

[19] 中华人民共和国国家知识产权局

[51] Int. Cl.  
H04L 12/56 (2006.01)



# [12] 发明专利申请公布说明书

[21] 申请号 200380111052.6

[43] 公开日 2007年2月7日

[11] 公开号 CN 1910868A

[22] 申请日 2003.12.23  
[21] 申请号 200380111052.6  
[86] 国际申请 PCT/EP2003/014788 2003.12.23  
[87] 国际公布 WO2005/064861 英 2005.7.14  
[85] 进入国家阶段日期 2006.8.21  
[71] 申请人 艾利森电话股份有限公司  
地址 瑞典斯德哥尔摩  
[72] 发明人 P·施利瓦-伯特林  
H·埃克斯特伦 R·路德维希

[74] 专利代理机构 中国专利代理(香港)有限公司  
代理人 杨凯 王忠忠

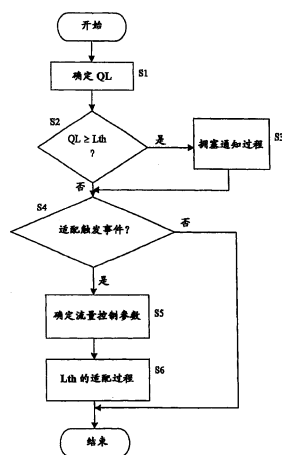
权利要求书 7 页 说明书 13 页 附图 4 页

## [54] 发明名称

用于控制队列缓冲器的方法及装置

## [57] 摘要

控制队列缓冲器的方法及装置，其中，如果队列长度参数超过门限值，则执行拥塞通知过程，所述门限值通过考虑流量控制参数的过程来自动适配。



1. 一种控制队列缓冲器(20)的方法, 所述队列缓冲器(20)连接到链路(40), 并且设置成在队列(21)中对流的数据单元(30)进行排队, 包括

确定(S1)与所述队列的长度相关的长度参数的值(QL;QL<sub>av</sub>),

把所述值(QL;QL<sub>av</sub>)与长度门限值(L<sub>th</sub>)进行比较(S2), 并且如果所述值(QL;QL<sub>av</sub>)大于所述长度门限值(L<sub>th</sub>), 则执行(S3)拥塞通知过程, 以及

自动门限适配过程(S6),

其特征在于

所述自动门限适配过程(S6)包括用于根据一个或多个流量控制参数(LIM1,LIM2;rwnd;rlfr)来调节所述长度门限值(L<sub>th</sub>)的过程。

2. 如权利要求 1 所述的方法, 其特征在于, 所述一个或多个流量控制参数(LIM1,LIM2)为预定值。

3. 如权利要求 2 所述的方法, 其特征在于, 所述预定值与数据单元发送方和数据单元接收方中的一个或两者的已知流量控制过程关联。

4. 如权利要求 1 至 3 其中之一所述的方法, 其特征在于, 还包括用于从所述队列(21)中排队的所述流的发送方和接收方其中之一所引入的流量控制参数(rwnd;rlfr)中确定所述一个或多个流量控制参数(LIM1,LIM2;rwnd;rlfr)中的一个或多个的过程(S5)。

5. 如权利要求 4 所述的方法, 其特征在于, 所述流量控制参数由所述接收方引入, 并被插入从所述接收方发送给所述发送方以用于确认数据单元的正确接收的确认数据单元中。

6. 如权利要求 5 所述的方法, 其特征在于, 所述缓冲器设置在连接所述发送方和所述接收方的通信网络的网络节点中, 其中, 用于确定所述流量控制参数的所述过程包括在所述网络节点上从所述

确认数据单元中提取所述流量控制参数。

7. 如权利要求 5 或 6 所述的方法, 其特征在于, 所述缓冲器设置在连接所述发送方和所述接收方的通信网络的第一网络节点中, 其中, 用于确定所述流量控制参数的所述过程包括在与所述第一网络节点不同的第二网络节点上从所述确认数据单元中提取所述流量控制参数, 并把所述流量控制参数从所述第二网络节点发送给所述第一网络节点。

8. 如权利要求 1 至 7 其中之一所述的方法, 其特征在于, 对于所述队列(21)中的所述流执行的流量控制是基于窗口的, 其中一个或多个流量控制参数之一是控制窗口。

9. 如权利要求 8 所述的方法, 其特征在于, 所述控制窗口由所述接收方引入, 并且表示接收方可处理的数据单元数量的限制。

10. 如权利要求 8 所述的方法, 其特征在于, 所述控制窗口由所述发送方引入, 并且表示发送方可发送的数据单元数量的限制。

11. 如权利要求 1 至 7 其中之一所述的方法, 其特征在于, 对于所述队列(21)中的所述流执行的流量控制是基于速率的, 其中一个或多个流量控制参数之一是控制速率。

12. 如权利要求 10 所述的方法, 其特征在于, 所述控制速率由所述接收方引入, 并且表示接收方可处理的到达数据单元的数据速率限制。

13. 如权利要求 10 所述的方法, 其特征在于, 所述控制速率由所述发送方引入, 并且表示对于发送方可发送的数据单元的速率的数据速率限制、当前发送速率和目标发送速率其中之一。

14. 如权利要求 1 至 13 其中之一所述的方法, 其特征在于, 所述自动门限适配过程包括

估算链路容量值(LC),

在所述长度门限值( $L_{th}$ )设置成等于所述估算链路容量值(LC)的条件下, 分析拥塞通知过程的执行是否会因所述发送方对拥塞通知的

反应而导致所述链路的利用不足, 以及

通过如果所述分析步骤表明没有利用不足, 则把所述长度门限值( $L_{th}$ )设置成等于所述估算链路容量值(LC), 否则把所述长度门限值( $L_{th}$ )设置成大于所述估算链路容量值(LC), 根据所述分析步骤的结果来适配所述长度门限值( $L_{th}$ )。

15. 如权利要求 14 所述的方法, 其特征在于, 如果所述分析步骤表明利用不足, 则所述长度门限值( $L_{th}$ )设置成根据所述流量控制参数(LIM1,LIM2;rwnd;rlfr)其中之一导出的值。

16. 如权利要求 14 或 15 所述的方法, 其特征在于, 所述队列(21)中的所述流的发送方以预定序列来发送所述数据单元, 所述队列(21)中的所述流的接收方向所述发送方发送用于确认所述数据单元的正确接收的确认消息, 其中, 每个确认消息标识在所述序列中正确接收的最后一个数据单元, 以及所述接收方向所述发送方发送表示所述接收方可处理的数据单元数量的限制的第一窗口值, 基于窗口的由所述发送方执行的流量控制使用发送窗口, 所述发送窗口被选作所述第一窗口值与第二窗口值中的最小值, 使得所述发送方不得发送具有比最高确认序列号与发送窗口之和更高的序列号的数据单元, 以及所述发送方将所述第二窗口值除以二来作为对拥塞通知的反应, 然后对于它所接收的每个复制确认消息把第二窗口增加预定增量,

