

(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
2017 年 7 月 6 日 (06.07.2017)



(10) 国际公布号
WO 2017/114378 A1

- (51) 国际专利分类号:
H04L 12/851 (2013.01)
- (21) 国际申请号: PCT/CN2016/112270
- (22) 国际申请日: 2016 年 12 月 27 日 (27.12.2016)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:
201511005543.3 2015 年 12 月 28 日 (28.12.2015) CN
- (71) 申请人: 新华三技术有限公司 (NEW H3C TECHNOLOGIES CO., LTD) [CN/CN]; 中国浙江省杭州市滨江区长河路 466 号, Zhejiang 310052 (CN)。
- (72) 发明人: 赵丽娜 (ZHAO, Lina); 中国北京市海淀区上地信息产业基地创业路 2 号东方电子大厦 211 室, Beijing 100085 (CN)。 李杰 (LI, Jie); 中国北京市海淀区上地信息产业基地创业路 2 号东方电子大厦 211 室, Beijing 100085 (CN)。 柳佳佳 (LIU, Jiajia); 中国北京市海淀区上地信息产业基地创业路 2 号东方电子大厦 211 室, Beijing 100085 (CN)。
- (74) 代理人: 北京博思佳知识产权代理有限公司 (BEIJING BESTIPR INTELLECTUAL PROPERTY LAW CORPORATION); 中国北京市海淀区上地三街 9 号嘉华大厦 B 座 409, Beijing 100085 (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, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, 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, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, 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,

[见续页]

(54) Title: FORWARDING OF IP DATA PACKET

(54) 发明名称: IP 数据包的转发

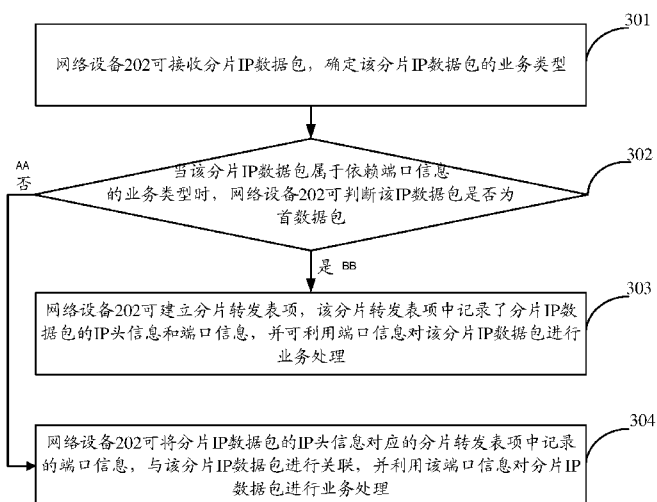


图 3

- 301 A NETWORK DEVICE 202 RECEIVES A FRAGMENTED IP DATA PACKET AND DETERMINES A SERVICE TYPE OF THE FRAGMENTED IP DATA PACKET
- 302 IF THE FRAGMENTED IP DATA PACKET BELONGS TO A SERVICE TYPE RELYING ON PORT INFORMATION, THE NETWORK DEVICE 202 DETERMINES WHETHER THE FRAGMENTED IP DATA PACKET IS A FIRST DATA PACKET
- 303 THE NETWORK DEVICE 202 ESTABLISHES A FRAGMENT FORWARDING ENTRY, IP HEADER INFORMATION AND PORT INFORMATION OF THE FRAGMENTED IP DATA PACKET BEING RECORDED IN THE FRAGMENT FORWARDING ENTRY, AND USES THE PORT INFORMATION TO PERFORM A SERVICE PROCESSING ON THE FRAGMENTED IP DATA PACKET
- 304 THE NETWORK DEVICE 202 ASSOCIATES THE PORT INFORMATION RECORDED IN THE FRAGMENT FORWARDING ENTRY CORRESPONDING TO THE IP HEADER INFORMATION OF THE FRAGMENTED IP DATA PACKET WITH THE FRAGMENTED IP DATA PACKET, AND USES THE PORT INFORMATION TO PERFORM SERVICE PROCESSING ON THE FRAGMENTED IP DATA PACKET
- AA NO
- BB YES

(57) Abstract: In an example, the method comprises: a network device receives a fragmented IP data packet and determines a service type of the fragmented IP data packet; if the fragmented IP data packet belongs to a service type relying on port information, determining whether the fragmented IP data packet is a first data packet; if so, then establishing a fragment forwarding entry, IP header information and port information of the fragmented IP data packet being recorded in the fragment forwarding entry, and using the port information to perform a service processing on the fragmented IP data packet; if not, associating the port information recorded in the fragment forwarding entry corresponding to the IP header information of the fragmented IP data packet with the fragmented IP data packet, and using the port information to perform service processing on the fragmented IP data packet.

(57) 摘要: 在示例中, 网络设备可接收分片 IP 数据包, 并确定所述分片 IP 数据包的业务类型; 当所述分片 IP 数据包属于依赖端口信息的业务类型时, 判断所述分片 IP 数据包是否为首数据包; 如果是, 则建立分片转发表项, 所述分片转发表项中记录了所述分片 IP 数据包的 IP 头信息和端口信息, 利用所述端口信息对所述分片 IP 数据包进行业务处理; 如果否, 则将所述分片 IP 数据包的 IP 头信息对应的分片转发表项中记录的端口信息, 与所述分片 IP 数据包进行关联, 并利用所述端口信息对所述分片 IP 数据包进行业务处理。

WO 2017/114378 A1



RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, **本国际公布:**

CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, — 包括国际检索报告(条约第 21 条(3))。
TG)。

IP 数据包的转发

相关申请的交叉引用

5 [01]本专利申请要求 2015 年 12 月 28 日提交的、申请号为 201511005543.3、发明名称为“一种 IP 数据包的处理方法和装置”的中国专利申请的优先权，该申请的全文以引用的方式并入本文中。

背景技术

10 [02]MTU (Maximum Transmission Unit, 最大传输单元)指数据链路层上能够传输的数据包的最大长度。在 IP (Internet protocol) 数据包经由路由寻址后要交由数据链路层进行传输时，可以根据 MTU 对 IP 数据包进行分片，以确保 IP 数据包能够被传输。

[03]如图 1 所示，如果包含 IP 头和 IP 数据的 IP 数据包的长度大于 MTU，则可将该 IP 数据包中的 IP 数据拆分成 IP 数据 1 和 IP 数据 2，然后将 IP 数据包中的 IP 头分别与 IP 数据 1 和 IP 数据 2 重新封装成 IP 数据包 1 和 IP 数据包 2。这样，IP 数据包 1 的长度不大于 MTU，且 IP 数据包 2 的长度不大于 MTU。其中，可将拆分得到的第一个 IP 数据包 (例如，IP 数据包 15 1) 称为 IP 数据包的首数据包，将第一个 IP 数据包之外的其它 IP 数据包 (例如，IP 数据包 2) 称为 IP 数据包的后续数据包。

