置为 down; 撤销向远端 PE 发布的路由。备用 PE 从自身连接网络设备的接口接收到故障检测应答, 则将该接口的状态置为 up, 并向远端 PE 发布路由。

20

以下以网络设备为 MGW 为例, 对本发明较佳实施例的方法进一步详细说明。

25

图 3 为本发明实施例中 MGW 接入 IP/MPLS 网络的示意图。如图 3 所示, MGW 工作在主备方式, 与 IP/MPLS 网络的两个 PE 设备: PE1 和 PE2 直连。MGW 上的主端口通过主链路 1 和 2 与 PE1 连接,

备用端口通过备用链路 3 和 4 与 PE2 连接,链路 3 和 4 分别作为主链路 1 和 2 的备用链路。MGW 的备用端口不工作,不收发报文,因此备用链路不进行数据流的收发处理。

当 MGW 和 PE1 开始正常工作时,以主链路 1 为例,PE1 通过主
5 链路 1 周期性地向 MGW 发送故障检测报文, MGW 从主链路 1 对应的主端口接收到该故障检测报文之后,回复故障检测应答。PE1 在预定的时间段内,例如故障检测的三个周期内,接收到该故障检测应答,则将自己的出接口,即自身通过主链路 1 与 MGW 主端口连接的接口置为 up 状态,同时向远端 PE 发送到达 MGW 的路由 1 (如图中虚线
10 所示)。

PE2 通过备用链路周期性地向 MGW 发送故障检测报文,由于 MGW 的备用端口不工作,因此无法接收到该故障检测报文,也就无法给 PE2 回复故障检测应答。PE2 在预定的时间段内,例如故障检测的三个周期内,没有接收到故障检测应答,则将自己的出接口,即自
15 身通过备用链路与 MGW 连接的接口置为 down,从而抑制路由的发布。

在正常工作过程中,PE1 在预定的时间段内从主链路接收到故障检测应答,则不进行处理,即保持自身与该主链路连接的出接口为 up 状态。MGW 通过主端口及主链路由 PE1 发送流量,由 PE1 转发
20 到 IP/MPLS 网络。对于从远端 PE 返回的流量,由于 PE2 没有向远端 PE 发布路由,因此远端 PE 只能通过 PE1 发布的路由 1,将报文路由到 PE1,再由 PE1 通过主链路和 MGW 的主端口,发送到 MGW。

当主链路出现故障时,例如主链路 1 出现故障时, MGW 在预定的时间段内没有从连接该主链路的主端口接收到故障检测报文,从而
25 感知到主链路故障,则触发主备切换,将原主端口切换为备用端口,将原备用端口切换为主端口,相应的,原主链路变成备用链路,原备用链路变成主链路。

图 4 为本发明实施例中主链路出现故障时 MGW 接入 IP/MPLS 网络的示意图。如图 4 所示,此时,由于 MGW 将原备用端口切换为

主端口，因此可以从切换后的主链路接收到 PE2 发送的故障检测报文，并向 PE2 回复故障检测应答；PE2 在预定的时间段内接收到该故障检测应答之后，将自己的出接口，即自身连接该切换后的主链路的接口置为 up 状态，同时向远端 PE 发布对应该切换后主链路的路由 1，
5 如图 4 中虚线所示。由于 MGW 将原来的主端口切换为备用端口，所以无法再接收到 PE1 发送的故障检测报文，也就无法回复故障检测应答，因此 PE1 在预定的时间段内，没有接收到 MGW 回复的故障检测应答，则将自己的出接口置为 down 状态，同时，撤销向远端 PE 发送的路由。

10 MGW 通过切换后的主端口向 PE2 发送流量，由 PE2 向 IP/MPLS 网络转发该流量。对于从远端 PE 返回的流量，由于 PE1 发布的路由已撤销，而 PE2 又发布了新的路由，因此远端 PE 根据 PE2 发布的路由，将返回的报文路由到 PE2，由 PE2 发送到 MGW。

