



(12)发明专利

(10)授权公告号 CN 104040932 B

(45)授权公告日 2018.10.02

(21)申请号 201380004762.2

(22)申请日 2013.01.16

(65)同一申请的已公布的文献号
申请公布号 CN 104040932 A

(43)申请公布日 2014.09.10

(30)优先权数据
1250025-2 2012.01.17 SE

(85)PCT国际申请进入国家阶段日
2014.07.03

(86)PCT国际申请的申请数据
PCT/EP2013/050735 2013.01.16

(87)PCT国际申请的公布数据
W02013/107767 EN 2013.07.25

(73)专利权人 网络洞察力知识产权公司
地址 瑞典斯德哥尔摩省

(72)发明人 G·让格 马丁·卡尔森

(74)专利代理机构 北京派特恩知识产权代理有限公司 11270
代理人 徐川 武晨燕

(51)Int.Cl.
H04L 1/22(2006.01)

(56)对比文件
CN 1501640 A,2004.06.02,
CN 101345705 A,2009.01.14,
审查员 陈秀英

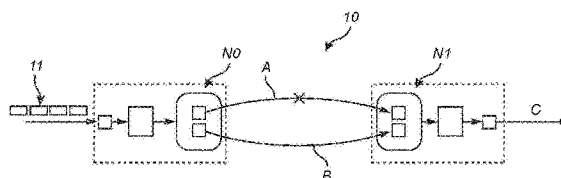
权利要求书3页 说明书9页 附图9页

(54)发明名称

一种网络通信冗余方法

(57)摘要

本申请提供了一种改进的用于通信网络的冗余方法。用于节点到节点通信的方法包括：为待传送的数据流(11)中的每个数据包提供各自的标识符；然后，至少两条链路(A、B)上将该数据流中的经复制的数据从发送节点(N0)传递到接收节点(N1)；在该接收节点处，接收并且缓冲来自至少两条链路的数据包，以便后续对数据流进行转发。标识符对应于数据包在数据流中的相对位置，并且缓冲的步骤包括：使用从至少两条链路中的至少一条链路接收的数据包，在公用索引缓冲区中重构数据流的预定部分。本发明是建立在以下的基础上：通过在接收节点处使用单个公用索引播放缓冲区来进行无中断故障切换，在索引播放缓冲区中，利用在多条链路上接收到的数据流中的单独的块(即，帧或数据包)，来选择性地重构已传递的数据流。



1. 一种用于节点到节点通信的方法,包括:
 - 为待传送的数据流(11)中的每个数据包提供各自的标识符;
 - 在至少两条链路(A、B)上将所述数据流的经复制的数据从发送节点(N0)传递到接收节点(N1);在所述接收节点处:
 - 接收并且缓冲来自所述至少两条链路的数据包,以便后续对所述数据流进行转发;
 - 其中,所述标识符对应于数据包在所述数据流中的相对位置,并且其中,所述缓冲的步骤包括:使用从所述至少两条链路中的至少一条链路接收的数据包,在单个公用索引缓冲区中重构所述数据流的预定部分;
 - 其中,所述方法进一步包括:
 - 在所述接收节点处,针对每个所接收到的数据包:
 - 记录数据包间到达时间;
 - 其中,所述数据流的转发是以一输出速率来进行的,所述输出速率基于所记录的数据包间到达时间数据来确定。
2. 根据权利要求1所述的方法,其中,所述重构的步骤包括:
 - 仅选择所接收到的具有特定标识符的一个数据包,以用于写入所述缓冲区中的每个索引位置。
3. 根据权利要求1所述的方法,进一步包括:
 - 针对每个所接收到的数据包,
 - 基于所接收到的数据包的标识符来确定相应的缓冲区索引;
 - 基于所确定的缓冲区索引来检验数据是否已经被写入所述相应的缓冲区索引的位置,并且如果数据没有被写入所述相应的缓冲区索引的位置,则将所接收到的数据包或者与所接收到的数据包相对应的指针写入所确定的缓冲区索引的位置。
4. 根据权利要求2所述的方法,进一步包括:
 - 针对每个所接收到的数据包,
 - 基于所接收到的数据包的标识符来确定相应的缓冲区索引;
 - 基于所确定的缓冲区索引来检验数据是否已经被写入所述相应的缓冲区索引的位置,并且如果数据没有被写入所述相应的缓冲区索引的位置,则将所接收到的数据包或者与所接收到的数据包相对应的指针写入所确定的缓冲区索引的位置。
5. 根据权利要求2至4中任一项所述的方法,其中,支持同时选择所接收到的数据包并写入所述缓冲区,所接收到的数据包具有不同的标识符并且在不同的链路上被接收。
6. 根据权利要求2至4中任一项所述的方法,进一步包括:针对每个所接收到的数据包,设置与经确定的缓冲区索引(b_n)和链路(A、B、 \dots 、N)相对应的有效位条目(V_A 、 V_B 、 \dots 、 V_N)。
7. 根据权利要求5所述的方法,进一步包括:针对每个所接收到的数据包,设置与经确定的缓冲区索引(b_n)和链路(A、B、 \dots 、N)相对应的有效位条目(V_A 、 V_B 、 \dots 、 V_N)。
8. 根据权利要求6所述的方法,进一步包括:基于所述有效位条目来监视所述节点到节点通信的当前冗余。
9. 根据权利要求7所述的方法,进一步包括:基于所述有效位条目来监视所述节点到节点通信的当前冗余。
10. 根据权利要求6所述的方法,进一步包括:将在一所接收到的数据包被记录之后到