附图说明

[04]图 1 是根据 MTU 的值对 IP 数据包进行分片的示意图。

[05]图 2 是本公开一示例中的应用场景示意图。

20 [06]图 3 是本公开一示例中的 IP 数据包的转发方法的示意性流程图。

[07]图 4 是本公开另一示例中的 IP 数据包的转发方法的示意性流程图。

[08]图 5 是本公开一示例中的网络设备的示意性硬件结构图。

[09]图 6 是本公开一示例中的 IP 数据包的转发装置的示意性结构图。

具体实施方式

25 [10]下面将结合附图对本公开中的技术方案进行清楚、完整地描述，显然，所描述的示例仅

仅是本公开一部分示例，而不是全部的示例。基于本公开中的示例，本领域普通技术人员在没有做出创造性劳动前提下所获得的所有其他示例，都属于本公开保护的范围。

[11]本公开示例中提出一种 IP 数据包的转发方法，该方法可以应用在源设备与目的设备之间的网络设备上。图 2 为本公开示例的应用场景示意图，该方法可以应用在网络设备 201 或者
5 网络设备 202 上，以网络设备 202 为例。在源设备 21 向目的设备 22 发送 IP 数据包时，可以根据 MTU 判断是否对 IP 数据包进行分片。在 IP 数据包要被分片的情况下，以图 1 为例，源设备 21 可将 IP 数据包分成 IP 数据包 1 和 IP 数据包 2。通过对 IP 数据包分片得到的各个数据包，例如 IP 数据包 1 和 IP 数据包 2 可以称为“分片 IP 数据包”。网络设备 201 分别收到 IP 数据包 1 和 IP 数据包 2，并可分别将 IP 数据包 1 和 IP 数据包 2 发送给网络设备 202。在
10 上述应用场景下，以网络设备 202 对 IP 数据包 1 和 IP 数据包 2 的处理为例。如图 3 所示，该 IP 数据包的处理方法包括以下步骤：

[12]步骤 301，网络设备 202 可接收分片 IP 数据包，确定该分片 IP 数据包的业务类型。

[13]步骤 302，在所述分片 IP 数据包属于依赖端口信息的业务类型时，网络设备 202 可判断该分片 IP 数据包是否为首数据包。如果是，执行步骤 303；如果否，执行步骤 304。

15 [14]其中，分片 IP 数据包中还可以包含分片标记、分片偏移量等信息。分片标记表示当前 IP 数据包是经过分片处理的 IP 数据包。分片偏移量表示当前 IP 数据包是首数据包、或者第二个 IP 数据包、或者最后一个 IP 数据包等。基于分片偏移量，可以判断当前 IP 数据包是否为首数据包。

[15]例如，首数据包的分片偏移量为 0，如果首数据包的长度为 100 字节，则第二个 IP 数据
20 包的分片偏移量为 101 字节，如果第二个 IP 数据包的长度为 100 字节，则第三个 IP 数据包的分片偏移量为 201 字节，以此类推。基于此，如果当前 IP 数据包的分片偏移量是 0，则可确定当前 IP 数据包是首数据包，如果当前 IP 数据包的分片偏移量不是 0，则可确定当前 IP 数据包不是首数据包。

[16]网络设备 202 在接收到 IP 数据包 1 时，确定 IP 数据包 1 是首数据包。网络设备 202 在接
25 收到 IP 数据包 2 时，确定 IP 数据包 2 不是首数据包。IP 数据包 1 包括 IP 头和 IP 数据 1，IP 数据包 2 中包括 IP 头和 IP 数据 2。其中，IP 数据包 1 的 IP 头与 IP 数据包 2 的 IP 头相同，均为分片前的 IP 数据包中的 IP 头。在实际应用中，未经分片处理的初始 IP 数据包可包含 IP 头和 IP 数据，IP 头中的信息可称为 IP 头信息，而紧跟着 IP 头的 IP 数据部分可包含端口信息。基于此，经过分片处理后，该端口信息可被包含在首数据包 IP 数据中。

[17]本公开示例中，IP 头信息具体可以包括但不限于以下任意一种或者多种：标识符、源 IP 地址、目的 IP 地址、协议标识。端口信息具体可以包括但不限于以下任意一种或者多种：目的端口标识、源端口标识。

[18]其中，依赖端口信息的业务类型可以为 NAT（Network Address Translation，网络地址转换）业务，包过滤业务等。网络设备 202 可以根据 IP 头信息和端口信息进行业务处理。例如，针对 NAT 业务，网络设备 202 可以基于 IP 数据包的 IP 头信息中包括的源 IP 地址和端口信息中包括的源端口标识查询 NAT 映射表，并将源 IP 地址和源端口标识修改为 NAT 映射表中记录的 IP 地址和端口标识，该 NAT 业务的处理过程不再详加赘述。针对包过滤业务，网络设备 202 可以基于 IP 数据包的 IP 头信息中包括的源 IP 地址、目的 IP 地址、以及 IP 数据包的端口信息中包括的源端口标识、目的端口标识查询过滤表，以确定是否丢弃 IP 数据包，该包过滤业务的处理过程不再详加赘述。

[19]步骤 303，网络设备 202 可建立分片转发表项，该分片转发表项中记录了分片 IP 数据包的 IP 头信息和端口信息，并可利用端口信息对该分片 IP 数据包进行业务处理。

[20]步骤 304，网络设备 202 可将分片 IP 数据包的 IP 头信息对应的分片转发表项中记录的端口信息，与该分片 IP 数据包进行关联，并利用该端口信息对分片 IP 数据包进行业务处理。

[21]针对步骤 303，网络设备 202 在接收到 IP 数据包 1 时，可建立分片转发表项，该分片转发表项记录了 IP 数据包 1 中包括的 IP 头信息和端口信息。网络设备 202 可以直接利用该端口信息对 IP 数据包 1 进行 NAT 业务处理、包过滤业务处理等过程。

[22]针对步骤 304，网络设备 202 在接收到 IP 数据包 2 时，由于 IP 数据包 2 中不包含端口信息，因此，网络设备 202 无法直接利用端口信息对 IP 数据包 2 进行 NAT 业务处理、包过滤业务处理等过程。网络设备 202 可通过查找 IP 数据包 2 的 IP 头信息对应的分片转发表项（即包含 IP 数据包 1 中的 IP 头信息和端口信息的分片转发表项），确定对应的端口信息（即 IP 数据包 1 中的端口信息），并将该端口信息与该 IP 数据包 2 进行关联，并利用该端口信息对 IP 数据包 2 进行业务处理。

[23]本公开示例中，针对步骤 304，将分片 IP 数据包的 IP 头信息对应的分片转发表项中记录的端口信息，与该分片 IP 数据包进行关联的过程，具体可以包括但不限于如下方式：查询是否存在与该分片 IP 数据包的 IP 头信息对应的分片转发表项。如果存在与该 IP 头信息对应的分片转发表项，则直接将该分片转发表项中的端口信息与该分片 IP 数据包进行关联。如果不存在与该 IP 头信息对应的分片转发表项，则先缓存该分片 IP 数据包，并等待首数据包到来。

当首数据包到来时，可根据首数据包中的 IP 头信息和端口信息建立分片转发表项，然后将该端口信息与缓存的分片 IP 数据包进行关联。

[24]本公开示例中，将端口信息与该分片 IP 数据包进行关联的过程，具体可以包括但不限于如下方式：方式一、在本地的数据结构中记录端口信息与该分片 IP 数据包之间的对应关系；
5 或者，方式二、将端口信息添加到该分片 IP 数据包对应的存储区域中。