其中，MGW 与 PE 之间的故障检测机制可以是地址解析协议
15 （ARP，Address Resolution Protocol）检测机制，对应的故障检测报文为 ARP 请求报文，故障检测应答为 ARP 应答报文。PE 设备只要在预定的时间段内接收到 ARP 应答报文，就会将自己的出接口，即自身连接接收 ARP 应答报文的链路的接口置为 up 状态，向远端 PE 发布路由；在预定的时间内没有接收到 ARP 应答报文，则将自己的
20 出接口置为 down，并撤销曾经向远端 PE 发布的路由。

MGW 与 PE 之间的故障检测机制也可以是双向转发检测（BFD，
Bidirectional Forwarding Detection）机制，对应的处理过程为：PE 设备向 MGW 周期性发送 BFD 控制报文，MGW 向 PE 设备回送 BFD 控制报文。当 PE 设备接收到返回的 BFD 控制报文后，经过相互协商
25 就可以创建 BFD 会话，从而将自己的出接口置为 up 状态。使用该故障检测机制时，对于出接口的状态控制，以及状态变化时发布路由的过程与使用 ARP 故障检测机制相同，在此不再赘述。

因此，本发明实施例实现了 MGW 与 PE 之间的链路故障检测与出接口联动，使得 PE 设备的主备状态与 MGW 的主备状态保持一致，

向远端 PE 发送的流量以及从远端 PE 返回的流量都可以通过与主端口对应的 PE 设备进行转发，从而减少了流量的转发时间，提高了业务收敛速度。

同时，由 PE 与 MGW 之间的 ARP 检测或 BFD 检测代替 PE 之间的 VRRP 检测，减少了 PE 感知故障的时间，使得备用 PE 可以随 MGW 的主备切换，迅速改变自己的出接口状态，以控制路由的发布。从而减少了故障切换时的业务丢包。

为了进一步减少故障切换期间的返程流量丢失，在正常情况下，PE2 可以向 PE1 发布一条出接口为 NULL0 的路由，如图 3 和图 4 中虚线所示的路由 2。正常情况下，在 PE1 的路由表中保存的路由 2 是优先级最低的路由，因此对于由远端 PE 路由到 PE1 的返程流量，PE1 不会选择路由 2。但是当出现故障时，PE1 与 MGW 之间的路由出现问题，PE1 就会选择一条优先级较低的有效路由，因此，PE1 可以通过路由 2 将返程流量转发到 PE2，等到 MGW 进行主备切换，MGW 与 PE2 之间的备用链路可用后，由 PE2 将来自 PE1 的流量转发到 MGW。因此，进一步减少了故障切换期间的业务丢包。

图 5 为本发明实施例中网络设备接入 IP/MPLS 网络的系统结构图。如图 5 所示，该系统包括网络设备 510、主 PE 520 及备用 PE 530。

其中，网络设备 510，用于确定在预定的时间段内从主端口接收到主 PE 520 发送的故障检测报文，则向主 PE 520 回复故障检测应答；否则将与该主端口对应的备用端口切换为主端口；并从切换后的主端口接收备用 PE 530 发送的故障检测报文，通过切换后的主端口向备用 PE 530 回复故障检测应答。

主 PE 520，用于通过自身连接网络设备 510 的接口，向网络设备 510 发送故障检测报文，确定在预设的时间段内，从该接口接收到故障检测应答，则将该接口的状态置为 up，并向远端 PE 发布路由；否则将该接口的状态置为 down，并撤销给远端 PE 发布的路由。

备用 PE 530，用于通过连接网络设备 510 的接口，向网络设备 510 发送故障检测报文；从该接口网络设备 510 返回的故障检测应答，

则将该接口的状态置为 up, 并向远端 PE 发布路由。

备用 PE 530 进一步用于向主 PE 520 发布 NULL0 的路由。从而使得主 PE 520 在确定主链路发生故障时, 即在预设的时间段内没有接收到网络设备 510 返回的故障检测应答时, 通过 NULL0 的路由, 5 将从远端 PE 接收的流量发送到备用 PE 530, 因此进一步减少了故障切换时的业务丢包。

