



(12)发明专利申请

(10)申请公布号 CN 105812287 A

(43)申请公布日 2016.07.27

(21)申请号 201610389666.X

H04L 29/14(2006.01)

(22)申请日 2012.04.28

H04W 28/20(2009.01)

(30)优先权数据

13/096,194 2011.04.28 US

(62)分案原申请数据

201280020755.7 2012.04.28

(71)申请人 微软技术许可有限责任公司

地址 美国华盛顿州

(72)发明人 J·E·埃尔森 E·B·南丁格尔

(74)专利代理机构 上海专利商标事务所有限公司 31100

代理人 杨丽

(51)Int.Cl.

H04L 12/925(2013.01)

H04L 29/06(2006.01)

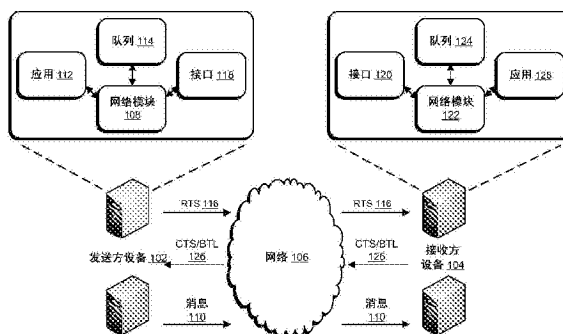
权利要求书1页 说明书12页 附图5页

(54)发明名称

分组交换网络中的有效电路

(57)摘要

本文描述了通过分组交换网络在发送方设备和接收方设备之间创建有效电路。为了建立有效电路,发送方设备通过分组交换网络向接收方设备发送请求。该请求与来自接收方设备用于从发送方设备接收消息的带宽保留相关联。接收方设备从多个发送方设备接收到多个请求,并且为这些发送方设备中的至少一个发送方设备保留带宽。接收方设备随后向该至少一个发送方设备发送响应,该响应提供了使用保留带宽向接收方设备发送消息的许可,该请求和响应从而建立了有效电路。接收方设备还可拒绝其他发送方设备的请求,从而致使该其他发送方设备向其他接收方设备发送其他请求。



1. 一种发送方设备,包括:
处理器;以及
被配置成由所述处理器操作以执行以下操作的网络模块,所述操作包括:
向第一接收方设备发送第一请求发送(RTS)第一消息,所述第一RTS向所述第一接收方设备寻求用于从所述发送方设备接收所述第一消息的第一带宽保留;
从所述第一接收方设备接收到拒绝所述第一RTS的响应;以及
基于所述响应,向第二接收方设备发送第二RTS,所述第二RTS向所述第二接收方设备寻求用于从所述发送方设备接收第二消息的第二带宽保留;
至少部分地基于所述第一接收方设备拒绝所述第一RTS,排定所述第一RTS的优先级高于第三RTS;以及
在稍后的时间从所述发送方设备向所述第一接收方设备重新发送所述RTS。
2. 如权利要求1所述的发送方设备,其特征在于,还包括存储供传送到接收方设备的消息的一个或多个队列。
3. 如权利要求2所述的发送方设备,其特征在于,所述网络模块被进一步配置成从所述一个或多个队列中选择所述队列中未被定向到所述第一接收方设备且其RTS尚未被发送的消息作为所述第二消息。
4. 一种用于接收方设备的方法,所述方法包括:
从多个发送方设备接收多个请求发送消息;
对所述多个请求排定优先级以供处理;
从所述多个请求中确定指示最高优先级的请求;
为所述发送方设备中与指示最高优先级的请求相关联的至少一个发送方设备保留带宽;以及
向所述发送方设备中的所述至少一个发送方设备发送响应,所述响应提供了使用保留带宽向所述接收方设备发送一个或多个消息的许可。

分组交换网络中的有效电路

[0001] 本申请是2012年4月28日提交的申请号为201280020755.7的题为“分组交换网络中的有效电路”的申请的分案申请。

技术领域

[0002] 本发明涉及分组交换网络,尤其涉及分组交换网络中的有效电路。

背景技术

[0003] 消息常常使用传输控制协议/网际协议(TCP/IP)协议套件通过分组交换网络来传送。TCP/IP使用反复试验的方法确定用于发送机设备进行消息传输的路径带宽。发送机设备发送消息的速率增大直到分组丢失发生。TCP/IP通过注意到未收到针对已发送消息的确认来检测分组丢失。分组丢失常常源自诸如在中介或端点设备处发生的缓冲器溢出等事件。在检测到分组丢失后,TCP/IP降低传输速率并且以逐渐增大的速率来传送消息直到再次检测到分组丢失。尽管这种方法针对在设备对之间传送大量消息的情况可以良好地工作,但对于传送小的消息或者少量消息是不高效的,因为在TCP/IP有机会检测到分组丢失之前消息可能就全部被传送了。如果所使用的初始速率太过,则可能需要重传已发送的全部消息。

[0004] 除了重传之外,分组丢失还可导致来自发送机设备的消息传输的临时暂停。当TCP/IP在一段时间内检测到某一数量的分组丢失时,就认为网络拥塞并且TCP/IP增大从正在经历分组丢失的发送机设备传送消息之间的延迟。这种延迟对于整体性能来说在成本上可能远高于重传。具有允许向许多接收方并行传送许多小的消息的充足带宽的发送机设备在使用TCP/IP时可能面临传输暂停的特定风险。

[0005] 完全二等分带宽网络的使用可以缓解核心中的拥塞,但分组丢失可能仍然在终端站处发生。如果TCP使用丢失作为拥塞信号,则可能仍然发生不必要的重传延迟。

[0006] 用于消息传送的另一种方法是在跨网络的两个设备之间建立电路。电路通常涉及在设备本身处以及在每一个中介设备处对设备之间的通信信道的带宽分配。然而,由于建立和拆除电路的高成本,作为许多网络应用的典范的短消息可能无法成为电路建立成本的正当理由。

发明内容

[0007] 通过分组交换网络进行通信的发送方和接收方设备通过交换请求来发送消息以及提供用于发送该消息的许可的响应来在他们自身之间建立有效电路。通过限制要并行传送的消息数量来确保发送方设备处的充足带宽,并且响应于接收方设备接收到请求传送消息在接收方设备处保留充足带宽。发送方和接收方设备之间的分组交换网络可以是完全二等分带宽网络或高二等分带宽网络,从而使得在中介设备处保留带宽变得不必要。接收方设备还可被配置成在接收到的请求将利用比接收方设备可用带宽更多的带宽时拒绝发送方请求。响应于其请求被拒绝,发送方设备可选择定向到不同接收方设备的不同消息,并且

可以向该不同接收方设备传送要发送该消息的请求。通过切换至不同的消息和不同的接收方设备,发送方设备确保了它不会处于空闲而是仍然有要传送的消息。

[0008] 提供本概述以便以简化形式介绍将在以下详细描述中进一步描述的一些概念。本概述并不旨在标识所要求保护主题的关键特征或必要特征,也不旨在用于限制所要求保护主题的范围。