[25]为了方便描述，后续以方式二为例进行说明。针对方式二，在实际应用中，网络设备可针对每个收到的 IP 数据包，分配对应的存储区域，在此基础上，本公开示例中，可以将端口信息添加到 IP 数据包对应的存储区域中，而在对 IP 数据包进行业务处理时，可以直接从该存储区域中得到该 IP 数据包对应的端口信息，继而利用该端口信息对 IP 数据包进行业务处
10 理。

[26]假设网络设备 202 可先接收到 IP 数据包 1，后接收到 IP 数据包 2。在接收到 IP 数据包 2 时，由于已经基于 IP 数据包 1 建立了分片转发表项，因此，可以查询到 IP 数据包 2 的 IP 头信息对应的分片转发表项。可将该分片转发表项中记录的端口信息添加到 IP 数据包 2 对应的存储区域。之后可根据该端口信息对 IP 数据包 2 进行业务处理。

15 [27]假设网络设备 202 可先接收到 IP 数据包 2，后接收到 IP 数据包 1。在接收到 IP 数据包 2 时，由于当前还没有基于 IP 数据包 1 建立分片转发表项，因此，无法在 IP 数据包 2 对应的存储区域中添加端口信息，先缓存 IP 数据包 2。在接收到 IP 数据包 1，并基于 IP 数据包 1 建立分片转发表项之后，则可以查询到 IP 数据包 2 的 IP 头信息对应的分片转发表项。因此，可以将该分片转发表项中记录的端口信息添加到 IP 数据包 2 对应的存储区域。之后可根据该
20 端口信息对 IP 数据包 1 和 IP 数据包 2 进行业务处理。

[28]针对步骤 303，网络设备 202 在对 IP 数据包 1 进行业务处理后，可以直接将 IP 数据包 1 发送给目的设备 22。针对步骤 304，网络设备 202 在对 IP 数据包 2 进行业务处理后，可以直接将 IP 数据包 2 发送给目的设备 22。

25 [29]基于上述技术方案，本公开示例中，在接收到分片 IP 数据包后，可以区分分片 IP 数据包的业务类型。如果分片 IP 数据包属于依赖端口信息的业务类型，可以先等待首数据包。当首数据包到来时，可将首数据包的端口信息，与后续数据包进行关联。以这种方式，可获得所有分片 IP 数据包的端口信息，并可利用各自对应的端口信息对各分片 IP 数据包进行业务处理。在满足业务处理需求的同时，不需要对所有分片 IP 数据包进行重组，尽可能减少分片 IP 数据包的缓存，提升系统整体处理效率和转发效率。

[30]以下结合图 4 所示的流程图对上述 IP 数据包的转发方法进行详细说明。

[31]步骤 401, 网络设备可接收分片 IP 数据包, 确定该分片 IP 数据包的业务类型。当该分片 IP 数据包属于依赖端口信息的业务类型时, 则执行步骤 402; 当该分片 IP 数据包属于依赖内容信息的业务类型时, 则执行步骤 405; 当该分片 IP 数据包属于不用进行业务处理的业务类型时, 则执行步骤 408。

[32]其中, 依赖端口信息的业务类型, 可以如 NAT 业务, 包过滤业务等。依赖内容信息 (如 TCP (Transmission Control Protocol, 传输控制协议) 端口、HTTP (Hyper Text Transfer Protocol, 超文本传输协议) 内容等) 的业务类型, 可以如 DPI (Deep Packet Inspection, 深度包检测) 业务等。不用进行业务处理, 即 IP 数据包可以直接进行转发。

[33]其中, 当网络设备上没有配置依赖端口信息的业务类型的业务, 且没有配置依赖内容信息的业务类型的业务时, 则可以认为 IP 数据包不用进行业务处理而被直接转发。

[34]在实际应用中, 假设 IP 数据包的业务类型同时为多个, 则可以按照各业务类型的优先级选择 IP 数据包的业务类型。例如, 业务类型的优先级顺序可以为: 依赖内容信息的业务类型 > 依赖端口信息的业务类型 > 不用进行业务处理。当 IP 数据包同时具有依赖内容信息的业务类型和依赖端口信息的业务类型时, 如对 IP 数据包进行 NAT 业务和 DPI 业务, 则可选择 IP 数据包的业务类型为依赖内容信息的业务类型。

[35]步骤 402, 当该分片 IP 数据包属于依赖端口信息的业务类型时, 网络设备可判断该 IP 数据包是否为首数据包。如果是, 可执行步骤 403; 如果否, 可执行步骤 404。

[36]其中, 分片 IP 数据包中还可以包含分片标记、分片偏移量等信息。分片标记表示当前 IP 数据包是经过分片处理的 IP 数据包。分片偏移量表示当前 IP 数据包是首数据包、或者第二个 IP 数据包、或者最后一个 IP 数据包等。基于分片偏移量, 可以判断当前 IP 数据包是否为首数据包。

[37]例如, 首数据包的分片偏移量为 0, 如果首数据包的长度为 100 字节, 则第二个 IP 数据包的分片偏移量为 101 字节, 如果第二个 IP 数据包的长度为 100 字节, 则第三个 IP 数据包的分片偏移量为 201 字节, 以此类推。基于此, 如果当前 IP 数据包的分片偏移量是 0, 则可确定当前 IP 数据包是首数据包, 如果当前 IP 数据包的分片偏移量不是 0, 则可确定当前 IP 数据包不是首数据包。

[38]步骤 403, 网络设备可建立分片转发表项, 该分片转发表项中记录了分片 IP 数据包的 IP 头信息和端口信息, 并可利用该端口信息对该分片 IP 数据包进行业务处理。

[39]步骤 404, 网络设备可将分片 IP 数据包的 IP 头信息对应的分片转发表项中记录的端口信息, 与该分片 IP 数据包进行关联, 并利用该端口信息对分片 IP 数据包进行业务处理。

[40]本公开示例中, IP 头信息具体可以包括但不限于以下任意一种或者多种: 标识符、源 IP 地址、目的 IP 地址、协议标识。端口信息具体可以包括但不限于以下任意一种或者多种: 目的端口标识、源端口标识。

[41]本公开示例中, 针对步骤 404, 将分片 IP 数据包的 IP 头信息对应的分片转发表项中记录的端口信息, 与该分片 IP 数据包进行关联的过程, 具体可以包括但不限于如下方式: 利用该分片 IP 数据包的 IP 头信息查询分片转发表项。如果存在对应该 IP 头信息的分片转发表项, 则可直接将该分片转发表项中的端口信息与该分片 IP 数据包进行关联。如果不存在对应该 IP 头信息的分片转发表项, 则可先缓存该分片 IP 数据包, 并等待首数据包到来。当首数据包到来时, 可根据首数据包中的 IP 头信息和端口信息建立分片转发表项, 然后将该端口信息与缓存的分片 IP 数据包进行关联。

[42]本公开示例中, 将端口信息与该分片 IP 数据包进行关联的过程, 具体可以包括但不限于如下方式: 方式一、在本地的数据结构中记录端口信息与该分片 IP 数据包之间的对应关系; 或者, 方式二、将端口信息添加到该分片 IP 数据包对应的存储区域中。