具体来说, 网络设备 510 包括主备端口切换模块 512 和故障检测模块 511。故障检测模块 511, 用于确定在预定的时间段内从主端口接收到主 PE 520 发送的故障检测报文时, 向主 PE 520 回复故障检测 10 应答; 否则通知主备切换模块 512 进行主备切换; 并通过切换后的主端口向接收备用 PE 530 发送的故障报文, 并回复故障检测应答。主备切换模块 512, 用于在故障检测模块 511 的控制下, 将与该主端口对应的备用端口切换为主端口。

主 PE 520 包括故障检测模块 521、接口状态设置模块 523 和路由发布模块 522。故障检测模块 521, 用于通过连接网络设备 510 的接口, 向网络设备 510 发送故障检测报文在预定的时间段内, 接收到网络设备 510 发送的故障检测应答, 则通知接口状态设置模块 523 将该接口的状态置为 up, 并通知路由发布模块 521 向远端 PE 发布路由; 否则, 通知接口状态设置模块 523 将该接口的状态置为 down; 并通 20 知路由发布模块 521 撤销向远端 PE 发布的路由。接口状态设置模块 523, 用于在故障检测模块 521 的控制下, 将该接口状态置为 up 或置为 down。路由发布模块 522, 用于在故障检测模块 521 的控制下, 向远端 PE 发布路由, 或撤销向远端 PE 发布的路由。

相应地, 备用 PE 包括故障检测模块 531、接口状态设置模块 533 25 和路由发布模块 532。故障检测模块 531, 用于通过连接网络设备 510 的接口, 向网络设备 510 发送故障检测报文, 接收网络设备 510 发送的故障检测应答, 则通知接口状态设置模块 533 将该接口的状态设置为 up; 并通知路由发布模块 532 向远端 PE 发布路由。接口状态设置模块 533, 用于在故障检测模块 531 的控制下, 将该接口的状态设置

为 up。路由发布模块 532，用于在故障检测模块 531 的控制下，向远端 PE 发布路由。

路由发布模块 532 进一步用于向主 PE 520 发送 NULL0 的路由。

由以上所述可以看出，本发明实施例所提供的技术方案中，连接
5 网络设备主端口和备用端口的 PE 分别通过连接网络设备的接口，向网络设备发送故障检测报文，如果在预设的时间段内，从接口接收到网络设备回复的故障检测应答，则将接口的状态置为 up，向远端 PE 发布路由；否则将接口的状态置为 down，撤销向远端 PE 发布的路由。因此，本发明实施例的技术方案具有如下优点：

10 第一，PE 设备连接网络设备的出接口为 up 时，向远端 PE 发布路由，连接网络设备的出接口由 up 变化为 down 时，撤销向远端 PE 发布的路由，从而使得在正常情况下或在故障切换完成后，远端 PE 都能将返回的流量直接发送到连接网络设备主端口的 PE 设备，因此流量不需要在 PE 设备之间进行转发，从而减少了由于 PE 设备之间
15 链路出现问题时造成的业务丢包。

第二，PE 设备向网络设备发送故障检测报文，由网络设备来判断主链路及 PE 设备是否出现故障，这样的检测机制代替了现有技术中的 PE 上运行的 VRRP 检测。使得在主链路发生故障，MGW 进行了主备切换时，实现了 PE 设备出接口和发布的路由与 MGW 主备切
20 换的联动，从而在发生故障时，向分组交换网络发送的流量和从远端 PE 设备返回的流量，可以直接发送到连接网络设备主端口的 PE 设备，由该 PE 设备转发到进行转发，从而减小了流量的转发时间，提高了业务收敛速度。

第三，本发明实施例中，由 PE 设备与网络设备之间的故障检测
25 代替 PE 设备之间的 VRRP 检测，使得 PE 设备不需要同时运行 VRRP 和 VPLS，因此不需要使用环回板，从而降低了 PE 设备成本、其实现复杂度以及其管理成本。