其中

所述一个或多个流量控制参数(LIM1,LIM2;rwnd;rlfr)其中之一是所述第一窗口值, 以及所述长度门限值( $L_{th}$ )最初设置成等于所述估算链路容量值(LC), 以及

所述自动门限适配过程包括在所述第一窗口值大于所述估算链路容量值(LC)与所述长度门限值( $L_{th}$ )的瞬时值之和的 1.5 倍时, 把所述长度门限值( $L_{th}$ )设置成等于所述估算链路容量值(LC)。

17. 如权利要求 16 所述的方法, 其特征在于, 所述自动门限适

配过程包括在所述第一窗口值大于或等于所述估算链路容量值(LC)与所述长度门限值( $L_{th}$ )的瞬时值之和的 1.5 倍时,把所述长度门限值( $L_{th}$ )设置成等于所述估算链路容量值(LC)。

18. 如权利要求 16 或 17 所述的方法,其特征在于,如果所述第一窗口值不满足把所述长度门限值( $L_{th}$ )设置成等于所述估算链路容量值(LC)的条件,则把所述长度门限值( $L_{th}$ )设置成等于所述第一窗口值的函数。

19. 如权利要求 18 所述的方法,其特征在于,所述函数是所述第一窗口值与预定减少值之间的差值。

20. 一种计算机程序产品,包括设置成在连接到包含所述链路的通信网络的可编程数据处理装置上运行时执行权利要求 1 至 19 其中之一的方法的计算机程序。

21. 一种队列缓冲控制器(10),用于控制连接到链路(40)并且设置成在队列(21)中对流的数据单元(30)进行排队的队列缓冲器(20),包括

队列长度确定器(101),用于确定与所述队列的长度相关的长度参数的值( $QL;QL_{av}$ ),

比较器(102),用于把所述值与长度门限值进行比较,

拥塞通知器(103),用于在所述值大于所述长度门限值时执行拥塞通知过程,以及

门限适配器(104),用于自动适配所述长度门限值,

其特征在于

所述门限适配器(104)设置成用于根据一个或多个流量控制参数(LIM1,LIM2;rwnd;rlfr)来调节所述长度门限值。

22. 如权利要求 21 所述的队列缓冲控制器(10),其特征在于,所述一个或多个流量控制参数(LIM1,LIM2)为预定值。

23. 如权利要求 22 所述的队列缓冲控制器(10),其特征在于,所述预定值存储在所述队列缓冲控制器(10)中,并且与数据单元接收方

的已知流量控制过程关联。

24. 如权利要求 19 至 21 其中之一所述的队列缓冲控制器(10), 其特征在于, 还包括流量控制参数确定器(105), 用于从所述队列(21)中排队的所述流的发送方和接收方其中之一所引入的流量控制参数(rwnd;rlfr)中确定所述一个或多个流量控制参数(LIM1,LIM2;rwnd;rlfr)中的一个或多个。

25. 如权利要求 24 所述的队列缓冲控制器(10), 其特征在于, 所述流量控制参数由所述接收方引入, 并且被插入从所述接收方发送给所述发送方以用于确认数据单元的正确接收的确认数据单元中, 以及所述队列缓冲器(20)设置在连接所述发送方和所述接收方的通信网络(3)的网络节点中, 其中所述流量控制参数确定器(105)设置成用于在所述网络节点上从所述确认数据单元中提取所述流量控制参数。

26. 如权利要求 21 所述的队列缓冲控制器(10), 其特征在于, 所述流量控制参数由所述接收方引入, 并且被插入从所述接收方发送给所述发送方以用于确认数据单元的正确接收的确认数据单元中, 以及其中所述队列缓冲器(20)设置在连接所述发送方和所述接收方的通信网络(3)的第一网络节点中, 其中所述流量控制参数确定器(105)设置成用于接收来自第二网络节点的所述流量控制参数, 所述流量控制参数在所述第二网络节点被提取。

27. 如权利要求 21 至 26 其中之一所述的队列缓冲控制器(10), 其特征在于, 所述门限适配器(104)包括

估算器, 用于估算链路容量值(LC),

分析器, 用于在所述长度门限值设置成等于所述估算链路容量值(LC)的条件下, 分析拥塞通知过程的执行是否会因所述发送方对拥塞通知的反应而导致所述链路的利用不足, 以及

适配器, 用于通过在所述分析步骤表明没有利用不足时把所述长度门限值设置成等于所述估算链路容量值(LC), 否则把所述长度门

限值设置成大于所述估算链路容量值(LC)，根据所述分析步骤的结果来适配所述长度门限值。

28. 如权利要求 27 所述的队列缓冲控制器(10)，其特征在于，所述门限适配器(104)设置成在所述分析器表明利用不足时把所述长度门限值设置成根据所述流量控制参数(LIM1,LIM2;rwnd;rlfr)其中之一导出的值。

29. 如权利要求 27 或 28 所述的队列缓冲控制器(10)，其特征在于，所述队列中的所述流的发送方以预定序列来发送所述数据单元，所述队列中的所述流的接收方向所述发送方发送用于确认所述数据单元的正确接收的确认消息，其中，每个确认消息标识在所述序列中正确接收的最后一个数据单元，以及所述接收方向所述发送方发送表示所述接收方可处理的数据单元数量的限制的第一窗口值，基于窗口的由所述发送方执行的流量控制使用发送窗口，所述发送窗口被选作所述第一窗口值与第二窗口值中的最小值，使得所述发送方不得发送具有比最高确认序列号与发送窗口之和更高的序列号的数据单元，以及所述发送方将所述第二窗口值除以二来作为对拥塞通知的反应，然后对于它所接收的每个复制确认消息把第二窗口增加预定增量，

其中

所述一个或多个流量控制参数(LIM1,LIM2;rwnd;rlfr)其中之一是所述第一窗口值，以及所述门限适配器(104)设置成最初把所述长度门限值设置为等于所述估算链路容量值，以及如果所述第一窗口值大于所述估算链路容量值(LC)与所述长度门限值( $L_{th}$ )的瞬时值之和的 1.5 倍，则把所述长度门限值设置成等于所述估算链路容量值。

30. 如权利要求 29 所述的队列缓冲控制器(10)，其特征在于，所述门限适配器(104)设置成用于在所述第一窗口值大于或等于所述估算链路容量值(LC)与所述长度门限值( $L_{th}$ )的瞬时值之和的 1.5 倍时，把所述长度门限值设置成等于所述估算链路容量值。