[43]为了方便描述, 后续以方式二为例进行说明。针对方式二, 在实际应用中, 网络设备可针对每个收到的分片 IP 数据包分配对应的存储区域。在此基础上, 本公开示例中, 可以将端口信息添加到分片 IP 数据包对应的存储区域中。而在对分片 IP 数据包进行业务处理时, 可以直接从该存储区域中得到该分片 IP 数据包对应的端口信息, 继而利用该端口信息对分片 IP 数据包进行业务处理。

[44]步骤 402 至步骤 404 的处理与步骤 302 至步骤 304 的处理类似, 在此不再赘述。

[45]基于步骤 402 至步骤 404 的技术方案, 在接收到分片 IP 数据包后, 可以区分该分片 IP 数据包的业务类型。如果该分片 IP 数据包属于依赖端口信息的业务类型, 可以先等待首数据包。当首数据包到来时, 可将首数据包的端口信息与后续数据包进行关联, 使得能够获得所有分片 IP 数据包的端口信息, 并可利用各自对应的端口信息对各分片 IP 数据包进行业务处理。在满足业务处理需求的同时, 不需要对所有分片 IP 数据包进行重组, 尽可能减少分片 IP 数据包的缓存, 提升系统整体处理效率和转发效率, 尽可能快的转发分片 IP 数据包, 缩短了时延。

[46]步骤 405, 当该分片 IP 数据包属于依赖内容信息的业务类型时, 网络设备可判断与原 IP

数据包对应的所有分片 IP 数据包是否均已接收；如果否，可执行步骤 406；如果是，可执行步骤 407。

[47]在实际应用中，每个分片 IP 数据包中均会携带分片偏移量，且最后一个分片 IP 数据包中还会携带表示其是最后数据包的标记。基于各分片 IP 数据包的分片偏移量以及最后数据包的

5 标记，可以判断出当前是否已存在所有分片 IP 数据包。

[48]例如，假设已经收到携带表示最后数据包的标记的分片 IP 数据包，该分片 IP 数据包的分片偏移量为 301 字节，如果当前已经收到分片偏移量为 0、数据包的长度为 100 字节的分片 IP 数据包，分片偏移量为 101 字节、数据包的长度为 100 字节的分片 IP 数据包，分片偏移量为 201 字节、数据包的长度为 100 字节的分片 IP 数据包，则这 4 个分片 IP 数据包正好组成

10 一个完整的 IP 数据包，因此，判断出当前已存在所有分片 IP 数据包。

[49]步骤 406，网络设备可缓存分片 IP 数据包，继续接收其它分片 IP 数据包。

[50]步骤 407，网络设备可对原 IP 数据包对应的所有分片 IP 数据包进行重组，并对重组后的 IP 数据包进行业务处理。

[51]例如，网络设备 202 在先接收到 IP 数据包 1 尚未接收到 IP 数据包 2 时，可缓存 IP 数据包 1。在 IP 数据包 2 也被接收到时，由于当前已经存在原 IP 数据包对应的所有分片 IP 数据包，因此网络设备 202 可对所有分片 IP 数据包（IP 数据包 1 和 IP 数据包 2）进行重组，得到重组后的 IP 数据包，并对重组后的 IP 数据包进行业务处理。

15

[52]其中，针对依赖内容信息的业务类型，例如，DPI 业务，网络设备 202 在仅接收到 IP 数据包 1 或者 IP 数据包 2 时，由于 IP 数据包 1 或者 IP 数据包 2 均只包含部分 IP 数据，因此无法单独基于 IP 数据包 1 或者 IP 数据包 2 进行业务处理。在此基础上，需要将 IP 数据包 1 和 IP 数据包 2 重组为未经分片处理的 IP 数据包，并对重组后的 IP 数据包进行业务处理。

20

[53]其中，在对 IP 数据包 1 和 IP 数据包 2 进行重组的过程中，可以根据 IP 数据包 1 中包含的源 IP 地址、目的 IP 地址、协议标识、分片标记、分片偏移量，IP 数据包 2 中包含的源 IP 地址、目的 IP 地址、协议标识、分片标记、分片偏移量等信息，对 IP 数据包 1 和 IP 数据包

25 2 进行重组，重组后生成的 IP 数据包与未经分片处理的 IP 数据包内容相同。

[54]针对步骤 407，网络设备 202 在对重组后的 IP 数据包进行业务处理后，可以直接将重组后的 IP 数据包发送给目的设备 22。或者，网络设备 202 也可以根据 MTU 将重组后的 IP 数据包分片成 IP 数据包 1 和 IP 数据包 2，并分别将 IP 数据包 1 和 IP 数据包 2 发送给目的设备 22。

[55]步骤 408, 当该分片 IP 数据包属于不用进行业务处理的业务类型时, 可利用该分片 IP 数据包的 IP 头信息转发该分片 IP 数据包。其中, 该 IP 头信息具体可以包括但不限于以下一种或者多种: 标识符、源 IP 地址、目的 IP 地址、协议标识。

[56]例如, 网络设备 202 在接收到 IP 数据包 1 之后, 可以利用 IP 数据包 1 中携带的标识符、源 IP 地址、目的 IP 地址、协议标识, 将 IP 数据包 1 发送给目的设备 22。网络设备 202 在接收到 IP 数据包 2 之后, 可以利用 IP 数据包 2 中携带的源 IP 地址、目的 IP 地址、协议标识, 将 IP 数据包 2 发送给目的设备 22。

[57]本公开示例中还提供了一种 IP 数据包的转发装置, 该 IP 数据包的转发装置可以应用在网络设备上。其中, 该 IP 数据包的转发装置可以通过软件实现, 也可以通过硬件或者软硬件结合的方式实现。以软件实现为例, 作为一个逻辑意义上的装置, 是通过其所在的网络设备的处理器, 读取非易失性存储介质中对应的机器可读程序指令形成的。如图 5 所示, 为本公开提出的 IP 数据包的转发装置所在的网络设备的一种硬件结构图。除了图 5 所示的处理器 501、非易失性存储介质 502 外, 网络设备还可以包括其他硬件, 如负责处理报文的转发芯片、网络接口、内存等。从硬件结构上来讲, 该网络设备还可能是分布式设备, 可能包括多个接口卡, 以便进行报文处理的扩展。

[58]如图 6 所示, 为本公开提出的 IP 数据包的转发装置的结构图, 所述装置可包括: 确定模块 11, 用于接收分片 IP 数据包, 并确定所述分片 IP 数据包的业务类型; 判断模块 12, 用于当所述分片 IP 数据包属于依赖端口信息的业务类型时, 判断所述分片 IP 数据包是否为首数据包; 处理模块 13, 用于在所述分片 IP 数据包是首数据包的情况下, 建立分片转发表项, 所述分片转发表项中记录了所述分片 IP 数据包的 IP 头信息和端口信息, 并利用所述端口信息对所述分片 IP 数据包进行业务处理; 在所述分片 IP 数据包是首数据包以外的后续数据包的情况下, 将所述分片 IP 数据包的 IP 头信息对应的分片转发表项中记录的端口信息, 与所述分片 IP 数据包进行关联, 并利用所述端口信息对所述分片 IP 数据包进行业务处理。