第四，由于 VRRP 故障检测时间较长，通常为 3s，所以在主 PE 设备出现故障时，备用 PE 感知故障的时间较长，VRRP 进行主备切

换较慢，从而会在故障切换时造成严重的业务丢包。而本发明实施例中，使用 PE 设备与网络设备之间的 ARP 检测机制或 BFD 机制代替 PE 设备之间的 VRRP 检测，由于 ARP 检测机制或 BFD 机制的故障检测时间小于 VRRP 的故障检测时间，使得主 PE 设备出现故障，

5 MGW 进行主备切换时，备用 PE 马上将自己的出接口切换为 up 状态，提高了 PE 进行故障切换的速度，从而减小了故障切换时的业务丢包。而且 PE 之间不存在 VRRP 链路，也就不会出现无法正常进行主备切换的问题。

第五，本发明实施例中，备用 PE 向主 PE 发送一条 NULL0 的路由，使得在故障切换过程中，发送到主 PE 报文可以经过该 NULL0

10 的路由，转发到备用 PE，由备用 PE 在故障切换之后，转发到网络设备，从而进一步减小了故障切换过程中的业务丢包。

通过以上的实施方式的描述，本领域的技术人员可以清楚地了解到本发明可以通过硬件实现，也可以借助软件加必要的通用硬件平台的方式来实现基于这样的理解，本发明的技术方案可以以软件产品的形式体现出来，该软件产品可以存储在一个非易失性存储介质（可以是 CD-ROM，U 盘，移动硬盘等）中，包括若干指令用以使得一台计算机设备（可以是个人计算机，服务器，或者网络设备）执行

15 本发明各个实施例所述的方法。

20 总之，以上所述仅为本发明的较佳实施例而已，并非用于限定本发明的保护范围。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

1、一种网络设备接入分组交换网络的方法，应用于所述网络设备通过以主备方式连接提供商边界路由器 PE 接入分组交换网络的系
5 统中，其特征在于，该方法包括：

主 PE 和备用 PE 分别通过连接所述网络设备的接口，向所述网络设备发送故障检测报文；

主 PE 判定在预设的时间段内，从自身连接所述网络设备的接口接收到故障检测应答，则将自身所述接口的状态置为 up，向远端 PE
10 发布路由；否则将自身所述接口的状态置为 down，撤销发布给远端 PE 的路由；

备用 PE 从自身连接所述网络设备的接口接收到故障检测应答，则将自身所述接口的状态置为 up，并向远端 PE 发布路由。

2、如权利要求 1 所述的方法，其特征在于，所述故障检测报文
15 为地址解析协议 ARP 请求报文，所述故障检测应答为 ARP 应答报文；

或者，所述故障检测报文和所述故障检测应答为双向转发检测 BFD 控制报文。

3、如权利要求 1 所述的方法，其特征在于，该方法进一步包括：
连接所述网络设备备用端口的 PE 向连接所述网络设备主端口的 PE
20 发送 NULL0 的路由。

4、一种网络设备接入分组交换网络的方法，应用于所述网络设备通过以主备方式连接提供商边界路由器 PE 接入分组交换网络的系统中，其特征在于，该方法包括：

网络设备确定在预定的时间段内从主端口接收到主 PE 发送的故障检测报文，则向所述主 PE 回复故障检测应答；
25

否则将与所述主端口对应的备用端口切换为主端口；从切换后主端口接收备用 PE 发送的故障检测报文，通过切换后的主端口向所述备用 PE 回复故障检测应答。

5、如权利要求 4 所述的方法，其特征在于，所述故障检测报文

为地址解析协议 ARP 请求报文,所述故障检测应答为 ARP 应答报文;
或者,所述故障检测报文和所述故障检测应答为双向转发检测 BFD 控制报文。

6、一种网络设备接入分组交换网络的系统,所述网络设备以主备方式连接提供商边界路由器 PE 接入所述分组交换网络,其特征在于,该系统包括网络设备、主 PE 和备用 PE,