31. 如权利要求 29 或 30 所述的队列缓冲控制器(10), 其特征在于, 所述门限适配器(104)在所述第一窗口值不满足把所述长度门限值设置成等于所述估算链路容量值的条件时, 把所述长度门限值设置成等于所述第一窗口值的函数。

32. 如权利要求 31 所述的队列缓冲控制器(10), 其特征在于, 所述函数是所述第一窗口值与预定减少值之间的差值。



## 用于控制队列缓冲器的方法及装置

### 技术领域

本发明涉及用于控制队列缓冲器的方法及装置，其中，所述队列缓冲器连接到链路，并在队列中对流的数据单元排队。

### 背景技术

在基于数据单元的通信、即要发送的信息被分为多个单元并且各个单元通过通信网络发送的通信中，已知的是沿网络在链路上提供队列缓冲器，使得可缓冲在这种链路上传送的数据单元。缓冲器可能是发送或输出缓冲器(即用于要通过出局链路发送的数据单元的缓冲器)或者接收或输入缓冲器(即用于已经通过入局链路发送的数据单元的缓冲器)。

用于传送数据的这类单元可带有各种名称，诸如协议数据单元、帧、分组、段、信元等，取决于具体上下文、所使用的具体协议和其它某些惯例。在本文献的上下文中，数据的所有这类单元一般将称作数据单元。

把数据单元放入队列中、使它们在队列中前移以及从队列中删除数据单元的过程被称作队列管理或队列控制。

队列管理的许多概念是已知的，包括监测队列长度参数(例如队列的绝对长度或平均长度)并与长度门限值比较、以便在超过门限值时执行拥塞通知过程的这类概念。这种拥塞通知过程可包括丢弃数据单元和/或标记数据单元(例如在数据单元中设置拥塞通知标志)。在EP-A-1249972的介绍中描述了例如全满时丢弃或随机早期检测(RED)等的这类管理概念的各种实例。因此，在此不需要重复。

EP-A-1249972 提出一种方案，其中，要与队列长度参数进行比

较的长度门限值根据排队数据单元要通过其发送的链路的一个或多个特性自动更新。通过使自动门限适应一个或多个链路特性，获得一种非常灵活的主动队列管理形式，它可提供改进的吞吐量和减小的延迟，特别是在具有时变特性的链路、比如无线链路上。

## 发明目的

本发明的一个目的是提供队列缓冲器管理的改进方法及装置。

## 发明内容

这个目的通过本申请的独立权利要求中所述的方法及装置来实现。在从属权利要求中描述有利的实施例。

根据本发明，在系统中，队列缓冲器控制或管理包括确定与队列长度相关的长度参数的值，把这个值与门限进行比较，并在超过门限时执行拥塞通知过程，以及提供自动门限适配过程，它包括用于根据一个或多个流量控制参数来调节长度门限值的过程。

通过考虑与排队数据单元流量关联的一个或多个流量控制参数，队列缓冲器管理变得更为灵活，并且更好地能够以使得对流量控制的缓冲器管理动作或判定的影响不产生不希望的效果的方式进行，和/或更好地能够以比先有技术系统更直接和智能的方式对流量控制作出反应。

一个或多个流量控制参数可通过任何适当或希望的方式来确定。作为一个实例，它们可能是缓冲控制器中存储的预定值。优选地，它们在这个情况中与数据单元接收方的已知流量控制过程相关联。换言之，如果已知某些流量控制过程(例如由包括流量控制例程的广泛使用的计算机程序提供)采用预定流量控制参数，则这些已知流量控制参数可用作调节长度门限值的基础。

作为对于根据预定值来导出一个或多个流量控制参数中的至少一部分的替代或补充方式，可提供用于从所述队列中排队的所述流的

发送方和接收方之一引入的流量控制参数中确定所述一个或多个流量控制参数中的至少一部分的过程。

缓冲器管理可通过关于给定网络上下文、流量控制原则以及预期效应(或者希望避免的负效应)所希望的任何方式来调谐。根据一个优选实施例,队列缓冲器管理对已确定流量控制参数的适配用来避免链路利用不足。链路利用不足意味着存在链路空闲(即没有正传送的数据单元)的周期。更具体来说,在这个优选实施例中,检查预期长度门限值是否可能导致又可能因发送数据单元的发送方的反应而引起链路利用不足的拥塞通知操作,以及长度门限值以不同方式适配,以便避免这种链路利用不足。因而更有效地利用网络资源。

## 附图说明

通过以下参照附图对优选实施例的详细说明,本发明的其它方面和细节将变得明显,附图包括:

图 1 说明根据本发明的一个实施例的缓冲器和缓冲控制器的示意框图表示,

图 2 说明本发明的一个方法实施例的流程图,

图 3 说明根据本发明的另一个实施例的缓冲器和缓冲控制器的示意框图表示,以及

图 4 说明本发明的另一个方法实施例的流程图。

## 具体实施方式

虽然本发明的以下优选实施例通常参照采用基于窗口的流量控制的系统,更具体来说是参照传送 IP(因特网协议)数据单元的系统,但是本发明既不是限制到基于窗口的流量控制也不是限制到基于 IP 的系统。本发明可适用于任何类型的数据单元的排队而与具体协议无关,并且可适用于任何类型的流量控制,例如还适用于基于速率的流量控制,只要能够让缓冲器管理过程确定流量控制参数。

此外，本发明的方法可适用于任何已知的队列管理方案，在其中，队列长度相关参数与至少一个长度门限值进行比较，并且在超过门限时执行拥塞通知过程。也就是说，该方法例如可适用于任何已知的 RED 方案，可适用于在队列全满时丢弃数据单元、例如尾端丢弃、随机丢弃或前端丢弃的方案，以及可适用于执行作为对丢弃的替代或补充的显式拥塞通知的任何已知的方案。

优选的是结合 EP-A-1249972 中公开的主动队列管理来应用本发明的方法及装置。本文献的完整公开及其美国相应申请通过引用结合于本文中。

图 1 说明队列缓冲控制器 10 的示意表示，它能够实现用于控制队列缓冲器 20 中的数据单元的管理的本发明。参考标号 3 表示通信网络，数据单元 30 通过它到达队列缓冲器 20，以便在通过链路 40 发送之前被放入队列 21。

在队列 21 中排队的数据单元 30 属于一个或多个流。流一般由源和目标主机地址、源和目标服务接入点(SAP)标识符和协议标识符来标识。流的定义和概念是本领域例如从 TCP/IP 众所周知的，在该情况中，SAP 标识符称作端口号，以及源和目标主机地址称作 IP 地址，使得在此不需要进一步说明。