[59]所述处理模块 13, 用于在将所述分片 IP 数据包的 IP 头信息对应的分片转发表项中记录的端口信息, 与所述分片 IP 数据包进行关联的过程中, 查询是否存在对应所述分片 IP 数据包的 IP 头信息的分片转发表项; 在存在所述 IP 头信息对应的分片转发表项的情况下, 将所述分片转发表项中的端口信息与所述分片 IP 数据包进行关联; 在不存在所述 IP 头信息对应的分片转发表项的情况下, 缓存所述分片 IP 数据包, 并等待所述首数据包的接收, 在所述首数据包被接收时, 将所述首数据包中的的端口信息与缓存的所述分片 IP 数据包进行关联。

[60]所述判断模块 12, 还用于在所述分片 IP 数据包属于依赖内容信息的业务类型时, 判断与

原 IP 数据包对应的所有分片 IP 数据包是否均被接收;

[61]所述处理模块 13, 还用于在尚有与原 IP 数据包对应的分片 IP 数据包未被接收的情况下, 缓存所述分片 IP 数据包, 继续接收与原 IP 数据包对应的其它分片 IP 数据包; 在与原 IP 数据包对应的所有分片 IP 数据包均被接收的情况下, 对与原 IP 数据包对应的所有分片 IP 数据包进行重组, 并对重组后的 IP 数据包进行业务处理。

[62]所述处理模块 13, 还用于当所述分片 IP 数据包属于不用进行业务处理的业务类型时, 利用所述 IP 数据包的 IP 头信息转发所述 IP 数据包。

[63]本公开示例中, 所述 IP 头信息包括以下任意一种或者多种: 标识符、源 IP 地址、目的 IP 地址、协议标识; 所述端口信息包括以下任意一种或者多种: 目的端口标识和、源端口标识。

10 [64]其中, 本公开装置的各个模块可以集成于一体, 也可以分离部署。上述模块可以合并为一个模块, 也可以进一步拆分成多个子模块。

[65]通过以上示例的描述, 本领域的技术人员可以清楚地了解到本公开可借助软件加通用硬件平台的方式来实现, 当然也可以通过硬件来实现。基于这样的理解, 本公开的技术方案本质上或者说对现有技术做出贡献的部分可以以软件产品的形式体现出来, 该计算机软件产品存储在一个存储介质中, 包括若干指令用以使得一台机器可执行设备 (可以是个人计算机, 服务器, 或者网络设备等) 执行本公开各个示例所述的方法。本领域技术人员可以理解附图只是一个优选示例的示意图, 附图中的模块或流程并不一定是实施本公开所必须的。

15 [66]本领域技术人员可以理解示例中的装置中的模块可以按照描述分布于本示例的装置中, 也可以进行相应变化位于不同于本示例的一个或多个装置中。上述示例的模块可以合并为一个模块, 也可进一步拆分成多个子模块。上述本公开示例序号仅仅为了描述, 不代表示例的优劣。

[67]以上仅为本公开的示例, 但是, 本公开并非局限于此, 任何本领域的技术人员能思之的变化都应落入本公开的保护范围。

权利要求书

1. 一种互联网协议（IP）数据包的转发方法，包括：

网络设备接收分片 IP 数据包，并确定所述分片 IP 数据包的业务类型；

5 当所述分片 IP 数据包属于依赖端口信息的业务类型时，所述网络设备判断所述分片 IP 数据包是否为首数据包；

在所述分片 IP 数据包是首数据包的情况下，

所述网络设备建立分片转发表项，所述分片转发表项记录了所述分片 IP 数据包的 IP 头信息和端口信息，并且

所述网络设备利用所述端口信息对所述分片 IP 数据包进行业务处理；

10 在所述分片 IP 数据包是首数据包以外的后续数据包的情况下，

所述网络设备将所述分片 IP 数据包的 IP 头信息对应的分片转发表项中记录的端口信息，与所述分片 IP 数据包进行关联，并且

所述网络设备利用所述端口信息对所述分片 IP 数据包进行业务处理。

2. 根据权利要求 1 所述的方法，其特征在于，将所述分片 IP 数据包的 IP 头信息对应的
15 分片转发表项中记录的端口信息，与所述分片 IP 数据包进行关联，包括：

所述网络设备查询是否存在与所述分片 IP 数据包的 IP 头信息对应的分片转发表项；

在存在所述 IP 头信息对应的分片转发表项的情况下，

所述网络设备将所述分片转发表项中的端口信息与所述分片 IP 数据包进行关联；

在不存在所述 IP 头信息对应的分片转发表项的情况下，

20 所述网络设备缓存所述分片 IP 数据包，并等待所述首数据包的接收，

在所述首数据包被接收时，所述网络设备将所述首数据包中的端口信息与缓存的所述分片 IP 数据包进行关联。

3. 根据权利要求 1 所述的方法，所述方法还包括：

在所述分片 IP 数据包属于依赖内容信息的业务类型的情况下，所述网络设备判断与原 IP
25 数据包对应的所有分片 IP 数据包是否均被接收；

在尚有所述原 IP 数据包对应的分片 IP 数据包未被接收的情况下，

所述网络设备缓存所述分片 IP 数据包，继续接收所述原 IP 数据包对应的其它分片 IP 数据包；

在与原 IP 数据包对应的所有分片 IP 数据包均被接收的情况下，

30 所述网络设备对所述原 IP 数据包对应的所有分片 IP 数据包进行重组，并对重组后

的 IP 数据包进行业务处理。

4. 根据权利要求 1 所述的方法, 所述方法还包括:

在所述分片 IP 数据包属于不用进行业务处理的业务类型的情况下, 所述网络设备利用所述分片 IP 数据包的 IP 头信息转发所述分片 IP 数据包。

5 5. 根据权利要求 1 所述的方法, 其中,

所述 IP 头信息包括以下任意一种或者多种: 标识符、源 IP 地址、目的 IP 地址、协议标识; 所述端口信息包括以下任意一种或者多种: 目的端口标识、源端口标识。

6. 一种互联网协议 (IP) 数据包的转发装置, 包括处理器以及存储有机器可读指令的非暂时性存储介质, 通过执行所述机器可读指令, 所述处理器被使得:

10 接收分片 IP 数据包, 并确定所述分片 IP 数据包的业务类型;

当所述分片 IP 数据包属于依赖端口信息的业务类型时, 判断所述分片 IP 数据包是否为首数据包;

在所述分片 IP 数据包是首数据包的情况下,

15 建立分片转发表项, 所述分片转发表项中记录了所述分片 IP 数据包的 IP 头信息和端口信息, 并且

利用所述端口信息对所述分片 IP 数据包进行业务处理;

在所述分片 IP 数据包是首数据包以外的后续数据包的情况下,

将所述分片 IP 数据包的 IP 头信息对应的分片转发表项中记录的端口信息, 与所述分片 IP 数据包进行关联, 并且

20 利用所述端口信息对所述分片 IP 数据包进行业务处理。

7. 根据权利要求 6 所述的装置, 其中, 在将所述分片 IP 数据包的 IP 头信息对应的分片转发表项中记录的端口信息, 与所述分片 IP 数据包进行关联时, 所述机器可执行指令还使得所述处理器:

查询是否存在对应所述分片 IP 数据包的 IP 头信息的分片转发表项;

25 在存在所述分片 IP 数据包的 IP 头信息对应的分片转发表项的情况下,

将所述分片转发表项中的端口信息与所述分片 IP 数据包进行关联;