达的具有相同标识符的任何所接收到的数据包丢弃。

11. 根据权利要求7所述的方法,进一步包括:将在一所接收到的数据包被记录之后到达的具有相同标识符的任何所接收到的数据包丢弃。

12. 根据权利要求8所述的方法,进一步包括:将在一所接收到的数据包被记录之后到达的具有相同标识符的任何所接收到的数据包丢弃。

13. 根据权利要求9所述的方法,进一步包括:将在一所接收到的数据包被记录之后到达的具有相同标识符的任何所接收到的数据包丢弃。

14. 根据权利要求1所述的方法,其中,所述输出速率是逐个数据包来确定的,使得对于特定缓冲区索引的数据包,所述输出速率是基于所接收到的数据包中具有与所述特定缓冲区索引相对应的标识符的那个数据包的所记录的数据包间到达时间来选择的。

15. 根据权利要求1所述的方法,其中,转发所述数据流时的所述输出速率是基于来自所述至少两条链路中的其中一条链路的所记录的数据包间到达时间来选择的。

16. 根据权利要求1至4中任一项所述的方法,其中,所述数据流的转发是从已经在至少两条链路上接收到所述数据流中的第一数据包之后的一段预定时间开始的。

17. 根据权利要求5所述的方法,其中,所述数据流的转发是从已经在至少两条链路上接收到所述数据流中的第一数据包之后的一段预定时间开始的。

18. 根据权利要求6所述的方法,其中,所述数据流的转发是从已经在至少两条链路上接收到所述数据流中的第一数据包之后的一段预定时间开始的。

19. 根据权利要求7所述的方法,其中,所述数据流的转发是从已经在至少两条链路上接收到所述数据流中的第一数据包之后的一段预定时间开始的。

20. 根据权利要求8所述的方法,其中,所述数据流的转发是从已经在至少两条链路上接收到所述数据流中的第一数据包之后的一段预定时间开始的。

21. 根据权利要求9所述的方法,其中,所述数据流的转发是从已经在至少两条链路上接收到所述数据流中的第一数据包之后的一段预定时间开始的。

22. 根据权利要求10所述的方法,其中,所述数据流的转发是从已经在至少两条链路上接收到所述数据流中的第一数据包之后的一段预定时间开始的。

23. 根据权利要求11所述的方法,其中,所述数据流的转发是从已经在至少两条链路上接收到所述数据流中的第一数据包之后的一段预定时间开始的。

24. 根据权利要求12所述的方法,其中,所述数据流的转发是从已经在至少两条链路上接收到所述数据流中的第一数据包之后的一段预定时间开始的。

25. 根据权利要求13所述的方法,其中,所述数据流的转发是从已经在至少两条链路上接收到所述数据流中的第一数据包之后的一段预定时间开始的。

26. 根据权利要求14或15所述的方法,其中,所述数据流的转发是从已经在至少两条链路上接收到所述数据流中的第一数据包之后的一段预定时间开始的。

27. 根据权利要求1至4中任一项所述的方法,其中,所述方法被用于中间节点的无中断的断开,所述中间节点布置在所述至少两条链路的其中一条链路中。

28. 根据权利要求5所述的方法,其中,所述方法被用于中间节点的无中断的断开,所述中间节点布置在所述至少两条链路的其中一条链路中。

29. 根据权利要求6所述的方法,其中,所述方法被用于中间节点的无中断的断开,所述

中间节点布置在所述至少两条链路的其中一条链路中。

30. 根据权利要求7所述的方法,其中,所述方法被用于中间节点的无中断的断开,所述中间节点布置在所述至少两条链路的其中一条链路中。

31. 根据权利要求8所述的方法,其中,所述方法被用于中间节点的无中断的断开,所述中间节点布置在所述至少两条链路的其中一条链路中。