队列 21 可按照下列方式排列：所述队列 21 中的所有数据单元来自同一个流，或者它们可能来自多个流。可能的是，缓冲器 20 通过下列方式来管理：存在多个队列，每个队列与相应的流关联。

现在描述用于在队列缓冲控制器 10 中体现本发明的概念的具体元件。要注意，队列缓冲控制器一般包括不止这些元件、即用于处理所接收数据单元和管理缓冲器的已知元件，为了简洁起见，没有对它们进行明确说明。控制器 10 还可能具有具体实现如 EP-A-1249972 中所述的系统的元件。

参考标号 101 描述用于确定与队列 21 的长度相关的长度参数的值的队列长度确定器。此外，提供比较器 102，用于把已确定长度值

与设置成自动适配长度门限值的门限适配器 104 所提供的长度门限值进行比较。比较器 102 连接到在已确定长度值大于长度门限值时执行拥塞通知过程的拥塞通知器 103。下面将会更详细地继续说明，要确定的长度参数可通过任何适当或所希望的方式来选择，例如作为绝对长度队列长度或者平均队列长度，以及拥塞通知过程可同样以适当或所希望的方式来选择，例如作为数据单元丢弃过程和/或显式数据单元标记过程。

根据本发明的这个实施例，提供流量控制参数确定器 105 以用于确定在队列 21 中排队的流的发送方和接收方其中之一所引入的流量控制参数。此外，门限适配器 104 以下列方式设置：长度门限值根据已确定流量控制参数来调节。

要注意，上述元件 101-105 可作为硬件、软件或者硬件和软件的任何适当组合来提供。控制器 10 优选地是可编程数据处理器，以及元件 101 至 105 是软件元件、例如程序代码部分。

图 2 说明本发明的方法的一个方法实施例的流程图，该方法可采用图 1 所示的控制器 10 来执行。在第一步骤 S1，确定与队列 21 的长度相关的长度参数的值。这个队列长度相关参数可通过任何所希望或适当的方式与队列长度相关，例如可能是实际或瞬时队列长度 QL，或者从实际或瞬时队列长度导出的参数、比如平均值 QL<sub>av</sub>。

在图 2 的实例中，队列长度相关参数是实际队列长度 QL。如果希望采用平均队列长度 QL<sub>ac</sub>，则这个平均值可按照任何已知的适当求平均算法来确定，以及这样一种算法通常可包括通过计算旧平均值与第一加权因子相乘再加上瞬时队列长度与第二加权因子相乘之和来更新旧平均值。例如，QL<sub>av</sub> 可计算为

$$QL_{av} \text{ (new)} = QL_{av} \text{ (old)} \cdot (1 - 1/2^n) + (QL \cdot 1/2^n)$$

其中，QL 表示瞬时队列长度值，以及 n 是在 0 与 1 之间可调整的指数加权因子。

回到图 2，队列长度参数 QL 与长度门限值 L<sub>th</sub> 进行比较。如果超

过长度门限值  $L_{th}$ ，则执行拥塞通知过程 S3，否则省略拥塞通知过程 S3。

如以上所述，拥塞通知过程可通过任何适当或所希望的方式来选择。例如，它可包括从队列 21 中丢弃/标记一个或多个预定或者一个或多个随机选择的数据单元，或者包括在把一个或多个新到达的数据单元放入队列 21 之前将其丢弃/标记，或者与预定概率函数相关地执行这种丢弃/标记操作。标记可例如通过在排队数据单元的一个或多个中设置显式拥塞通知标志来实现。

在步骤 S2 或 S3 之后，过程转到步骤 S4，在其中确定是否已经发生适配长度门限参数  $L_{th}$  的适配触发事件。要注意，步骤 S4 是可选的。如果步骤 S4 判定已经发生适配触发事件，则过程转到步骤 S5，在其中确定队列 21 中排队的流的发送方和接收方其中之一所引入的流量控制参数。然后，过程转到步骤 S6，在其中进行  $L_{th}$  的适配过程，它根据步骤 S5 所确定的流量控制参数来调节长度门限值  $L_{th}$ 。

可能要注意，图 2 所示的步骤的具体排列只是一个实例。特别地，组成用于判定拥塞通知过程的执行的过程的步骤 S1、S2 和 S3 与步骤 S4、S4 和 S6 构成的适配过程无关。因此，步骤 S1、S2 和 S3 可与 S4、S5 和 S6 无关地排列，使得步骤 S4-S6 可在步骤 S1-S3 之前执行或者与其并行执行。可能特别要注意，一方面的步骤 S1-S3 以及另一方面的 S4-S6 一般包含在控制或管理队列缓冲器的更大的方法中，更大的方法具有更多步骤和过程，但是其中这些附加步骤和过程没有示出，因为它们不属于本发明。图 2 的方法可实现为软件，以及步骤 S1-S3 可例如在一个线程中实现，而 S4-S6 可在另一个独立线程中实现。但是，该方法也可直接以硬件的形式来实现。

要注意，步骤 S6 的  $L_{th}$  的适配过程一般不仅考虑已确定流量控制参数，而且还考虑与队列管理、例如 EP-A-1249972 提出的链路相关特性相干的其它参数。

此外还要注意，引起自动门限适配过程 S6 的发起的步骤 S4 中的

适配触发事件可按照适当或所希望的方式来选择。例如，可能定期发起自动门限适配过程，使得步骤 S4 中的适配触发事件可能是某个时间条件变为真或者是计数器达到预定值的事件。适配触发事件还可包括在适配过程 S6 中起一定作用的一个或多个参数的变化，例如测量的流量控制参数或者一个或多个链路特性参数。可能要注意，实际上，变化仅采用某个粒度来确定。换言之，不是参数中的任何非常小的变化都被认为是触发事件，而是在出现预定幅度的变化时才认为是触发事件。

确定流量控制参数的过程可确定所关注流的流量控制所涉及的一个或多个流量控制参数。优选地，一个流量控制参数被监测，以及该流量控制参数优选地由流的接收方引入。这例如可通过把流量控制参数插入从接收方发送给发送方用于确认数据单元的正确接收的确认数据单元中来进行。这个种类的方案是本领域众所周知的，例如 ARQ(自动重传请求)，使得在此不需要作进一步说明。

例如，如果流量控制基于窗口，则流量控制参数可能是接收方所引入的控制窗口，所述控制窗口表示接收方可处理的数据单元数量的限制。例如，所测量流量控制参数可能是从 TCP 已知的接收窗口或者通知窗口 *rwnd*。但是，所测量流量控制参数也可由发送方引入，并且表示发送方可发送的数据单元数量的限制。