在不存在所述分片 IP 数据包的 IP 头信息对应的分片转发表项的情况下,

缓存所述分片 IP 数据包, 并等待所述首数据包的接收,

30 在所述首数据包被接收时, 将所述首数据包中的端口信息与缓存的所述分片 IP 数据包进行关联。

8. 根据权利要求 6 所述的装置, 其中, 所述机器可执行指令还使得所述处理器:

在所述分片 IP 数据包属于依赖内容信息的业务类型的情况下, 判断与原 IP 数据包对应的所有分片 IP 数据包是否均被接收;

在尚有原 IP 数据包对应的分片 IP 数据包未被接收的情况下,

5 缓存所述分片 IP 数据包, 继续接收原 IP 数据包对应的其它分片 IP 数据包;

在与原 IP 数据包对应的所有分片 IP 数据包均被接收的情况下,

对与原 IP 数据包对应的所有分片 IP 数据包进行重组, 并对重组后的 IP 数据包进行业务处理。

9. 根据权利要求 6 所述的装置, 其中, 所述机器可执行指令还使得所述处理器:

10 在所述分片 IP 数据包属于不用进行业务处理的业务类型的情况下, 利用所述分片 IP 数据包的 IP 头信息转发所述分片 IP 数据包。

10. 根据权利要求 6 所述的装置, 其中,

所述 IP 头信息包括以下任意一种或者多种: 标识符、源 IP 地址、目的 IP 地址、协议标识; 所述端口信息包括以下任意一种或者多种: 目的端口标识、源端口标识。