附图说明

[0009] 参考附图来阐明详细描述,附图中,附图标记最左边的数字标识该附图标记首次出现的附图。在不同附图中使用同一附图标记指示相似或相同的项或特征。

[0010] 图1图解了根据各实施例的示例环境。

[0011] 图2a-2b示出了根据各实施例的发送方设备的示例队列。

[0012] 图3示出了根据各种实施例的接收方设备的示例队列。

[0013] 图4是示出了根据各实施例的用于通过分组交换网络与接收方设备建立有效电路的由发送方设备执行的多个操作的流程图。

[0014] 图5是示出了根据各实施例的用于基于对接收方设备带宽的保留通过分组交换网络与发送方设备建立有效电路的由接收方设备执行的多个操作的流程图。

[0015] 图6是根据各个实施例的示例计算机系统体系结构的框图。

具体实施方式

[0016] 本文描述了用于通过分组交换网络在发送方设备和接收方设备之间建立有效电路的技术。如本文所使用的,术语“有效电路”是指两个端点设备之间其中信道带宽被确保的通信信道。与传统电路相反,文本讨论的有效电路不要求用于建立的大量开销。通过利用完全二等分带宽或高二等分带宽的分组交换网络,在中介设备处进行带宽保留变得不必要。通过限制发送方设备处的消息传送以及在接收方设备处保留针对发送方设备的消息的带宽来确保有效电路的端点处的带宽。

[0017] 本公开利用了这样一个观测结论:在高或完全二等分带宽网络中,拥塞完全在终端站(发送方和接收方设备)处发生而不在网络核心中发生。因此,相比于观测分组丢失(作为来自网络核心的隐式信号)的传统方法,在网络边缘处进行带宽分配变得可行。

[0018] 为了保留接收方设备带宽,发送方设备向接收方设备传送针对该发送方设备需要发送的每一数据消息的请求发送(RTS)消息。发送方设备可以具有要传送到一个或多个接收方设备的多个数据消息,并且可以基于发送方设备用于传送数据消息的带宽来限制所发送的RTS消息的数量。在限制RTS消息的数量时,发送方设备可以对至少某些RTS消息采取消极响应并因而请求对原本共同将使用比发送方设备可用的带宽更高的带宽的数据消息的传送。在一些实施例中,发送方设备可以具有要传送给单个接收方设备的多个数据消息并且可以针对该多个数据消息发送一个RTS消息。

[0019] 接收方设备可以从多个发送方设备接收RTS消息。这些RTS消息可以请求对发送原本共同将使用比接收方设备可用的带宽更高的带宽的数据消息的许可。为了确保接收方设备只接收它有带宽来接收的那么多的同时或并发数据消息,接收方设备可以为一个或多个发送方设备保留用于传送他们的数据消息的带宽,并且用拒绝其他发送方设备请求发送数

据消息的消息来对他们的RTS消息作出响应(这些消极响应在此处被称为“繁忙稍后再试”(BTL)消息)。接收方设备用向发送数据消息提供许可(clearance)的消息来对已经为其保留了带宽的发送方设备作出响应(这些提供发送许可的消息在此处被称为“CTS”消息)。在选择要用CTS消息对哪些发送方设备作出响应时,接收方设备可以按照某种方式对RTS消息进行排序。例如,接收方设备可以基于RTS消息中包括的优先级或者基于先前发送给RTS消息的发送方设备的BTL消息的数量来对RTS消息进行排序。

[0020] 在接收到CTS消息后,在接收CTS消息的发送方设备和接收方设备之间建立有效电路。发送方设备随后可继续向接收方设备发送已经接收到对其的许可的数据消息。在一些实施例中,接收方设备可以仅仅用CTS消息对由发送方设备发送给该接收方设备的RTS消息的子集作出响应。

[0021] 响应于接收到BTL消息,发送方设备可以选择定向到不同接收方设备的不同数据消息,并且向该接收方设备发送针对该数据消息的RTS消息。在一些实施例中,数据消息可由发送方设备进行排队,并且发送方设备可以从该队列中选择被定向到不同接收方设备且先前尚未发送对其的RTS的第一消息。通过在一替换性的接收方设备处保留带宽,发送方设备确保了它不会具有空闲带宽而是仍然有要发送的数据消息。

[0022] 在一些实施例中,发送方设备不必针对每一个数据消息都发送RTS。作为控制消息的数据消息或者小于阈值大小的数据消息可被简单地发送给接收方设备而不建立有效电路。这进一步降低了该方案的总体开销。

[0023] 示例环境

[0024] 图1图解了根据各实施例的示例环境。如图1所示,发送方设备102可通过分组交换网络106与接收方设备104通信。发送方设备102可配备有网络模块108,该网络模块108从发送方设备102的应用112接收要传送到接收方设备104的消息110。网络模块108可将消息110排列在发送方设备102的队列114中。在将消息110传送到接收方设备104之前,网络模块108可经由发送方设备102的接口118向接收方设备104传送针对消息110的请求发送(RTS)消息116。这些RTS消息116随后可经由接收方设备104的接口120被接收并且由接收方设备104的网络模块122来处理。网络模块122可以将RTS消息116排列在接收方设备104的用于对RTS消息116进行排序的队列124中。接收方设备104随后可用提供消息110的发送许可(即,CTS消息)的消息126或者用拒绝RTS消息116的消息126(即,BTL消息)来对RTS消息116作出响应。接收到CTS消息126的发送方设备102随后可向接收方设备104的应用128传送消息110,消息110由发送方设备102的应用112被定向到那些应用128。接收到BTL消息126的发送方设备102可选择其他接收方设备104。

[0025] 尽管发送方设备102和接收方设备104在此处被示为分开的且不同的设备,但应当理解,任何计算设备可以既是一些消息110的发送方设备102又是其他消息110的接收方设备104。因而这样的计算设备将配备有发送方设备102的网络模块108、应用112和队列114以及配备有接收方设备104的网络模块122、队列124和应用128。被配置成既能够充当发送方设备102又能够充当接收方设备104的计算设备也可具有单个接口集(在这种情形中接口118和接口120将指代同一接口集),或者两个接口集(在这种情形中一个集合(接口118)可专用于发送消息110而另一集合(接口120)可专用于接收消息110)。

[0026] 在各个实施例中,发送方设备102和接收方设备104可以是任何类型的计算设备。

例如,发送方设备102和接收方设备104可以是或者包括个人计算机(PC)、膝上型计算机、服务器或服务器场、大型机、平板计算机、工作站、电信设备、个人数字助理(PDA)、媒体播放器、媒体中心设备、个人视频记录器(PVR)、电视机、或任何其他类型的一个或多个设备。在一些实现中,发送方设备102和接收方设备104中的一者或多者代表在一个或多个计算设备上实现的一个或多个虚拟机。本文将进一步描述发送方设备102和接收方设备102的模块和数据的特性和功能。用于实现发送方设备102和接收方设备104中的一者或两者的示例计算设备在图6中示出,并且在下文中将参考该附图进行更详细地描述。

[0027] 在各个实现中,网络106可以是任何一种或多种类型的网络,诸如广域网(WAN)、局域网(LAN)或因特网。而且,网络106可以是公共的、私有的,或者包括公共和私有网络两者。此外,网络106可以是有线的、无线的,或者包括有线和无线网络两者。网络106可利用任何一个或多个协议来进行通信,诸如TCP/IP、其他基于分组的协议、或其他协议。在一个实施例中,网络106可包括连接云计算环境的多个节点的云计算环境。在这样的一个实施例中,发送方设备102和接收方设备104可以是云计算环境的节点。