另一方面，在基于速率的流量控制系统中，所测量流量控制参数可能是接收方所引入的控制速率，例如它表示接收方可处理的到达数据单元的数据速率限制，它也可称作 *rlfr*(接收方限制流速率)。另一方面，流量控制参数之一还可能是流 13 的发送方所公告的数据速率，例如表示发送方可发送的数据单元的速率的数据速率限制、所述发送方的当前发送速率或者所述发送方的目标发送速率。目标发送速率例如可由发送视频信息并打算以某个质量发送的应用来通告。

图 1 所示的控制器 10 和缓冲器 20 一般在通信网络 3 的网络节点中提供。这样一种通信节点例如可能是网络 3 中的服务器和/或路由

器。

如果流量控制参数从上述确认数据单元中提取,则这可在其中配备了控制器 10 和缓冲器 20 的网络节点上进行,或者可在不同的网络节点上进行,在后一种情况中,所提取流量控制参数从所述另一个网络节点发送给其中配备了控制器 10 和缓冲器 20 的网络节点。

图 1 说明前一种情况的一个实例,在其中,另外假定缓冲器 20 还保持确认数据单元,使得流量控制参数确定器 105 可从缓冲的确认数据单元中提取预期流量控制参数。

流量控制参数的提取一般优选地在其中配备了缓冲器和缓冲控制器的节点中进行。但是,网络以下列方式设置是可行的:这例如因数据单元由具有嵌入其中的高层数据单元的缓冲器来处理而极为困难或者不可行,以及附加加密已经在嵌入的高层数据单元中执行,使得所述高层数据单元中包含的任何控制参数无法被提取。要注意,术语“嵌入”表示封装和/或分段。

在这种情况下,其中提取在配备了控制器和缓冲器的网络节点上不可行,应当实现一个过程,使得其中嵌入的高层数据单元可解密的不同网络节点执行预期流量控制参数的提取,并把所提取流量控制参数发送给所关注的网络节点。发送可通过专用连接来执行,或者通过提供从另一个网络节点发送给所关注网络节点的数据单元中的专用位置来执行。

在后一种情况中,流量控制参数确定器 105 本身连接到网络 3,以便接收所提取流量控制参数(图 1 中未示出)。

现在参照图 3 和图 4 来描述其它实施例。要注意,与图 1 和图 2 中包含的相同的参考标号在图 3 和图 4 中表示相应的元件,使得不需要重复说明。

图 3 所示的配置几乎与图 1 所示相同,但流量控制参数确定器 105 已经由用于存储基于已知流量控制过程所使用的流量控制参数的一个或多个预定值的存储装置 106 代替。图 4 说明一个流程图,它与



图 2 相同, 但步骤 S5 被省略, 以及 S6 中的  $L_{th}$  的适配过程设置成在适配  $L_{th}$  时考虑存储装置 106 中存储的一个或多个值。

流量控制过程中使用的参数例如可从以下事实来了解: 包含流量控制例程的某些计算机程序始终采用一个或多个具体的流量控制参数。那么, 不需要从数据单元或者确认消息中获取这类参数。作为一个实例, Microsoft Corporation 的众所周知的操作系统 Windows 2000 为 TCP 接收窗口、即 TCP 接收方向 TCP 发送方通告的窗口的计算指定特定值。这些值基于 16 KB 或 64 KB。初始条件中使用的精确值为 16384 字节, 以及对于以太网, 窗口通常设置为 17520 字节(16 KB 上舍入为十二个 1460-字节段, 参见 Dave MacDonald 和 Warren Barkley 的 Microsoft Windows 2000 TCP/IP 实现细节, 白皮书, 2000 年, 第 31 和 32 页)。因此, 值 16 KB 和 64 KB(以及可能 16384 字节和/或 17520 字节)可被存储, 作为预定流量控制值。

要注意, 图 1 和图 3 所示的实施例可结合, 即以利用预定存储流量控制参数以及由发送方和/或接收方所引入的并从数据单元或确认消息中得到的流量控制参数的一个实施例的形式。

如以上所述, 自动门限适配过程可通过任何适当或所希望的方式来设置。它优选地以下列方式来设置: 它包括估算链路容量值 LC, 在长度门限值设置成等于估算链路容量值的条件下, 分析拥塞通知过程的执行是否因发送方对拥塞通知的反应而导致这个估算链路容量的利用不足, 以及根据分析步骤的结果, 通过在分析步骤表明没有利用不足时把链路门限值设置成等于估算链路容量值, 否则把长度门限值设置成大于估算链路容量值, 来适配长度门限值。

链路容量值可被理解为在考虑中的流的发送方必需发出的数据的最小量, 使得链路 40 分配给所述流的带宽被完全利用。完全利用表示分配到给定流的链路带宽的份额始终被完全利用。以稍微不同的方式来说明, 如果考虑仅服务于单个流的链路 40 的简化实例, 则这意味着那个流的发送方发送许多数据, 使得链路 40 一直忙, 即不断

发送数据单元而在中间没有任何空闲时间。再换句话说，链路容量是发送方发送的数据量，使得被发送的任何附加数据不会增加吞吐量，因为附加数据被排队。

链路容量还可被理解为链路 40 提供给所关注流的数据速率与无负荷网络的情况的所述流关联的往返时间(RTT)相乘之积。无负荷网络表示没有排队延迟。因此，无负荷 RTT 的值等于流的实际 RTT 与所述流的所有排队延迟之间的差值。

因此，链路容量值的估算一般包括确定表示无负荷 RTT 的时间值，并把这个值与链路 40 为所关注流提供的数据速率 DR 相乘。这个无负荷 RTT 的确定可通过任何适当或所希望的方式进行。一个实例是计算常数  $RTT_{wc}$  与链路 40 所提供的 RTT 之和。 $RTT_{wc}$  是除了所关注链路 40 的 RTT 之外的整个无负荷 RTT 的最坏情况估算，并且可具有 150 到 300 ms、更优选为 150 到 250 ms 的值。估算给定流的无负荷 RTT 的一个备选可能性包括例如通过保持所缓冲数据单元 30 在队列 21 中花费的时间量的平均值，并计算流的实际 RTT 与这个排队延迟之间的差值，来计算缓冲器 20 上的排队延迟。流的实际 RTT 的值例如可由发送方插入所述流的数据单元中，并由控制器 10 读取。

要注意，链路容量的估算值 LC 一般与实际瞬时链路容量不相同。估算链路容量的过程优选地使估算值超过实际链路容量，即估算值是保守的。这可通过任何适当或所希望的方式来实现，例如通过采用以上所述的无负荷 RTT 的最坏情况估算值和/或通过把预定正因子加至估算链路容量时所使用的参数的一个或多个。换言之，可把预定因子加至无负荷 RTT 和/或链路的 RTT 和/或 DR，以及可把预定正因子加至计算结果，即，用  $LC+\epsilon$  来取代 LC 的计算值： $LC \leftarrow LC+\epsilon$ 。