32. 根据权利要求9所述的方法,其中,所述方法被用于中间节点的无中断的断开,所述中间节点布置在所述至少两条链路的其中一条链路中。

33. 根据权利要求10所述的方法,其中,所述方法被用于中间节点的无中断的断开,所述中间节点布置在所述至少两条链路的其中一条链路中。

34. 根据权利要求11所述的方法,其中,所述方法被用于中间节点的无中断的断开,所述中间节点布置在所述至少两条链路的其中一条链路中。

35. 根据权利要求12所述的方法,其中,所述方法被用于中间节点的无中断的断开,所述中间节点布置在所述至少两条链路的其中一条链路中。

36. 根据权利要求13所述的方法,其中,所述方法被用于中间节点的无中断的断开,所述中间节点布置在所述至少两条链路的其中一条链路中。

37. 根据权利要求14或15所述的方法,其中,所述方法被用于中间节点的无中断的断开,所述中间节点布置在所述至少两条链路的其中一条链路中。

38. 一种通信系统中的节点,其用于数据包的数据流在至少两条链路上的节点到节点的通信,其中,每个数据包被提供有相应的标识符,所述节点包括:

接收装置/接收器和单个公用索引缓冲区,用于接收并缓冲所述数据流的数据包;

发送装置/发送器,用于在所述至少两条链路上将所述数据流的经复制的数据传递到接收节点;

其中,所述标识符对应于数据包在所述数据流中的相对位置,并且其中,所述缓冲数据包的步骤包括:使用从所述至少两条链路中的至少一条链路接收的数据包,在所述缓冲区中重构所述数据流的预定部分;

其中,所述接收节点针对每个所接收到的数据包,记录数据包间到达时间;

其中,所述数据流的转发是以一输出速率来进行的,所述输出速率基于所记录的数据包间到达时间数据来确定。

一种网络通信冗余方法

技术领域

[0001] 本发明涉及用于网络通信的冗余方法的领域,并且更具体地,涉及一种基于在至少两条链路上进行的数据流的节点到节点传送的通信方法。

背景技术

[0002] 一种用以针对通信网络中的连接(例如,异步串行接口(ASI)连接或以太网连接)增加节点至节点通信中的冗余的标准方法是在通信节点之间使用经复制的传输路径、信道或链路,这在下文中被称为1+1连接。典型地,接收节点被布置成从发送节点接收数据流的副本和一连串的数据包。经复制的数据流经由一条主链路和一条备份链路被输送到该接收节点。要注意的是,一条链路可以是指一系列的中间的节点到节点通信。基于各自链路上通信的质量或状况,接收节点中的选择机制确定哪一条链路应当被选为主链路,该节点的输出数据流从该主链路被转发到与该节点相关联的另一节点或一个应用。因此,这种1+1连接的总体思路是这样的:如果主链路遭遇故障,能够向备份链路执行故障切换。然而,由于例如主链路和备份链路上的通信的不同的时间延迟,如果主链路突然发生故障,经转发的数据流中的数据丢失或复制数据形式的中断通常发生在故障切换时。

[0003] 作为对发生故障时切换整个通信链路的替代,已经表明能够基于包含在数据流的各个数据包中的标识符来选择哪些重复数据被转发。欧洲专利申请公布No. EP0854610A2公开了一种用于通过第一独立的以太网通信线路和第二独立的以太网通信线路彼此连接的多个节点之间的以太网通信的冗余方法。相同的数据包在两条通信线路上从发送节点被传送到接收节点,每个数据包都已经添加了用于识别数据内容的数据标识符。在接收处,在第一链路和第二链路上传送的数据被临时到存储到各自的缓冲区中,并且在接收节点中,根据数据标识符来确定在两条各自的通信线路上传送的哪些数据已经首先到达,以及哪些数据被接收为第一接收数据,以供后续使用。在保持连接中的冗余的同时,这提供了一种针对各自的数据部分具有最小引入延迟的冗余方法。