所述网络设备,用于确定在预定的时间段内从主端口接收到所述主 PE 发送的故障检测报文,则向所述主 PE 回复故障检测应答;否则将与所述主端口对应的备用端口切换为主端口;从切换后主端口接收所述备用 PE 发送的故障检测报文,通过切换后的主端口向所述备用 PE 回复故障检测应答;

所述主 PE,用于通过自身连接所述网络设备的接口,向所述网络设备发送故障检测报文,确定在预设的时间段内,从自身所述接口接收到故障检测应答,则将自身所述接口的状态置为 up,并向远端 PE 发布路由;否则将自身所述接口的状态置为 down,并撤销发布给远端 PE 的路由;

所述备用 PE,用于通过连接所述网络设备的接口,向所述网络设备发送故障检测报文;从自身连接所述网络设备的接口接收到故障检测应答,则将自身所述接口的状态置为 up,并向远端 PE 发布路由。

7、如权利要求 6 所述的系统,其特征在于,所述备用 PE 进一步用于向主 PE 发布 NULL0 的路由。

8、一种提供商边界路由器 PE,其特征在于,该 PE 包括:故障检测模块、接口状态设置模块和路由发布模块,

所述故障检测模块,用于通过连接网络设备的接口,向网络设备发送故障检测报文,接收到网络设备发送的故障检测应答,则通知所述接口状态设置模块将所述接口的状态置为 up;并通知所述路由发布模块向远端 PE 发布路由;

所述接口状态设置模块,用于在所述故障检测模块的控制下,将所述接口的状态置为 up;

所述路由发布模块，用于在所述故障检测模块的控制下，向远端 PE 发布路由。

- 9、如权利要求 8 所述的 PE，其特征在于，所述故障检测模块进一步用于在预定的时间段内没有接收到网络设备回复的故障检测应答，则通知所述接口设置模块将所述接口的状态置为 down，并通知所述路由发布模块撤销向远端 PE 发布的路由；

所述接口状态设备模块进一步用于在所述故障检测模块的控制下，将所述接口的状态置为 down；

- 所述路由发布模块进一步用于在所述故障检测模块的控制下，撤销向远端 PE 发布的路由。

10、如权利要求 8 所述的 PE，其特征在于，所述路由发布模块进一步用于向对应的主 PE 发送 NULL0 的路由。

11、一种网络设备，其特征在于，所述网络设备包括：主备端口切换模块和故障检测模块，

- 所述故障检测模块，用于确定在预定的时间段内从主端口接收到主 PE 发送的故障检测报文时，向所述主 PE 回复故障检测应答；否则，通知所述主备切换模块进行主备切换；并通过切换后的主端口接收备用 PE 发送的故障检测报文，向所述备用 PE 回复故障检测应答；