分析拥塞通知过程的执行是否将导致估算链路容量的利用不足的步骤取决于用于所考虑流的流量控制的具体类型。下面继续给出具体实例。在任何情况中，如果在长度门限值被考虑设置成等于估算链

路容量值的条件下,由于发送方对拥塞通知的流量控制反应而不会出现利用不足,则实际长度门限值可设置为这个估算链路容量。另一方面,如果分析表明利用不足、即其中可能出现分配给特定流的带宽的一个或多个空闲周期的情况,则长度门限值  $L_{th}$  设置成大于估算链路容量值。进行这种操作时,执行拥塞通知过程的趋势被降低,因为减小了超过长度门限值  $L_{th}$  的概率。因此,也降低了链路利用不足的可能性。因此,链路资源不太可能因利用不足而浪费。

如果分析步骤表明利用不足,则长度门限值  $L_{th}$  优选地设置成根据流量控制参数导出的值。

这将在一个实例的上下文中说明,在其中,发送方与接收方之间的流量控制是基于窗口的。更具体来说,发送方被设置成以预定序列发送数据单元,以及接收方被设置成发送确认数据单元的正确接收的确认消息。每个确认消息标识在序列中正确接收到的最后一个数据单元,以及接收方向发送方发送表示接收方可处理的数据单元数量的限制的窗口值。流的发送方采用被选择为接收方所通告的窗口与第二窗口值中的最小值的发送窗口来执行基于窗口的流量控制,使得发送方不得发送具有比最高确认序列号和发送窗口之和更高的序列号的数据单元。作为对拥塞通知的反应,发送方将第二窗口值除以  $k(k>1)$ ,此后再对于它所接收的每个复制确认消息将第二窗口增加一个预定增量。复制确认消息是标识已经在前一个确认数据单元中确认的序列号的正确接收的确认数据单元。

以这种方式工作的系统的一个实例是 TCP,在其中,通知窗口又称作  $rwnd$ ,上述第二窗口是拥塞窗口,以及对拥塞通知作出反应的过程称作快速恢复。对于 TCP,  $k=2$ 。但是要注意,当前实施例不是限制为 TCP,而是可适用于按照以上一般描述的方式工作的任何系统。

根据这个实施例,用于缓冲器管理的流量控制参数是确认数据单元中包含的第一窗口值,即表示接收方可处理的数据单元数量的限制

的窗口值。此外，长度门限值  $L_{th}$  最初(即在开始管理所关注流的数据单元时)设置为上述估算链路容量值  $LC$ 。然后，自动门限适配过程  $S6$  以下列方式来执行：如果第一窗口值(例如  $rwnd$ )大于最初等于估算链路容量值  $LC$  的三倍的( $L_{th}+2 \times LC$ )，则长度门限值  $L_{th}$  设置成等于估算链路容量值  $LC$ 。

其原因说明如下。在按照上述方式工作的流量控制系统中，如果队列长度门限  $L_{th}$  设置为估算链路容量值  $LC$ ，则在拥塞通知过程时的发送窗口大小将为  $L_{th}+LC=2 \times LC$ 。这是因为以下事实：如果发送窗口大小等于链路容量，则队列长度将为零。此外，为了在拥塞通知响应阶段期间(即，当第二窗口值除以二，然后对于每个复制确认消息增加一个预定增量时)使用链路，这个第二窗口必需能够扩充到在拥塞通知过程时的发送窗口大小的 1.5 倍。因此，如果发送窗口无法扩充到超过  $2 \times 1.5 \times LC$ ，则拥塞通知将导致链路利用不足。因此，比较接收方所通告的窗口与  $3 \times LC$  的步骤是用于分析是否将响应拥塞通知而出现链路利用不足的步骤的一个实例。

在一般化以上考虑时，可能要考虑， $L_{th}$  可能改变并且不再等于  $LC$ ，使得条件是把通知窗口与  $(L_{th}+LC) \times 1.5$  进行比较。

采用 TCP 作为实例，这个实施例也可用伪代码来表达：

```

If (  $rwnd > 1.5 \times (LC + L_{th})$  ) {
   $L_{th} = LC$ ;
} else {
   $L_{th} = X$ ;
}

```

要注意，在上述实例中，如果  $rwnd$  大于  $1.5 \times (LC+L_{th})$ ，则长度门限  $L_{th}$  设置成等于估算链路容量  $LC$ 。但是，如果  $rwnd=1.5 \times (LC+L_{th})$ ，设置  $L_{th}=LC$  也是可行的。

$X$  的值可通过任何适当或所希望的方式来选择，只要它大于  $LC$ 。例如，它可选择为第一窗口值(本例中为  $rwnd$ )的函数。它优选地设置成等于第一窗口值减去预定减少值，其中，所述预定减少值可

能为零，或者其本身可能是流的数据单元的最大段大小的函数。例如，预定减少值可能是最大段大小的两倍。

与流的接收方所通告的窗口值相关地设置长度门限值的好处在于，如果流由接收方限制，则不会达到长度门限值。如果没有达到长度门限值，则不会进行拥塞通知过程，因而不会不必要地产生链路利用不足。因此，实现链路资源的更好利用，同时没有负面影响在发送方和接收方执行的流量控制。

现在描述另一个 TCP 实例，它基于结合图 3 和图 4 所述的配置及方法。假定存储了两个已知的接收方窗口值，表示为 LIM1 和 LIM2。LIM1 大于 LIM2。例如，LIM1=64KB，LIM2=16KB，参见以上结合操作系统 Windows2000 所述的实例。LIM1 无疑也可能是基于 64KB 所选的值、例如 64KB 左右给定范围内的任何值，以及 LIM2 同样也可能是基于 16KB 所选的值、例如 16KB 左右给定范围内的任何值。以伪代码表达：

```
If  $1.5 \times (LC + L_{th}) > LIM1$  {  
   $L_{th} = X1$ ;  
} else if  $1.5 \times (LC + L_{th}) > LIM2$  {  
   $L_{th} = X2$ ;  
} else {  
   $L_{th} = LC$   
}
```

X1 大于 X2，并且两者均大于 LC。优选地，X1 为 LIM1 的函数，和/或 X2 为 LIM2 的函数。例如，函数可能是 LIM1 或 LIM2 减去预定减少值的值，其中，所述预定减少值可能为零，或者其本身可能是流的数据单元的最大段大小的函数。例如，预定减少值可能是最大段大小的两倍。