[0028] 如上文所提及的,网络106可以是分组交换网络。这样的网络106还可以是完全二等分带宽或高二等分带宽网络,从而允许网络的任何节点以全速与任何其他节点通信而没有中介节点引入的任何传输延迟。在其他实施例中,网络106可以是无限带宽网络。网络106还可包括一个或多个交换机。该交换机可以是任何类型的交换机,并且可以各自包括网络接口组件,诸如传入和传出网络接口组件,每一个网络接口组件具有带宽。例如,交换机可具有多个传入以太网端口和传入无线端口,以及传出以太网和无线端口。在一些实施例中,交换机的传入带宽与交换机的传出带宽成比例。例如,为网络分层结构中交换机“下方”的设备提供服务的网络接口的总传入带宽可以是每秒10GB,而上至核心交换机的网络接口组件的总带宽也可以是每秒10GB。通过使交换机的传入带宽与传出带宽成比例,避免引入与交换机相关联的瓶颈。这种具有成比例带宽的交换机在2009年3月25日提交的题为“Data Center Without Structural Bottlenecks(没有结构性瓶颈的数据中心)”的No.12/410,697美国专利申请、在2009年3月25日提交的题为“Data Center Interconnect and Traffic Engineering(数据中心互连和流量工程)”的No.12/410,745美国专利申请以及在2009年10月14日提交的题为“Agile Data Center Network Architecture(灵活的数据中心网络架构)”的No.12/578,608美国专利申请中更详细地描述。

[0029] 在一些实施例中,如图1中所示,发送方设备102和接收方设备104可各自包括应用。发送方设备102可包括应用112而接收方设备104可包括应用128。应用112和128可以是能够读取和写入数据(诸如数据消息110)的任何类型的应用。应用112和128可以是相同类型的应用或不同类型的应用。而且,应用112可以代表多个应用而应用128可以代表对应的多个应用。应用112可以是客户机、服务器或者应用128的对等体。从应用112和128的角度,交换的数据消息110被直接发送而不必知晓RTS/CTS消息交换或有效电路。

[0030] 在各个实施例中,数据消息110可包括任何种类的任何类型的数据,并且可以具有任意大小。每一个数据消息110可被分组化到一个或多个分组中以供由发送方设备102的应用112、网络模块108或某一其他组件(诸如发送方设备102的平台或网络堆栈)通过网络106来传送。包括数据消息110的分组的数量可以随着数据消息110的大小以及用于传送数据消息110的网络协议而变化。除了其数据之外,每一个数据消息110还可与元数据相关联。这样

的元数据可包括对数据消息110被定向到的接收方设备104的标识,并且如果数据消息110是数据消息110组的一部分,则元数据还可包括标识了数据消息110在该组内的位置的序列号。

[0031] 在一些实施例中,网络模块108可包括任何一个或多个模块、应用、进程、线程或功能。网络模块108可以是发送方设备102的平台或发送方设备102的网络堆栈的组件,或者可以独立于其他发送方设备102模块和应用。如上文所提及的,网络模块108可被配置成执行以下各项中的至少一项:接收数据消息110、使接收到的数据消息110分支、使数据消息110排队、选择供传输的数据消息110、对所选的数据消息110传送RTS消息116、作为响应接收CTS消息126或BTL消息126、响应于CTS消息126来传送数据消息110、或者响应于BTL消息126来选择替换性数据消息110。除了这些操作之外,网络模块108还可管理发送方设备102的接口118。这些接口118可由网络模块108选择以供传输RTS消息116和数据消息110。而且,网络模块108可并发地对任何数量的数据消息110执行这些操作。例如,网络模块108可从应用112接收数据消息110同时传送针对另一数据消息110的RTS消息116。

[0032] 在一些实施例中,在从应用112接收到数据消息110后,网络模块108可确定该数据消息110是控制消息还是大小小于阈值的消息以便对接收到的数据消息110进行分叉。网络模块108将作为控制消息或小于阈值的数据消息110提供给发送方设备102的网络堆栈以供使用发送方设备102的接口118进行传输。不是控制消息且大于阈值的数据消息110随后由网络模块108进行排队以供后续传输。

[0033] 在各个实施例中,网络模块108将接收到的数据消息110存储在一个队列114或若干队列114中。发送方设备102可以具有其中存储了所有接收方设备104的数据消息110的单个队列114,或者可以具有与每一个接收方设备104相关联的队列114。在其他实施例中,发送方设备102可利用与接收方设备104具有其他关联关系的其他数量的队列114。

[0034] 例如,图2a解说了用于存储定向到多个接收方设备104的数据消息110的队列114的示例。如图2a中所示,队列202(其可以作为队列114的示例)包括多个数据消息110的条目204。尽管这些条目204被示为表的各行,但应该理解,条目204也可被表示为各列或者采用某种非表格形式的项。每一个条目204包括消息标识符206、是否发送了针对由该条目204表示的数据消息110的RTS消息116的指示符208、以及目标接收方设备104的标识符210。每一个消息标识符206可以是唯一地标识了数据消息110的任何类型的数据值,诸如整数或串。每一个指示符208可以例如是指示了是否发送了RTS消息116的二进制值或其他值类型的有限集合之一。每一个对目标接收方设备105的标识符210可以是接收方设备104的网络标识符,诸如IP地址或其他形式的地址。

[0035] 图2b解说了用于存储数据消息110的多个队列114的示例,其中一个队列114针对多个接收方设备104中的每一个设备。如图2b中所示,队列212、214、和216中的每一个队列(其可以作为队列114的示例)包括一个或多个数据消息110的条目204。尽管这些条目204被示为表的各行,但应该理解,条目204也可被等同地表示为各列或者采用某种非表格形式的项。每一个条目204包括消息标识符206、以及是否发送了针对由该条目204表示的数据消息110的RTS消息116的指示符208。不需要存在表示数据消息110的目标接收方设备104的标识符210的字段,因为队列212、214和216中的每一个队列与特定的接收方设备104相关联。每一个消息标识符206可以是唯一地标识了数据消息110的任何类型的数据值,诸如整数或

串。每一个指示符208可以例如是指示了是否发送了RTS消息116的二进制值或其他值类型的有限集合之一。

[0036] 队列114可以是本领域公知的任何类型的数据结构,包括模块的数据结构,诸如网络模块108或其他文件或数据库。而且,队列114可以是如本领域所公知的先进先出(FIFO)数据结构,或者可以按照不同顺序将数据消息110添加到队列114或从队列114中移除。例如,数据消息110可被添加到队列114末端但可以按照不同顺序从队列114中被移除。在这样的一个示例中,数据消息110在队列中的位置可以是在选择数据消息110并从队列114中移除数据消息110以供传输时的多个因素之一。

[0037] 在各个实施例中,网络模块108响应于被添加到队列114的数据消息110、响应于估算了发送方设备102将完成其他数据消息的传输的时间的计时器、或者响应于来自发送方设备102的平台的输入来从队列114中的一个或多个队列中选择一个或多个数据消息110。网络模块108可以基于一个或多个因素来从队列114中选择数据消息110。这些因素可包括数据消息110在队列114中的位置、先前从哪个队列114中选择了数据消息110、数据消息110的大小、发送方设备接口118的可用带宽、将多个数据消息110链接在一起以供传输到单个接收方设备104的能力。这些因素仅仅出于说明的目的而提供,并且在替换性实施例中可以使用任何其他数量的因素。

[0038] 在一些实施例中,网络模块108首先确定发送方设备接口118的可用带宽。该可用带宽可以是网络模块108知晓的预设数字,诸如表示接口118的总带宽的百分之八十或百分之九十的数字。在其他实施例中,该可用带宽可以变化并且可以在不同时间被重复地确定。