- 所述主备切换模块，用于在所述故障检测模块的控制下，将与所述主端口对应的备用端口切换为主端口。

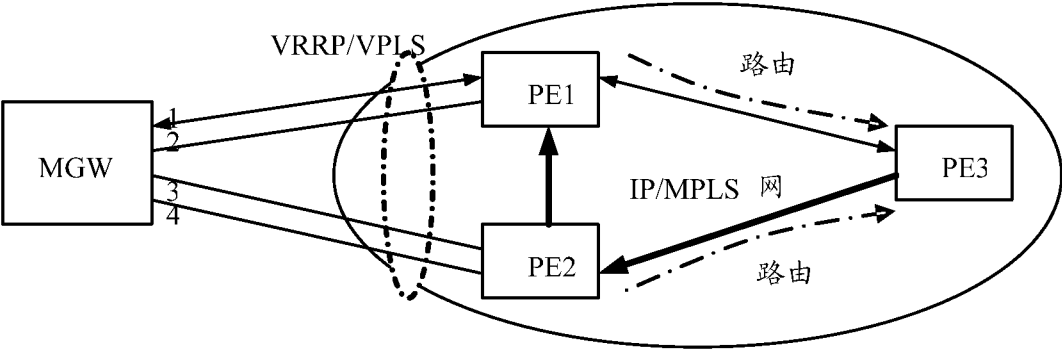


图 1

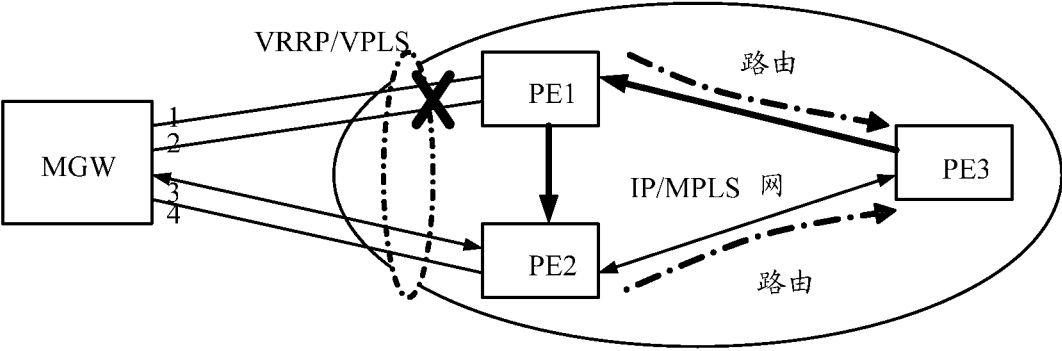


图 2

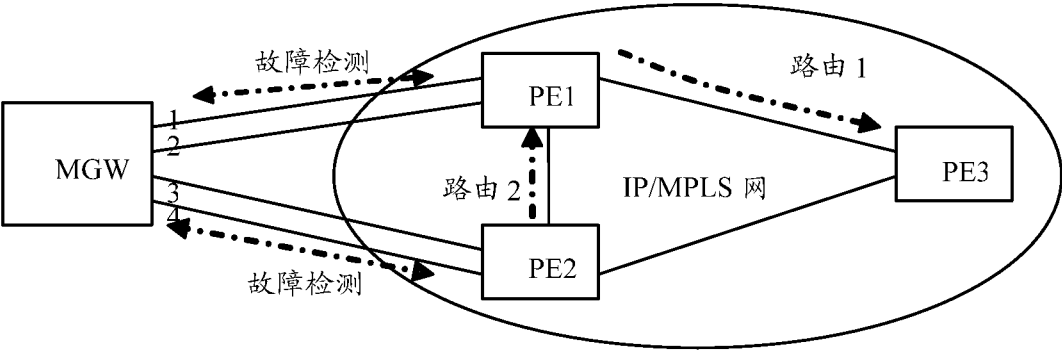


图 3

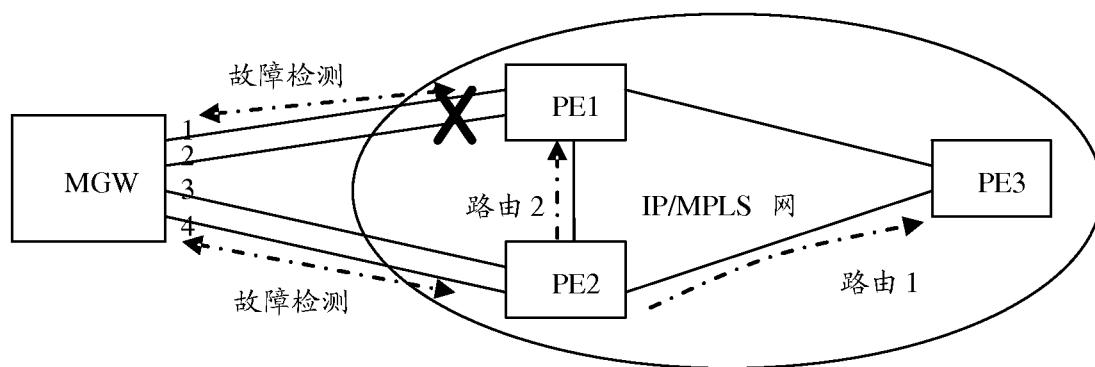


图 4

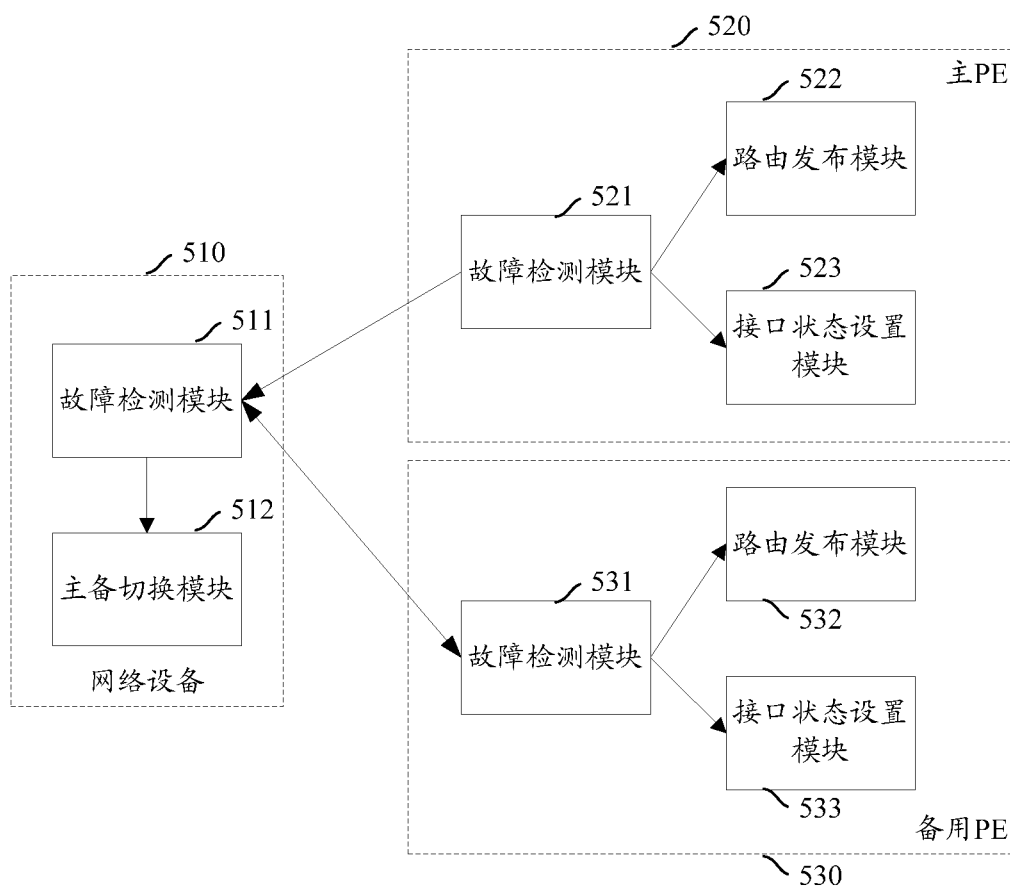


图 5

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2009/070067

A. CLASSIFICATION OF SUBJECT MATTER

H04W 92/04(2009.01)i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC:H04W 92/-; H04L 12/-

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

CPRS;CNKI;WPI;EPODOC;PAJ; network device packet exchange access router fault detection interface
message primary secondary

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	CN1303557A(ERICSSON TELEFON AB L M) 11 Jul.2001(11.07.2001) The whole document	1-11
A	CN1391749A(ERICSSON TELEFON AB L M)15 Jan.2003(15.01.2003) The whole document	1-11

☐ Further documents are listed in the continuation of Box C.☒ See patent family annex.