虽然已经通过具体实例描述了本发明，但这些不是要进行限制，因为本发明的范围由所附权利要求来定义。权利要求中的参考标号仅用来使权利要求更易于阅读，同样不是要具有任何限制作用。

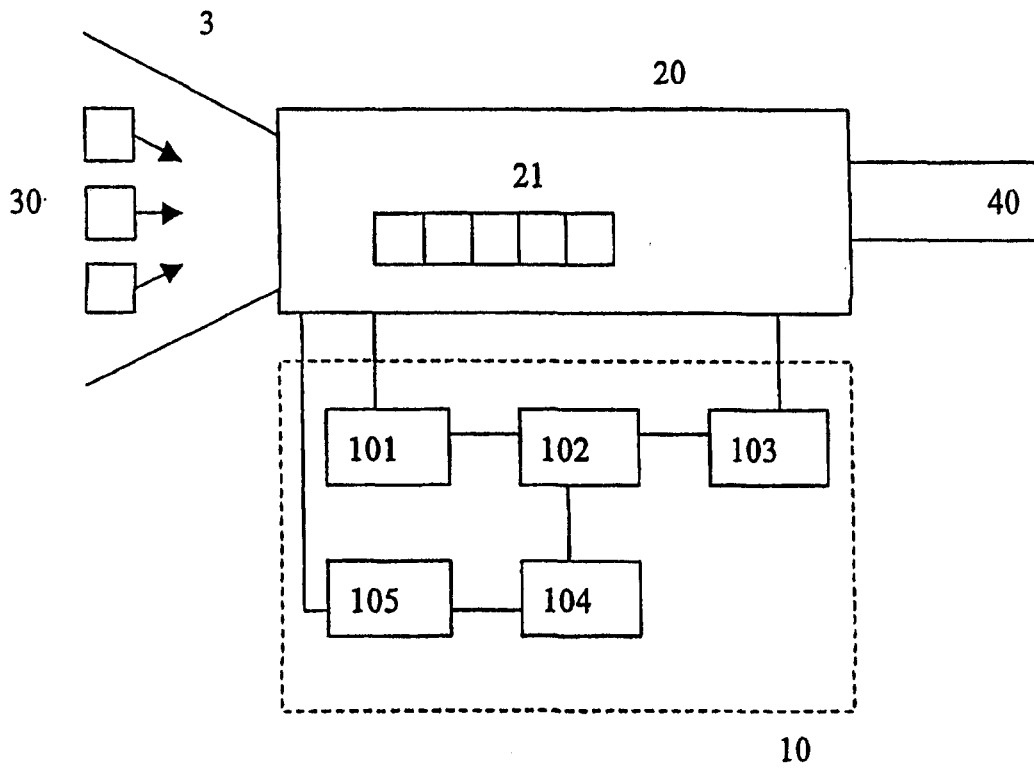


图 1