[0039] 网络模块108随后可以选择将尽可能多的利用可用带宽来进行传输的数据消息110。在一个实施例中,网络模块108从一个或多个队列114的头部选择数据消息110而不参考接收方设备104,直到所选数据消息110的共同大小逼近可用带宽为止。在发送方设备102具有针对多个接收方设备104的多个队列114的其他实施例中,网络模块108可以按照循环方式来向队列提供服务,从而从要向其提供服务的下一队列114中选择数据消息110。在其他实施例中,优选于选择单个接收方设备104的多个数据消息110,网络模块108可以将单个接收方设备104的数据消息110进行链接以便减少接收方设备所需的带宽保留的数目。在这样的实施例中,网络模块108跟踪已经向其传送了经链接数据消息110的接收方设备104以便阻止单个接收方设备104支配消息传输。在还有一些其他实施例中,网络模块108对先前已经发送过RTS消息116的数据消息110以及已经超过阈值时间量的数据消息110排定优先级。等待阈值时间量的原因在于允许提供了BTL消息126的接收方设备104完成其对其他数据消息110的接收并且释放其保留的带宽。

[0040] 在一些实施例中,通过在选择数据消息110时考虑接口118的可用带宽,网络模块108限制了可被并行传送的数据消息110的数量以便避免发送方设备102上的传输瓶颈。然而,在多个所选数据消息110的RTS消息116可被拒绝的情况下,网络模块108可以选择比用于传送的可用带宽更多数量的数据消息110以便确保对可用带宽的完全利用。

[0041] 在各个实施例中,在选择数据消息110后,网络模块108生成并传送所选数据消息110的RTS消息116。网络模块108可针对每一个所选数据消息110生成和传送RTS消息116,或者可以针对要被并行传送到单个接收方设备104的一组数据消息110生成和传送RTS消息116。RTS消息116可以是标识了与其相关联的特定数据消息110、与该RTS消息相关联的一个

或多个数据消息110的一个大小或多个大小、对目标接收方设备104的标识符和/或消息优先级的小消息。消息优先级可由网络模块108基于针对RTS消息的目标接收方设备104排队的数据消息110与针对其他接收方设备104排队的数据消息110之比以及一个或多个队列114的总大小来确定。在其中RTS消息116与数据消息110链相关联的实施例中,RTS消息116可标识针对该接收方设备104的数据消息110的数量。

[0042] 在生成RTS消息116后,网络模块108选择一个或多个可用接口118以供传输RTS消息116。接口118可以是本领域公知的任何类型的网络接口,诸如用于WAN或LAN的有线或无线接口、用于PAN或蓝牙®网络的无线接口、或任何其他类型的接口。例如,接口118可包括一个或多个网卡。这些接口118可包括由网络模块108或另一组件管理的接口池作为共享相同队列集和队列状态(即,队列114)的共同资源。在一个实施例中,接口118的带宽可被分配给不同进程,包括网络模块108的一个或多个进程。这可包括由多个进程对单个网络接口118的独立使用。只要这些进程提前静态地划分了网络带宽,这些进程就可使用所分配的带宽而无需协调。在选择一个或多个可用接口118后,网络模块108通过一个或多个可用接口118来传送RTS消息116。

[0043] 响应于发送RTS消息116,发送方设备102接收CTS消息126、BTL消息126或两者。对于发送方设备102发送的每一个RTS消息116,可向其返回一个CTS消息126或BTL消息126。

[0044] 在经由接口118接收到CTS消息126后,网络模块108确定CTS消息126向哪个数据消息110或哪些数据消息110提供许可。如上文所提及的,从接收方设备104接收到CTS消息126在发送方设备102和该接收方设备104之间建立了有效电路。CTS消息126可通过消息标识符206来标识特定数据消息110,或者可以指代特定RTS消息116,该特定RTS消息116可进而由网络模块108将其与特定数据消息110相关。CTS消息126还可标识被许可传送到发送了该CTS消息126的接收方设备104的多个数据消息110。例如,CTS消息126可响应于请求将若干数据消息110发送到接收方设备104的许可的RTS消息116。这样的CTS消息126可向RTS消息116中标识的一个、一些或全部数据消息110提供发送许可。

[0045] 在各个实施例中,在确定由CTS消息126标识的数据消息110后,网络模块108可选择用于传送数据消息110的接口118中的一个或多个接口并且经由所选接口118来传送数据消息110。在一个实施例中,用于传送数据消息110的接口118可以不同于用于传送针对该数据消息110的RTS消息116的接口118。在传送数据消息110后,网络模块108可将数据消息110从一个或多个队列114中移除。

[0046] 作为接收到一个或多个CTS消息126的补充或替代,发送方设备102可经由接口118接收到一个或多个BTL消息126。在接收到BTL消息126后,网络模块108确定该BTL消息126拒绝了哪个数据消息110或哪些数据消息110。网络模块108随后可在队列114中更新关于那些数据消息110的指示符208以反映针对那些数据消息110的RTS消息116已经被拒绝。

[0047] 网络模块108随后可继续选择以供传送的一个或多个替换性数据消息110来代替被拒绝的数据消息110。在一些实施例中,在选择替换性数据消息110时,网络模块108可利用用于选择现在拒绝的数据消息110的上述因素。在选择一个或多个替换性数据消息110时,网络模块108还可考虑被拒绝的数据消息110的大小或多个被拒绝的数据消息110的共同大小。而且,在选择一个或多个替换性消息时,网络模块108可过滤掉针对其的RTS消息116已经被拒绝的数据消息110以及定向到从其接收了BTL消息126的接收方设备104的数据

消息110。一旦选择了替换性数据消息110,网络模块108就生成该替换性数据消息110的RTS消息116,选择在传送该RTS消息116时使用的接口118,并且传送该RTS消息116。通过选择替换性数据消息110而非等待传送先前所选的数据消息110,网络模块108确保了发送方设备102不会处于空闲而是仍然有等待传送的数据消息110。

[0048] 在一些实施例中,如上文所提及的,发送方设备102可利用计时器来确定何时选择供传送的其他数据消息110。计时器可以是网络模块108的一部分或者是单独的组件。计时器要么在数据消息110被选择时要么在数据消息110的传输开始时被发起,并且计时器基于可用带宽和数据消息110的大小来估算用于传送数据消息110的时间。一旦到达该估算时间,计时器向网络模块108通知已经到达时间以使得网络模块108能够选择供传输的其他数据消息110。

[0049] 在各个实施例中,网络模块116可包括任何一个或多个模块、应用、进程、线程或功能。网络模块116可以是接收方设备104的平台或接收方设备104的网络堆栈的组件,或者可以独立于其他接收方设备104模块和应用。如上文所提及的,网络模块116可被配置成执行以下各项中的至少一项:接收RTS消息116、对RTS消息116排序、为RTS消息116的一个或多个发送方设备102保留带宽、生成CTS消息126并将CTS消息126发送到已经为其保留了带宽的发送方设备102、生成BTL消息126并且将BTL消息126发送到其他发送方设备102、跟踪BTL消息126已经被发送到一发送方设备102的次数、以及使用保留带宽来接收数据消息110。而且,网络模块116可并发地对任何数量的数据消息110执行这些操作。例如,网络模块116可接收数据消息110同时响应于另一数据消息110的RTS消息116来传送CTS消息126。

[0050] 在一个实施例中,当网络模块116经由接口120接收到RTS消息116时,网络模块116可将这些RTS消息在队列124中排队以进行排序。RTS消息116可接收自多个发送方设备102,并且接收那些RTS消息116的接口120可以是任何类型的接口,诸如与上述接口118相同或不同类型的接口。队列124可以是任何类型的数据结构,诸如网络模块116的数据结构。