* Special categories of cited documents:	“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
“A” document defining the general state of the art which is not considered to be of particular relevance	“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
“E” earlier application or patent but published on or after the international filing date	“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
“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)	“&”document member of the same patent family
“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	

Date of the actual completion of the international search
26 Mar.2009 (26.03.2009)Date of mailing of the international search report
16 Apr. 2009 (16.04.2009)Name and mailing address of the ISA/CN
The State Intellectual Property Office, the P.R.China
6 Xitucheng Rd., Jimen Bridge, Haidian District, Beijing, China
100088
Facsimile No. 86-10-62019451Authorized officer
HUANG, Yuan
Telephone No. (86-10)62411769

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No.

PCT/CN2009/070067

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
CN1303557A	11.07.2001	EP0948168A1	06.10.1999
		WO9950998A1	07.10.1999
		AU3602999A	18.10.1999
		EP1070407A1	24.01.2001
		JP2002510904T	09.04.2002
		AU752032B2	05.09.2002
		US6754228B1	22.06.2004
		EP1070407B1	24.11.2004
		DE69922180E	30.12.2004
		IL138303 A	25.09.2005
		DE69922180T2	01.12.2005
		CN1149793C	12.05.2004
		CA2326488C	12.08.2008
CN1391749A	15.01.2003	WO0137508A2	25.05.2001
		AU1563601A	30.05.2001
		EP1230762A2	14.08.2002
		KR20020067517A	22.08.2002
		TW522679B	01.03.2003
		CN1163026C	18.08.2004
		INMUMNP200200437E	16.06.2006
		CA2400999A1	25.05.2001

国际检索报告

国际申请号
PCT/CN2009/070067

A. 主题的分类

H04W 92/04(2009.01)i

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(标明分类系统和分类号)

IPC:H04W 92/-; H04L 12/-

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, 和使用的检索词(如使用))

CPRS;CNKI;WPI;EPODOC;PAJ;网络设备 分组交换 接入 路由器 故障检测 接口 报文 network device packet exchange access router fault detection interface message primary secondary

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
A	CN1303557A(艾利森电话股份有限公司) 11.7 月 2001(11.07.2001) 全文	1-11
A	CN1391749A(艾利森电话股份有限公司)15.1 月 2003(15.01.2003) 全文	1-11

☐ 其余文件在 C 栏的续页中列出。

☒ 见同族专利附件。

* 引用文件的具体类型:

“A” 认为不特别相关的表示了现有技术一般状态的文件

“E” 在国际申请日的当天或之后公布的在先申请或专利

“L” 可能对优先权要求构成怀疑的文件, 或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, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件

“X” 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性

“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, 要求保护的发明不具有创造性

“&” 同族专利的文件

国际检索实际完成的日期

26.3 月 2009 (26.03.2009)

国际检索报告邮寄日期

16.4 月 2009 (16.04.2009)

中华人民共和国国家知识产权局(ISA/CN)

中国北京市海淀区蓟门桥西土城路 6 号 100088

传真号: (86-10)62019451

受权官员

黄渊

电话号码: (86-10) **62411769**

国际检索报告
关于同族专利的信息

国际申请号
PCT/CN2009/070067

检索报告中引用的 专利文件	公布日期	同族专利	公布日期
CN1303557A	11.07.2001	EP0948168A1	06.10.1999
		WO9950998A1	07.10.1999
		AU3602999A	18.10.1999
		EP1070407A1	24.01.2001
		JP2002510904T	09.04.2002
		AU752032B2	05.09.2002
		US6754228B1	22.06.2004
		EP1070407B1	24.11.2004
		DE69922180E	30.12.2004
		IL138303 A	25.09.2005
		DE69922180T2	01.12.2005
		CN1149793C	12.05.2004
		CA2326488C	12.08.2008
		WO0137508A2	25.05.2001
		AU1563601A	30.05.2001
CN1391749A	15.01.2003	EP1230762A2	14.08.2002
		KR20020067517A	22.08.2002
		TW522679B	01.03.2003
		CN1163026C	18.08.2004
		INMUMNP200200437E	16.06.2006
		CA2400999A1	25.05.2001