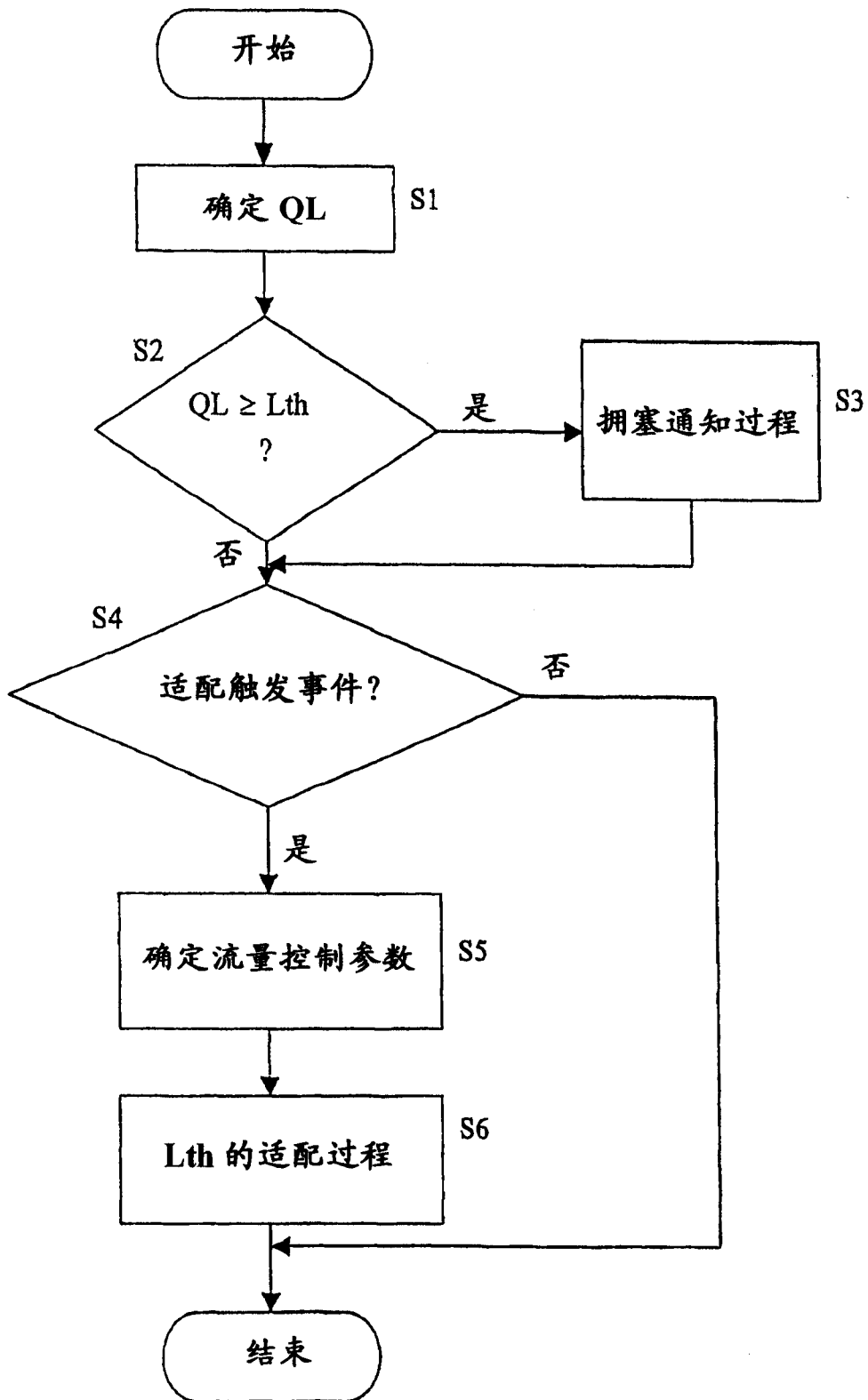


图 2

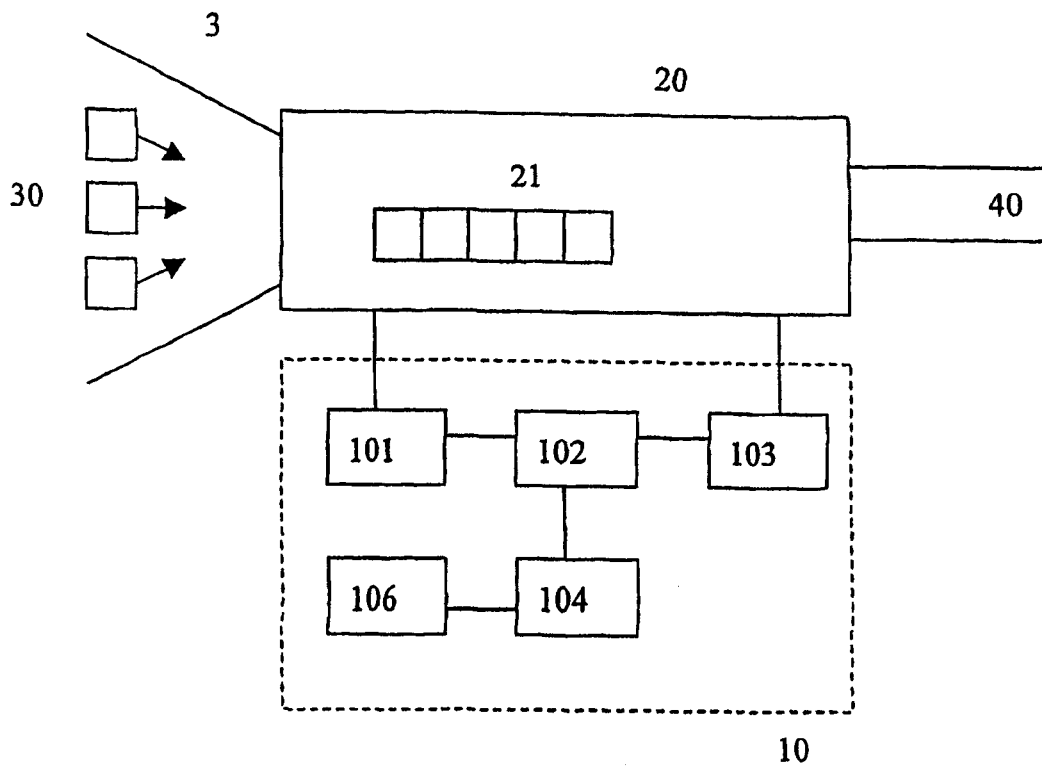


图 3



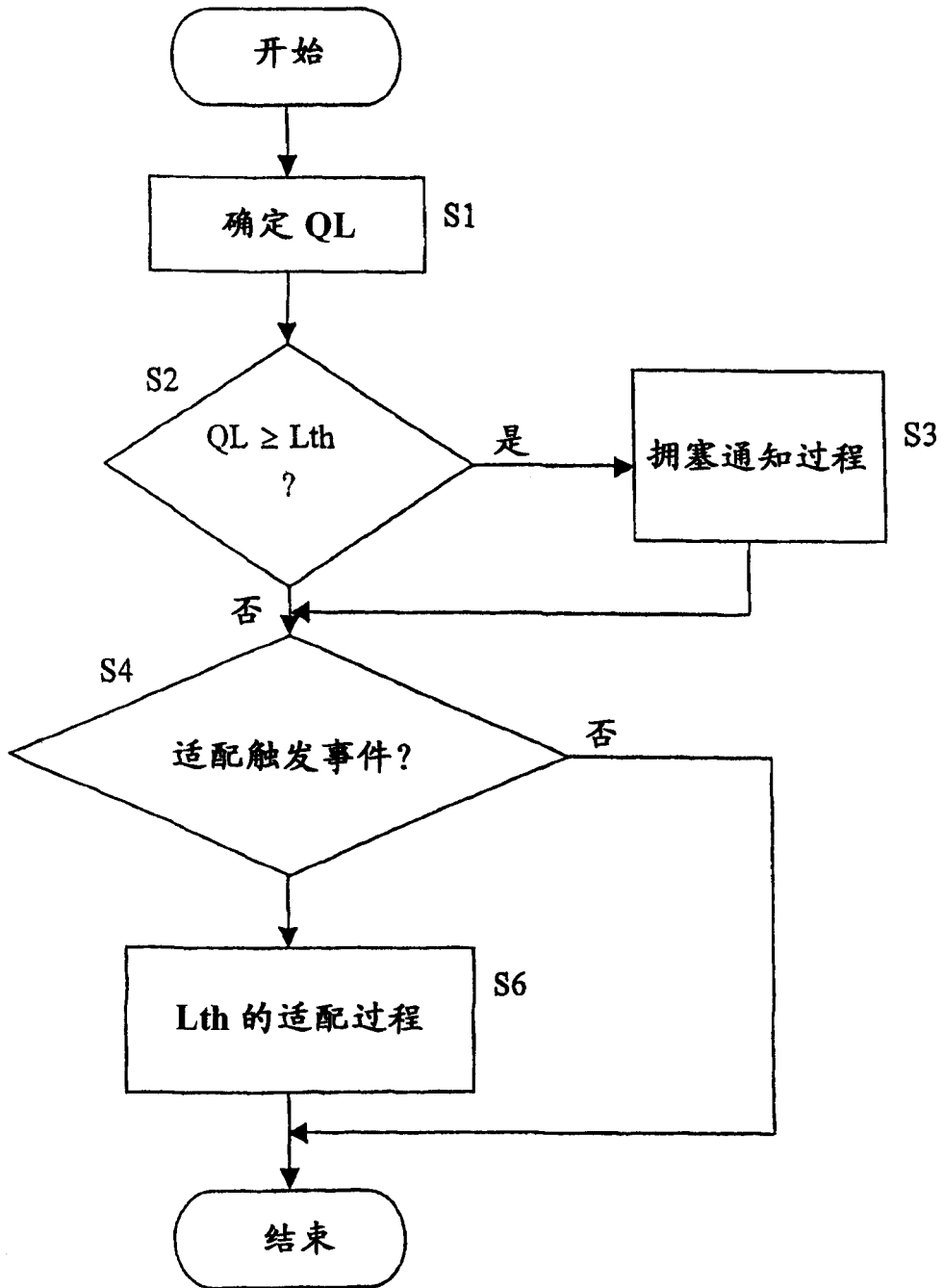


图 4