[0051] 例如,图3解说了包括用于对接收到的RTS消息116进行排序的RTS消息条目的队列124的示例。如图所示,队列302(其可以作为队列124的示例)包括多个RTS消息116的条目304。尽管这些条目304被示为表的各行,但应该理解,条目304也可被等同地表示为各列或者某种非表格形式的项。每一个条目304包括消息标识符306、响应于与特定数据消息110相关联的RTS消息116被发送到发送方设备102的BTL消息126的计数308、对与条目304的RTS消息116相关联的数据消息110的优先级的指示符310、以及对发送了条目304的RTS消息116的发送方设备102的标识符312。每一个消息标识符306可以是唯一地标识了RTS消息116或与RTS消息116相关联的数据消息110的任何类型的数据值,诸如整数或串。每一个计数308可以从接收方设备104的另一数据结构检索的整数或某一其他数值,该计数用于跟踪响应于与特定数据消息110相关联的RTS消息116而被发送的BTL消息126。当接收到RTS消息116时,网络模块116检查该数据结构以寻找与由接收到的RTS消息116所指代的数据消息110相关联的BTL消息126的数量。每一个指示符310可以例如是指示了是否被排定优先级的二进制值或其他值类型的有限集合之一。另选地,如果有若干优先级水平可用,则每一个条目304的指示符310可指示数据消息110的特定优先级水平。每一个对发送方设备102的标识符312可以是发送方设备102的网络标识符,诸如IP地址或其他形式的地址。

[0052] 在各个实施例中,网络模块116可基于任何数量的因素对队列124中的RTS消息116

进行排序。例如,网络模块116可基于消息优先级和/或基于上文提及的BTL消息126的计数来对RTS消息116排序。网络模块116还可考虑接收方设备可供保留的带宽以及与RTS消息116对应的数据消息110的大小。

[0053] 基于该排序,网络模块可为与接收到的RTS消息116相关联的一个或多个发送方设备102保留带宽。这些可以是与排序最高的RTS消息116对应的发送方设备102。在一些实施例中,网络模块116可为发送方设备102应用最大带宽保留从而使得没有发送方设备102可被分配到接收方设备102的整个可用带宽。该最大带宽保留可阻止发送方设备102垄断接收方设备带宽。

[0054] 为了使带宽保留有效,网络模块116将CTS消息126发送到为其保留了带宽的发送方设备102。CTS消息126可标识它们对其作出响应的RTS消息116或者向其提供许可的数据消息110。在一些实施例中,在RTS消息116标识多个数据消息110的情形中,CTS消息126还可标识多个数据消息110的数量或身份。

[0055] 网络模块116随后可响应于未导致带宽保留和CTS消息126响应的RTS消息116来生成和发送BTL消息126。BTL消息126可标识它们对其作出响应的RTS消息116或者对其拒绝许可的数据消息110。在一些实施例中,在RTS消息116标识多个数据消息110的情形中,BTL消息126还可标识多个数据消息110的数量或身份。在发送BTL消息126之后,网络模块116可更新如上所讨论的存储在接收方设备104的数据结构中的特定数据消息110的BTL消息的计数。在用CTS消息126或BTL消息126对队列124中的每一个条目304作出响应之后,网络模块116还可转储清除队列124中的条目304。

[0056] 在各个实施例中,网络模块116随后可使用被保留的带宽来接收数据消息110。因为接收方设备104仅接收它向其提供了发送许可的数据消息110,防止了接收方设备104处的缓冲器溢出的风险。在接收到数据消息110后,网络模块116将数据消息110提供给它们被定向到的应用128并且释放与数据消息110相关联的带宽保留。在一些实施例中,完成接收数据消息110和释放带宽保留会触发发送下一CTS消息126。

[0057] 示例操作

[0058] 图4和5是解说了各过程的示例操作的流程图。这些过程的示例操作在单独的框中示出,并且参考这些框来概括。这些过程被示为逻辑流程图,每个操作可表示可用硬件、软件或其组合实现的一组操作。在软件的上下文中,这些操作表示存储在一个或多个计算机可读存储介质上的计算机可执行指令,这些指令在由一个或多个处理器执行时执行既定操作。一般而言,计算机可执行指令包括执行特定功能或实现特定抽象数据类型的例程、程序、对象、组件、数据结构等。描述操作的次序并不旨在解释为限制,并且任何数量的所述操作可以按任何次序和/或并行组合以实现该过程。

[0059] 图4是示出了根据各实施例的用于通过分组交换网络与接收方设备建立有效电路的由发送方设备执行的多个操作的流程图。如框402所示,发送方设备可接收以供从该发送方设备的一个或多个应用传送到接收方设备的消息。发送方设备和接收方设备可由分组交换网络连接,诸如完全二等分带宽网络或高二等分带宽网络。在框404,在接收到消息后,发送方设备可确定该消息是否是控制消息或具有小于阈值大小的大小。在框406,如果该消息是控制消息或小于阈值大小,则发送方设备可将该消息传送到该消息所指定的接收方设备。在框408,如果消息不是控制消息或大于阈值大小,则发送方设备将消息排队以供后续

传送。

[0060] 在框410,发送方设备选择供传送到接收方设备的消息。发送方设备可从发送方设备队列中选择消息,诸如队列中的第一消息。在一些实施例中,发送方设备可选择以供传送的多个消息,诸如供链接传送到同一接收方设备的多个消息或到多个接收方设备的多个消息。在框410a,发送方设备基于发送方设备的可用传送带宽来限制所选消息的数量,从而使得全部所选消息可并发地被传送而不必等待任何其他所选消息完成发送。在框410b,发送方设备可从队列中选择与先前的发送请求(RTS)相关联的消息。发送方设备可具有针对该消息的先前发送的RTS,并且可接收到拒绝发送该消息的许可的响应。

[0061] 在框412,发送方设备发送针对所选消息的RTS。该RTS可包括消息的大小。在一些实施例中,发送包括:经由发送方设备的多个接口发送多个请求,该多个接口共享同一消息队列和队列状态集合。在一个实施例中,RTS可经由不同于消息的接口来发送。在框412a,发送方设备并行地将多个RTS发送到多个接收方设备。在框412b,发送方设备将前往单个接收方设备的RTS链接,并发地将经链接的RTS发送到接收方设备。在412c,链接可包括:在单个RTS中指示该发送方设备需要向该接收方设备发送的消息的数量。

[0062] 响应于该RTS,发送方设备在框414接收发送许可(CTS)或在框422接收拒绝该RTS的响应(此处被称为“繁忙稍后再试”(BTL)消息)。在框414,发送方设备从接收方设备接收CTS。RTS和CTS在通过分组交换网络在发送方设备和接收方设备之间建立了有效电路。接收方设备将CTS消息发送到为其保留了带宽的发送方设备。在框416,发送方设备确定哪些消息被CTS许可。例如,如果发送方设备在框412c指示发送方设备具有要传送的十条消息,则接收方设备可能仅用许可了消息中的五条消息的一个或多个CTS来作出响应。在框418,发送方设备将被许可的一个或多个消息发送给接收方设备,该接收方设备将使用被保留的带宽来接收该被许可的一个或多个消息。通过利用发送方设备的全部带宽以及接收方设备的保留带宽,发送方设备确保了消息在没有任何延迟、分组丢失或未使用的带宽的情况下被发送。

[0063] 在框420,发送方设备利用计时器来估算何时应该将下一消息传送到接收方设备。计时器估算要发送所选的一个或多个消息的时间以及因而估算消息传送将何时完成以及发送方设备带宽将何时可用的时间。