15

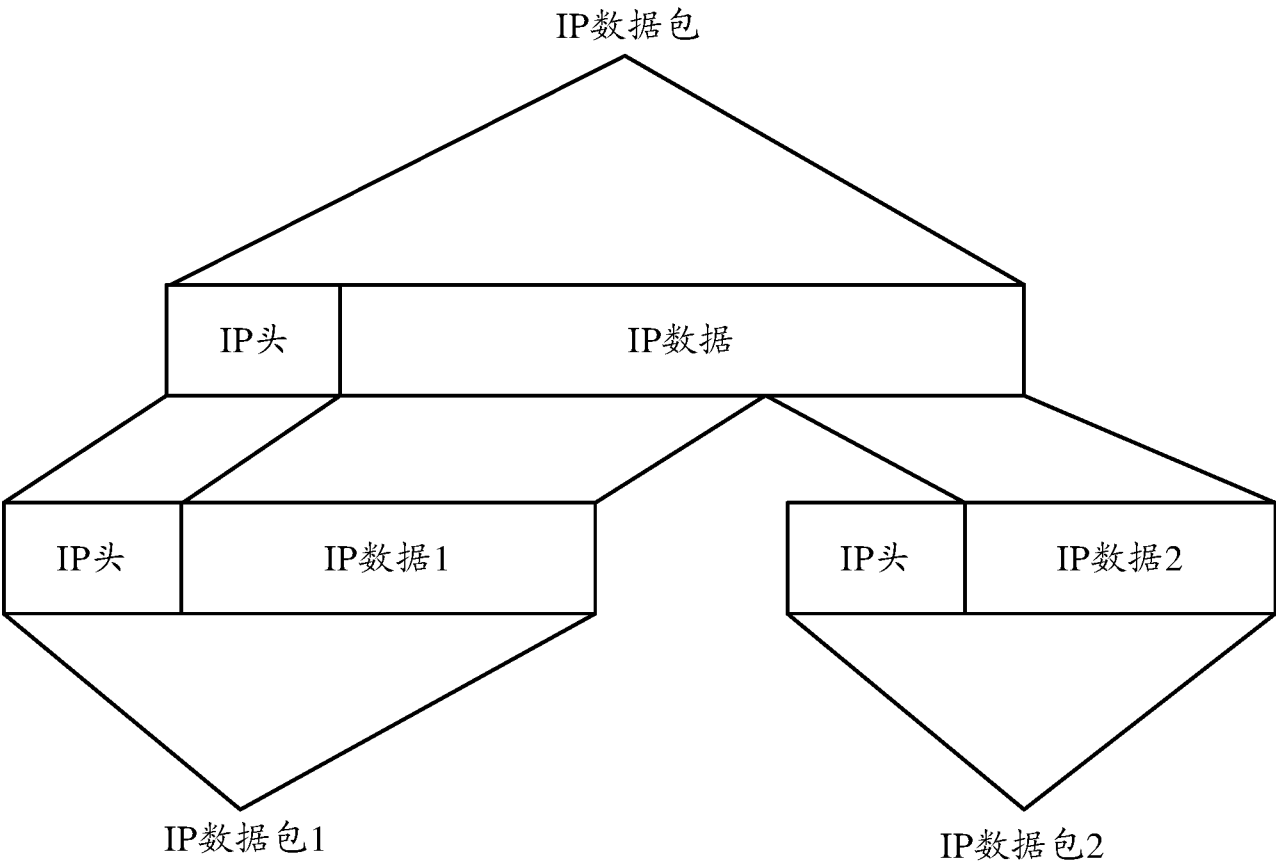


图 1

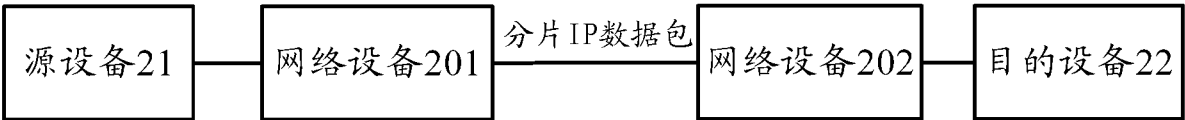


图 2

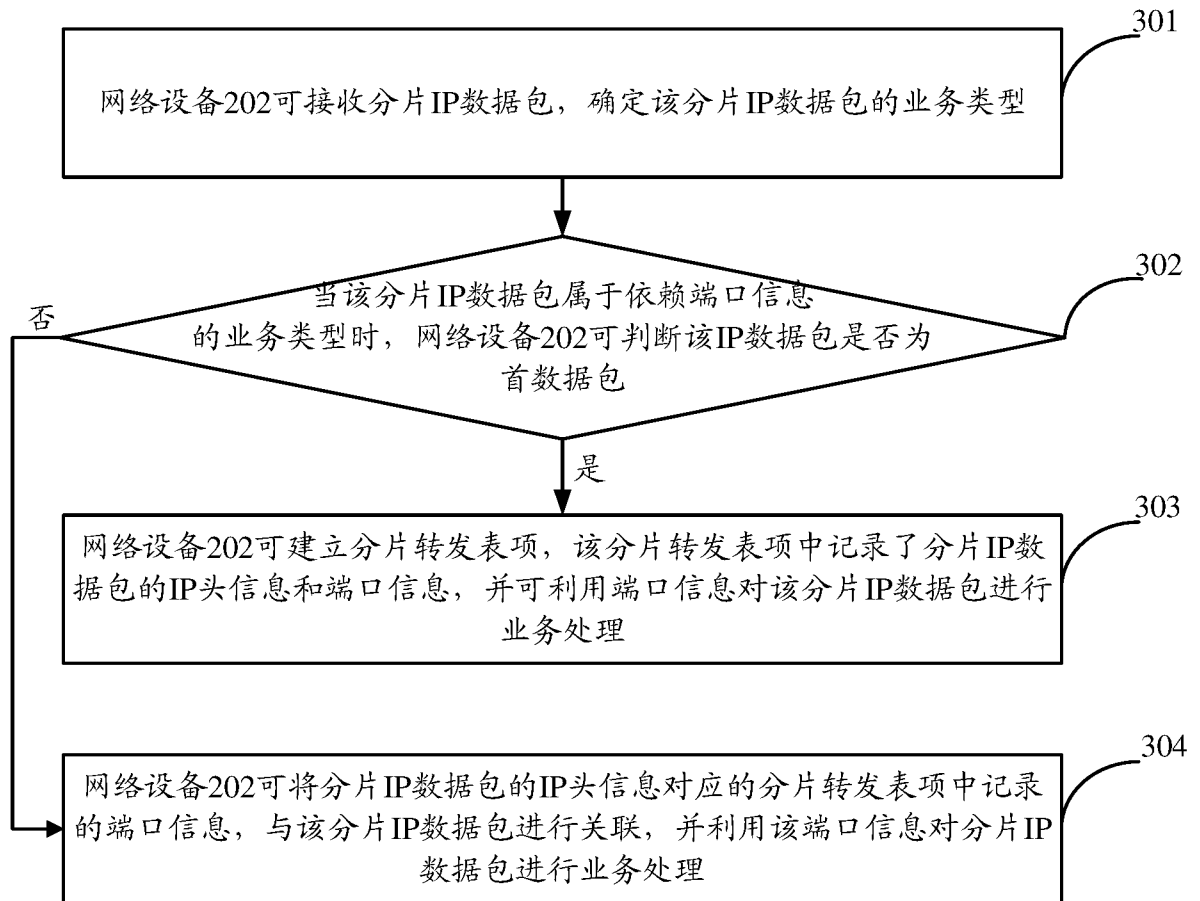


图 3

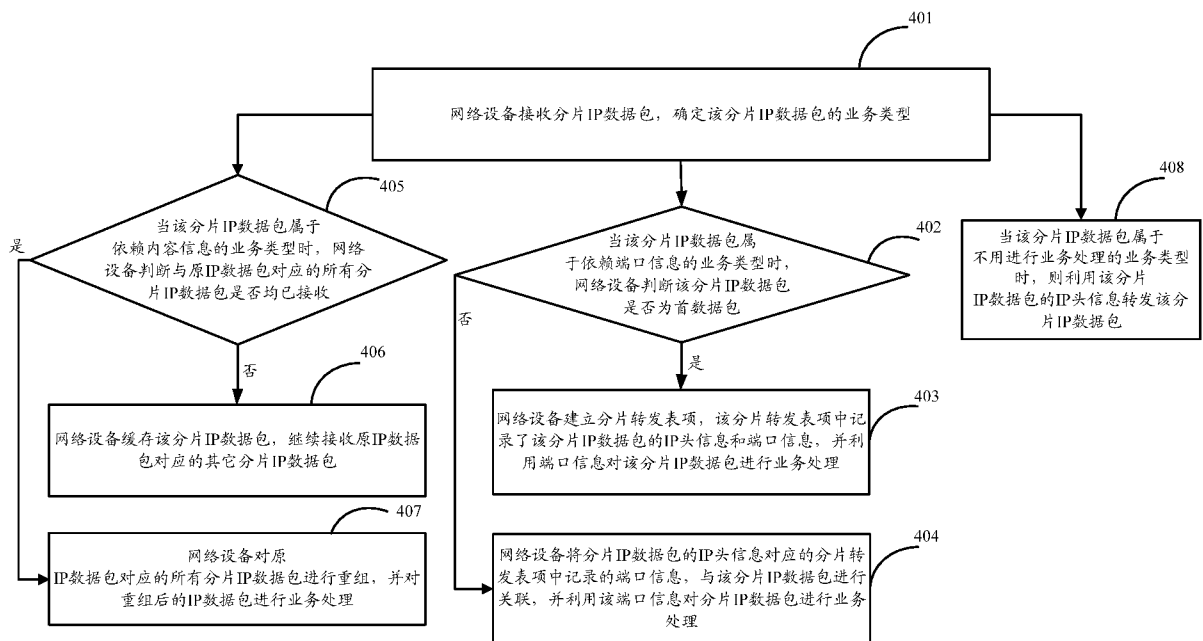


图 4

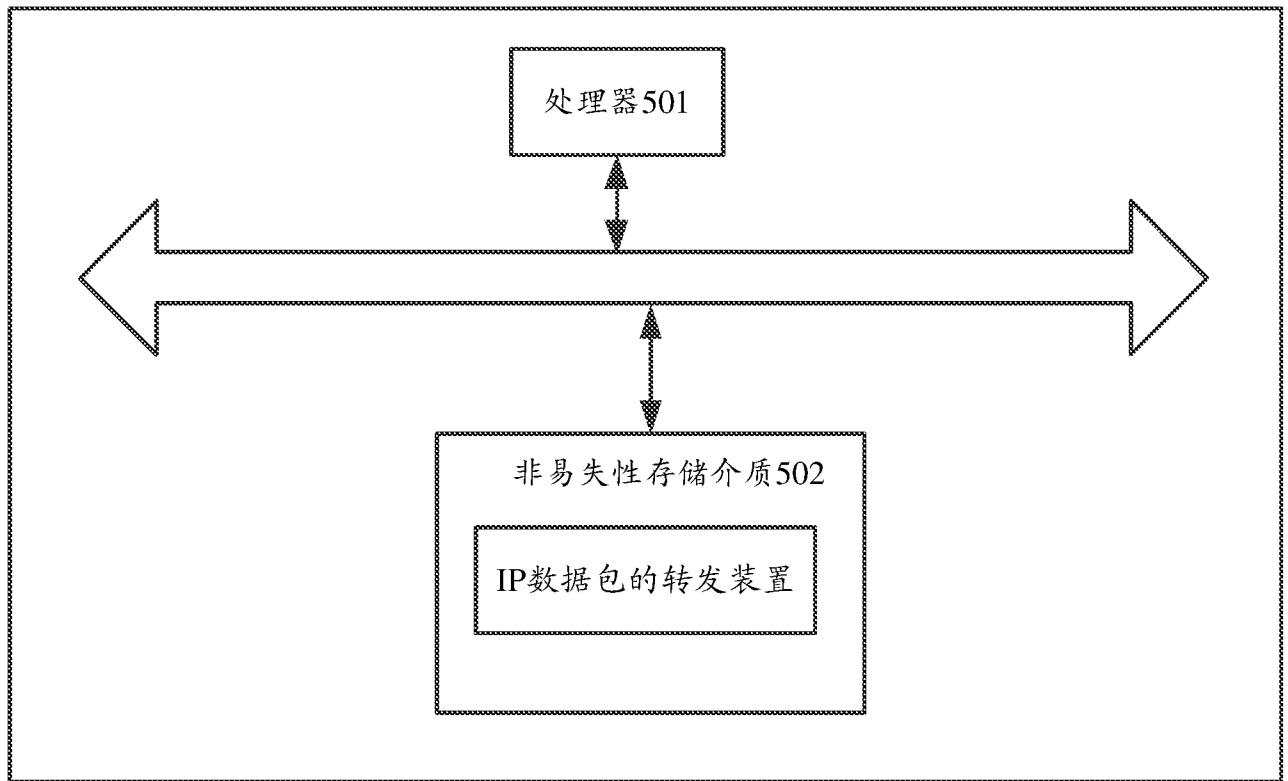


图 5

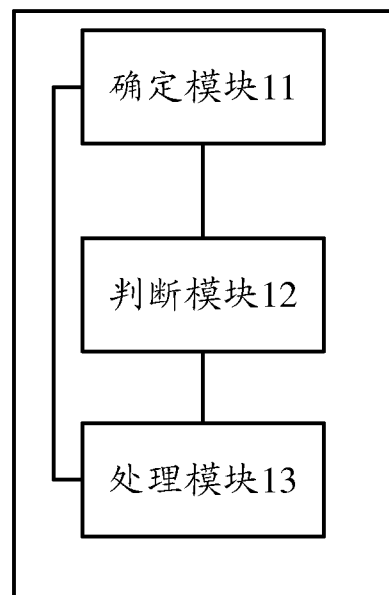


图 6

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2016/112270

A. CLASSIFICATION OF SUBJECT MATTER

H04L 12/851 (2013.01) 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)

H04L

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)

CNPAT; CNKI; WPI; EPODOC: IP, fragment, first, second, subsequent, NAT, packet filtering, SEGMENT+, NETWORK ADDRESS TRANSLATION, NETWORK ADDRESS CONVERS+, PORT, FORWARD+, MAP+, TABLE, SOURCE, DESTINATION

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	CN 1960316 A (HUAWEI TECHNOLOGIES CO., LTD.), 09 May 2007 (09.05.2007), description, page 1, line 6 to page 3, line 12 and page 5, line 15 to page 7, bottom line, and figures 1-3	1-10
X	CN 1863158 A (HUAWEI TECHNOLOGIES CO., LTD.), 15 November 2006 (15.11.2006), description, page 4, line 7 to page 5, line 8	1-10
X	CN 101783786 A (ZTE CORP.), 21 July 2010 (21.07.2010), description, paragraphs 0053-0062	1-10
A	US 2011211577 A1 (CONNECTION TECHNOLOGY SYSTEMS INC.), 01 September 2011 (01.09.2011), the whole document	1-10
A	US 7624195 B1 (CISCO TECHNOLOGY, INC.), 24 November 2009 (24.11.2009), the whole document	1-10

☐ 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 09 February 2017 (09.02.2017)	Date of mailing of the international search report 07 March 2017 (07.03.2017)
Name and mailing address of the ISA/CN: State Intellectual Property Office of the P. R. China No. 6, Xitucheng Road, Jimenqiao Haidian District, Beijing 100088, China Facsimile No.: (86-10) 62019451	Authorized officer XIN, Haiming Telephone No.: (86-10) 82246953

International application No.
PCT/CN2016/112270

Form PCT/ISA/210 (patent family annex) (July 2009)

A. 主题的分类 H04L 12/851 (2013.01) i 按照国际专利分类 (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) H04L 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)) CNPAT; CNKI; WPI; EPODOC; IP, 分片, 第一, 首, 第二, 后续, 网络地址转换, NAT, 包过滤, 端口, SEGMENT+, NETWORK ADDRESS TRANSLATION, NETWORK ADDRESS CONVERS+, PORT, FORWARD+, MAP+, TABLE, SOURCE, DESTINATION																				