[0064] 在框422,代替接收到CTS,发送方设备从接收方设备接收到拒绝请求发送消息的BTL。在框424,响应于BTL,发送方设备选择要发送给另一接收方设备的另一消息。在一些实施例中,发送方设备选择队列中其RTS尚未被发送且被定向到另一接收方设备的第一消息。随后在框412发送其他消息的RTS消息。发送方设备在RTS被拒绝的情况下选择另一消息进行传送以确保该发送方设备不会静止处于空闲而是仍然有要传送的消息以及传送那些消息的可用带宽。

[0065] 图5是示出了根据各实施例的用于基于接收方设备带宽的保留通过分组交换网络与发送方设备建立有效电路的由接收方设备执行的多个操作的流程图。如框502所示,接收方设备可从多个发送方设备接收到多个请求发送(RTS)消息。在一些实施例中,RTS消息包括发送方设备正请求要发送的消息的大小。在框504,在接收到这些RTS消息后,接收方设备可至少基于这些RTS消息中的一个或多个消息中包括的优先级或者发送方设备已经接收到拒绝其RTS消息的响应的次数中的一者来对RTS消息进行排序。

[0066] 在框506,接收方设备可为至少一个发送方设备保留用于从该至少一个发送方设备接收消息的带宽。该至少一个发送方设备可使其RTS消息选择自经排序的RTS消息。在一些实施例中,接收方设备保留带宽用于从接收方设备接收与多个RTS消息对应的多个消息。而且,在其他实施例中,为其保留的带宽的发送方设备的数量可以与正在请求发送的发送方设备的消息大小有关。在框508,接收方设备随后对该至少一个发送方设备的一个或多个RTS消息作出响应,从而向该至少一个发送方设备提供使用保留的带宽的一个或多个发送许可(CTS)消息。

[0067] 在框510,接收方设备用拒绝其他发送方设备的RTS消息的响应来对其他发送方设备作出响应。这些消极响应在此处被称为“繁忙稍后再试”(BTL)消息。该BTL消息向该其他发送方设备指示它们应该将RTS消息发送给其他接收方设备以便完全利用它们的发送方设备的带宽。在一些实施例中,当接收方设备将BTL消息发送给发送方设备时,它使发送给该发送方设备的BTL消息的数量表示递增。该表示随后被用于后续排序操作。

[0068] 在框512,接收方设备使用保留的带宽从该至少一个发送方设备接收一个或多个消息。

[0069] 示例系统体系结构

[0070] 图6是根据各个实施例的用于发送方设备102或接收方设备104的示例计算机系统体系结构的框图。如图所示,发送方设备102/接收方设备104可至少包括存储器602(包括高速缓存存储器)以及一个或多个处理单元(或处理器)604。处理单元604可在适当时用硬件、软件、固件、或其组合来实现。处理器604的软件或固件实现可包括用任何合适的编程语言编写的、用于执行所述各种功能的计算机或机器可执行指令。处理器604还可以或者替换地包括一个或多个图形处理单元(GPU)。

[0071] 存储器602可以存储可在处理器604上加载并执行的程序指令以及在执行这些程序期间生成的数据。取决于计算设备的配置和类型,存储器602可以是易失性的(诸如随机存取存储器(RAM))和/或非易失性的(诸如只读存储器(ROM)、闪存等)。计算设备或服务器还可包括附加可移动存储606和/或不可移动存储608,其包括但不限于磁盘、光盘和/或磁带。盘驱动器及其相关联的计算机可读介质可为计算设备提供对计算机可读指令、数据结构、程序模块以及其它数据的非易失性存储。在某些实现中,存储器602可包括多种不同类型的存储器,如静态随机存取存储器(SRAM)、动态随机存取存储器(DRAM)、或ROM。

[0072] 计算机可读介质包括至少两种类型的计算机可读介质,即计算机存储介质和通信介质。

[0073] 计算机存储介质包括用于存储诸如计算机可读指令、数据结构、程序模块或其他数据的任何方法或技术来实现的易失性和非易失性、可移动和不可移动的介质。计算机存储介质包括,但不限于,RAM、ROM、可擦除可编程只读存储器(EPROM)、闪存或者其他存储技术、紧致盘只读存储器(CD-ROM)、数字多功能盘(DVD)或者其他光存储、磁带盒、磁带、磁盘存储或者其他磁存储设备、或者能够用来存储信息以供计算设备访问的任何其他非传输介质。

[0074] 相反,通信介质可用诸如载波或其他传输机制等已调制数据信号来体现计算机可读指令、数据结构、程序模块或其他数据。如此所定义的,计算机存储介质不包括通信介质。

[0075] 发送方设备102/接收方设备104还可包含允许发送方设备102/接收方设备104与

所存储的数据库、另一计算设备或服务器、用户终端、和/或网络上的其它设备进行通信的通信连接610,诸如网络106。发送方设备102/接收方设备104还可包括输入设备612,如键盘、鼠标、笔、语音输入设备、触摸输入设备等等,和输出设备614,如显示器、扬声器、打印机等等。

[0076] 发送方设备102/接收方设备104可进一步包括用于跨网络106发送和接收数据的接口118/120。如上文所提及的,接口118/120可包括任何数量的网络接口,每一个网络接口具有用于传送和接收数据的相关联的带宽。

[0077] 转到更详细的存储器602的内容,存储器602可包括平台616。平台616可包括操作系统和/或一个或多个应用程序或服务。如果存储器602属于发送方设备102,则存储器602还可包括网络模块108和应用112以及一个或多个队列114,该网络模块108和应用112可各自表示任何一个或多个模块、应用、进程、线程或功能,该队列114可表示任何存储格式或数据结构。网络模块108、应用112和队列114在上文中更详细地进行了描述。如果存储器602属于接收方设备104,则存储器602还可包括网络模块122和应用128以及一个或多个队列124,该网络模块122和应用128可各自表示任何一个或多个模块、应用、进程、线程或功能,该队列124可表示任何存储格式或数据结构。网络模块122、应用128和队列124在上文中更详细地进行了描述。

[0078] 结语

[0079] 尽管已经用结构特征和/或方法动作专用的语言描述了本主题,但要理解,所附权利要求书中定义的主题不必限于所描述的具体特征或动作。相反地,这些具体特征和动作是作为实现该权利要求的示例性形式而公开的。

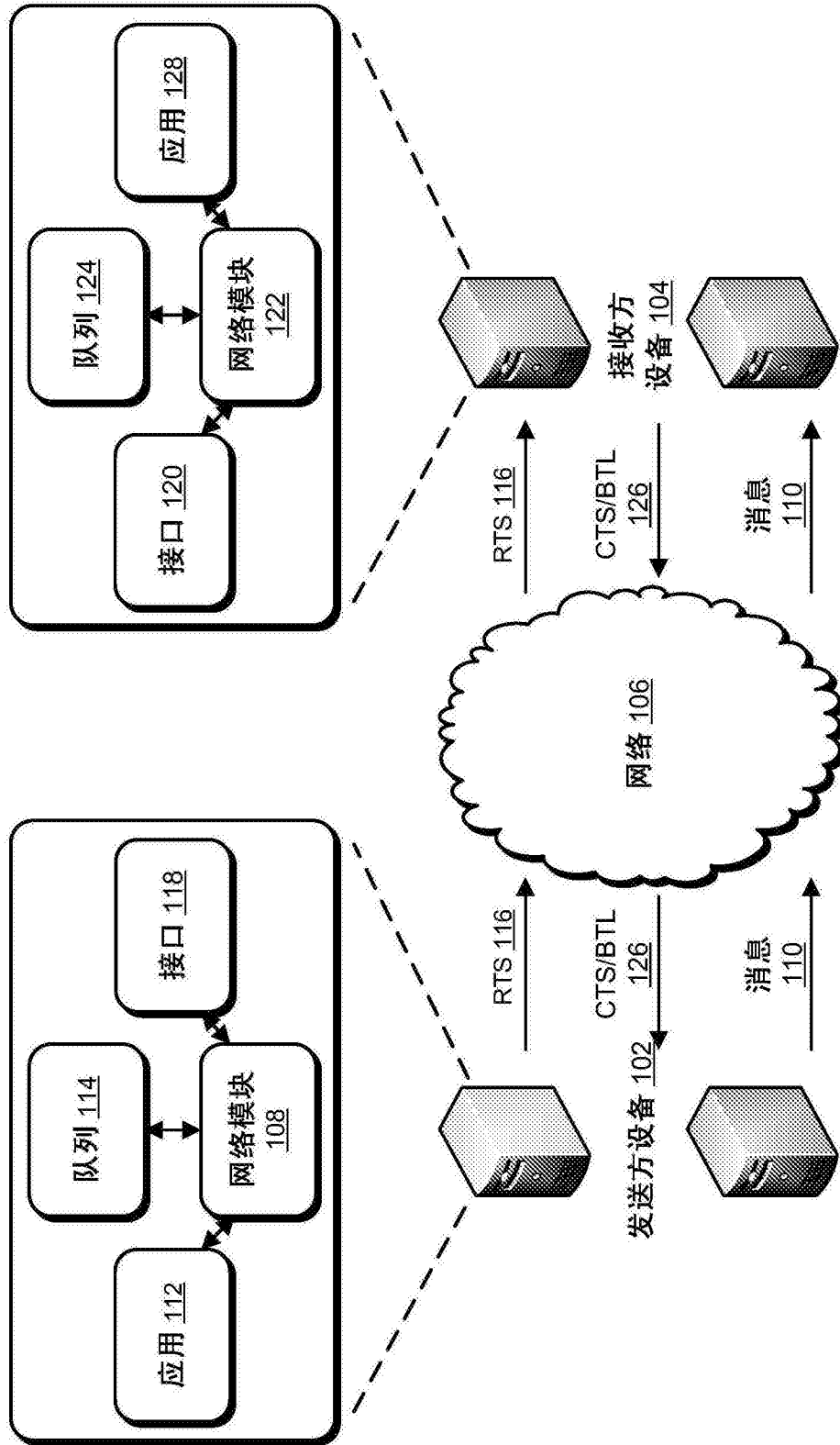


图1



图2a

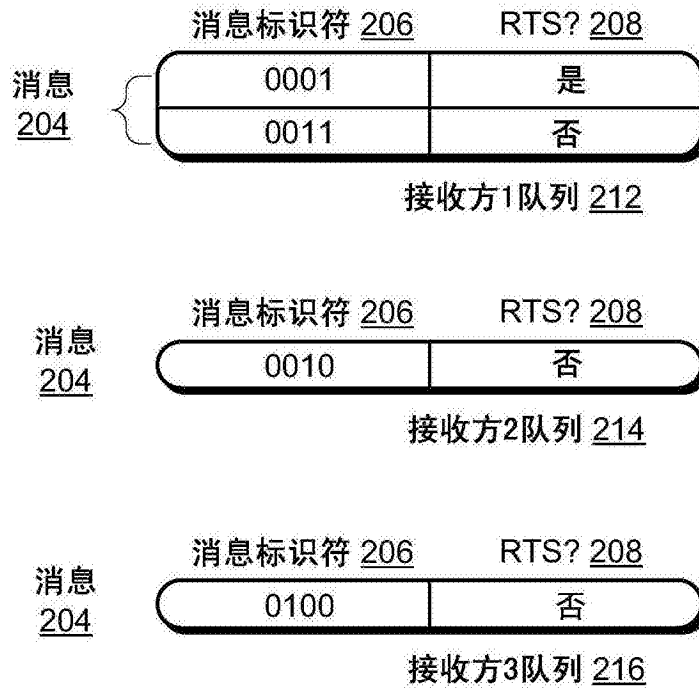


图2b



图3