C. 相关文件 <table border="1"> <thead> <tr> <th>类 型*</th> <th>引用文件, 必要时, 指明相关段落</th> <th>相关的权利要求</th> </tr> </thead> <tbody> <tr> <td>X</td> <td>CN 1960316 A (华为技术有限公司) 2007年 5月 9日 (2007 - 05 - 09) 说明书第1页第6行到第3页第12行, 第5页第15行到第7页最后一行, 图1-3</td> <td>1-10</td> </tr> <tr> <td>X</td> <td>CN 1863158 A (华为技术有限公司) 2006年 11月 15日 (2006 - 11 - 15) 说明书第4页第7行到第5页第8行</td> <td>1-10</td> </tr> <tr> <td>X</td> <td>CN 101783786 A (中兴通讯股份有限公司) 2010年 7月 21日 (2010 - 07 - 21) 说明书第0053-0062段</td> <td>1-10</td> </tr> <tr> <td>A</td> <td>US 2011211577 A1 (CONNECTION TECHNOLOGY SYSTEMS INC.) 2011年 9月 1日 (2011 - 09 - 01) 全文</td> <td>1-10</td> </tr> <tr> <td>A</td> <td>US 7624195 B1 (CISCO TECHNOLOGY, INC.) 2009年 11月 24日 (2009 - 11 - 24) 全文</td> <td>1-10</td> </tr> </tbody> </table>			类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求	X	CN 1960316 A (华为技术有限公司) 2007年 5月 9日 (2007 - 05 - 09) 说明书第1页第6行到第3页第12行, 第5页第15行到第7页最后一行, 图1-3	1-10	X	CN 1863158 A (华为技术有限公司) 2006年 11月 15日 (2006 - 11 - 15) 说明书第4页第7行到第5页第8行	1-10	X	CN 101783786 A (中兴通讯股份有限公司) 2010年 7月 21日 (2010 - 07 - 21) 说明书第0053-0062段	1-10	A	US 2011211577 A1 (CONNECTION TECHNOLOGY SYSTEMS INC.) 2011年 9月 1日 (2011 - 09 - 01) 全文	1-10	A	US 7624195 B1 (CISCO TECHNOLOGY, INC.) 2009年 11月 24日 (2009 - 11 - 24) 全文	1-10
类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求																		
X	CN 1960316 A (华为技术有限公司) 2007年 5月 9日 (2007 - 05 - 09) 说明书第1页第6行到第3页第12行, 第5页第15行到第7页最后一行, 图1-3	1-10																		
X	CN 1863158 A (华为技术有限公司) 2006年 11月 15日 (2006 - 11 - 15) 说明书第4页第7行到第5页第8行	1-10																		
X	CN 101783786 A (中兴通讯股份有限公司) 2010年 7月 21日 (2010 - 07 - 21) 说明书第0053-0062段	1-10																		
A	US 2011211577 A1 (CONNECTION TECHNOLOGY SYSTEMS INC.) 2011年 9月 1日 (2011 - 09 - 01) 全文	1-10																		
A	US 7624195 B1 (CISCO TECHNOLOGY, INC.) 2009年 11月 24日 (2009 - 11 - 24) 全文	1-10																		
☐ 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 (如具体说明的) “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 2017年 2月 9日		国际检索报告邮寄日期 2017年 3月 7日																		
ISA/CN的名称和邮寄地址 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 辛海明 电话号码 (86-10) 82246953																		

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2016/112270

检索报告引用的专利文件			公布日 (年/月/日)	同族专利			公布日 (年/月/日)
CN	1960316	A	2007年 5月 9日	无			
CN	1863158	A	2006年 11月 15日	无			
CN	101783786	A	2010年 7月 21日	无			
US	2011211577	A1	2011年 9月 1日	EP	2362608	A1	2011年 8月 31日
				TW	201130265	A	2011年 9月 1日
US	7624195	B1	2009年 11月 24日	US	7653745	B1	2010年 1月 26日