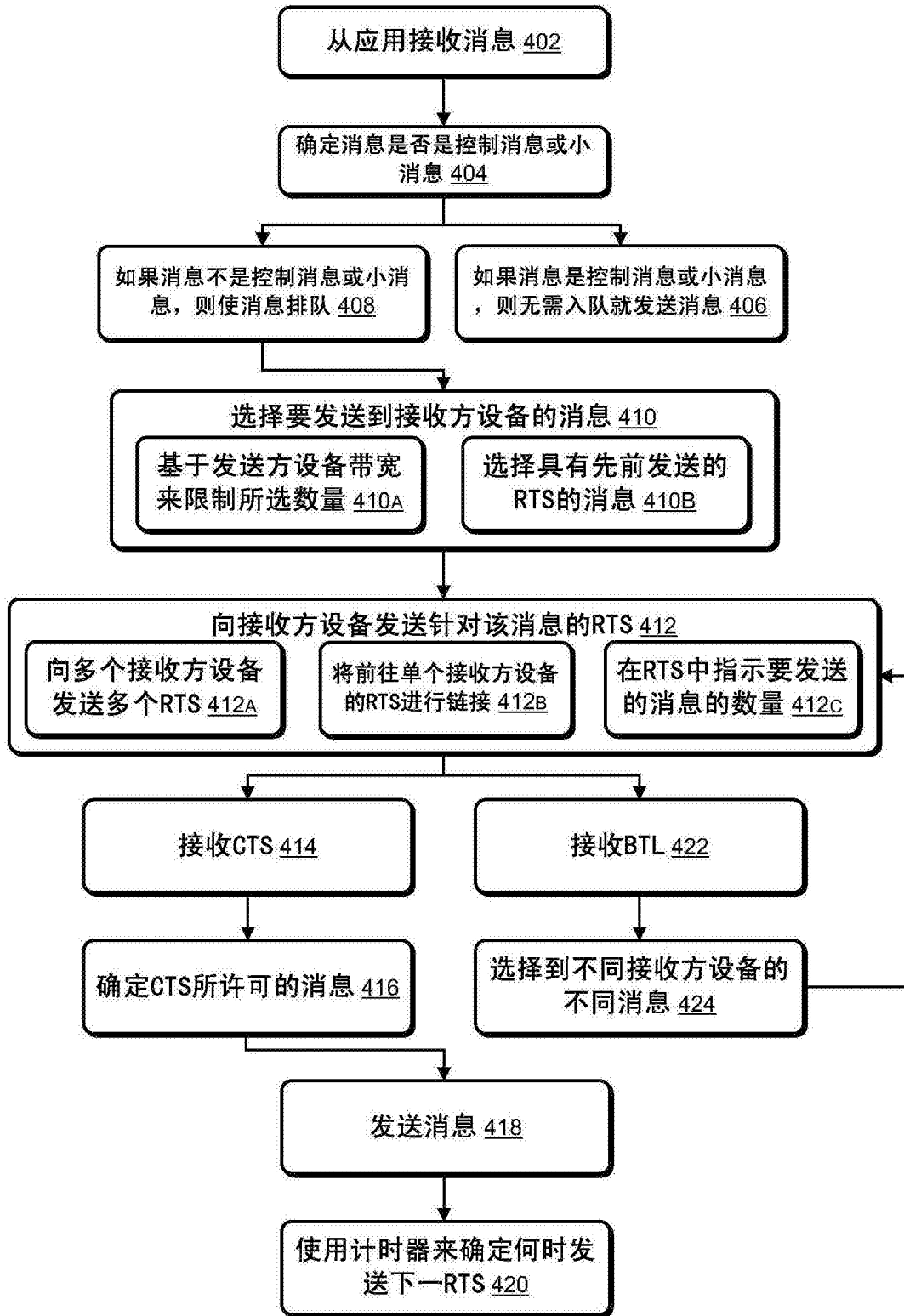


图4

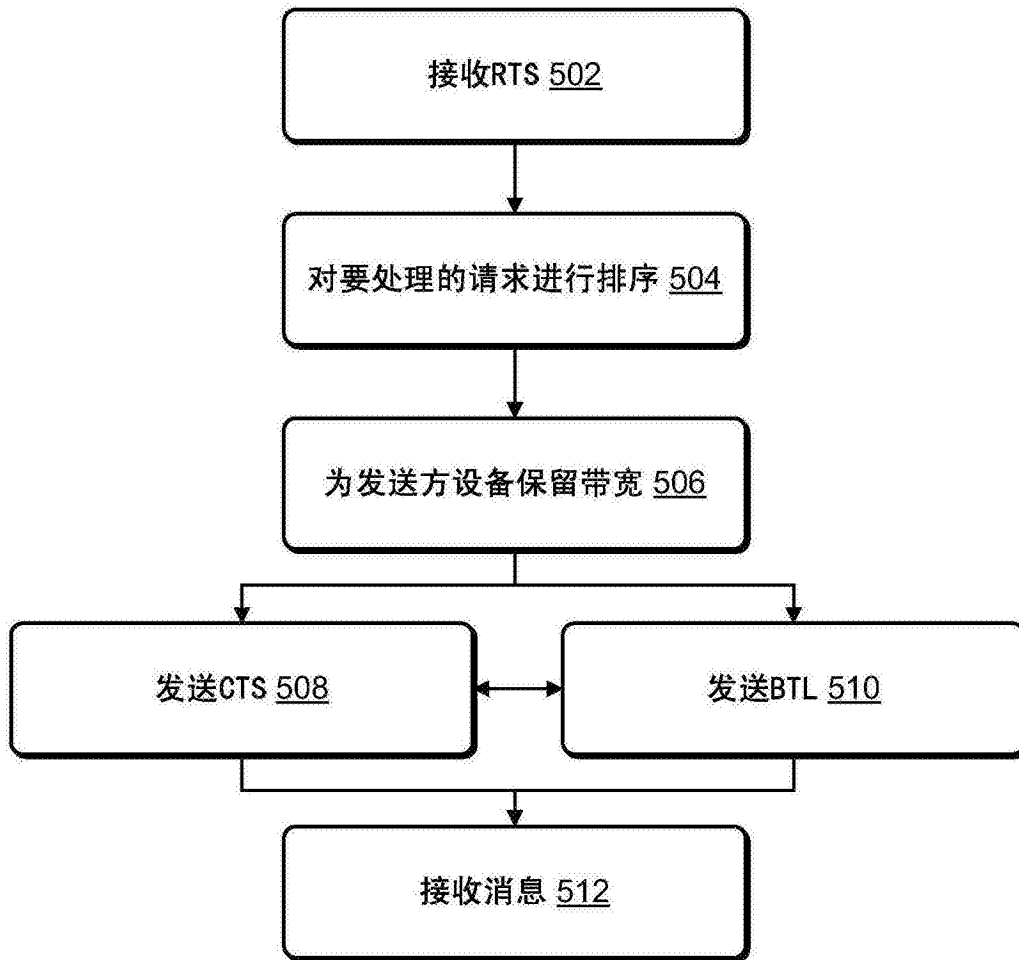


图5

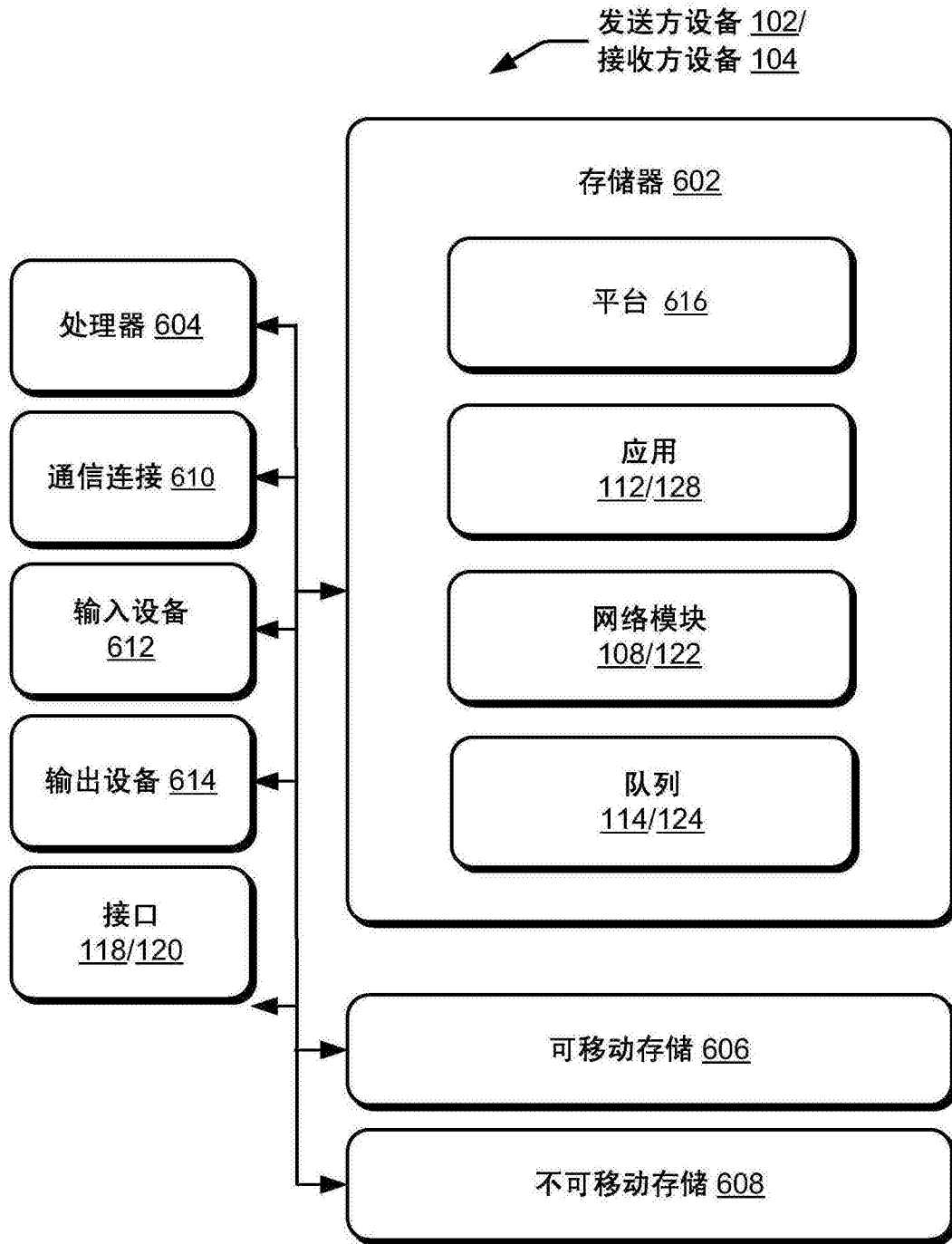


图6