发明内容

[0004] 鉴于以上内容,本发明的一个目标是提供一种改进的用于通信网络的冗余方法。特别地,一个目标是提供在多条链路上的无中断的节点到节点通信,就去往一个节点的传入链路的数量而言,多个链路是可扩展的,也就是说,在功能或者对该节点的实质性重建的需求没有损失的情况下,能够增加或者移除去往该节点的链路。

[0005] 这个目标是通过权利要求1中所限定的根据本发明的一种用于节点到节点通信的方法来实现的。本发明建立在以下基础上:通过在接收节点处使用单个的公用索引播放缓冲区,能够以精确方式执行无中断的故障切换,使得在节点到节点传送期间中没有数据被复制或丢失,在索引播放缓冲区中,利用在多条链路上接收到的数据流中的单独的块(即,帧或数据包),来选择性地重构已传递的数据流。

[0006] 因此,根据本发明的一个方面,提供了一种用于节点到节点通信的方法。本方法包

括:为待传送的数据流中的每个数据包提供各自的标识符;并且在至少两条链路上将所述数据流的经复制的数据从发送节点传递到接收节点。进一步地,本方法包括:在接收节点处,选择性地接收并且缓冲来自所述至少两条链路的数据包,以便后续对所述数据流进行转发。所述标识符对应于数据包在所述数据流中的相对位置,并且所述缓冲的步骤包括:使用从所述至少两条链路中的至少一条链路接收的数据包,在公用索引缓冲区中重构所述数据流的预定部分。

[0007] 本方法的优点在于,本方法能够在相对不稳定的网络中实现,例如,包含较差性能链路的有数据包丢失的以太网网络。由于数据流在多条链路上的节点之间进行传递,并且通过从并行的链路中的一条链路增加一条链路的数据流中的丢失数据包来在每个接收节点中的公用缓冲器中进行重建,因此对每条单独的链路的质量要求得以降低。要注意的是,归因于链路租赁价格,在三条以上不稳定的链路上输送数据比在两条或一条不稳定的链路上输送数据可能更划算。进一步地,利用公用缓冲区来重新创建已传送的数据流有利地提供了一种易于扩展的系统,在这种系统中,到一个接收节点的传入链路的数量能够在不需要对该接收节点进行重新配置的情况下改变。

[0008] 针对已传送的数据流的重构使用公用索引缓冲区的另一个优势在于:接收节点中所需的已分配的缓冲区空间被显著地减小。与已知的系统相比,对于具有两条传入链路的示例,当利用公用共享缓冲区时,数据流的重构所需的最低要求的缓冲区容量是其大小的一半。进一步地,所需的缓冲容量通常由传入链路的实际或估计的差分延迟(即,至少两条链路上的不同的经复制的数据流之间的时间偏移量)来确定。例如,如果具有特定标识符的数据包到达一条链路的时间比到达一条并行的链路的时间晚5毫秒,那么该连接就被称为具有5毫秒的差分延迟。这必须通过缓冲数据流的相应的时间部分来进行补偿,以提供无中断的连接。与对所有链路上所接收到的所有数据包进行缓冲的已知系统相比,对于以上具有两条传入链路的示例,当接收节点具有有限的总的缓冲区空间时,能够实现相对于用于补偿差分延迟的缓冲的两倍的无中断的容忍度。

[0009] 根据本方法的一个实施例,重构所述数据流的至少一部分的步骤包括:仅选择所接收的具有特定标识符的一个数据包,以用于写入所述缓冲区中的每个索引位置。具有特定标识符的多个数据包可以到达不同链路上的节点。然而,由于数据流的重构仅要求具有该特定标识符的数据包的一个副本,当一个副本已经被写入公用缓冲区时,该数据包的其他副本能够被丢弃。选择哪一个副本可以基于当前的选择策略来确定,例如,最先到达、最可靠的链路等。

[0010] 进一步地,这为多条链路上的无中断的节点到节点通信提供了传入链路之间具有接近于零的差分延迟的改进处理,以上描述的现有技术的解决方案在以无中断的方式处理差分延迟上可能具有困难。这是由于以下原因:当多条传入链路之间的差分延迟较小时,实际首先到达的链路可能在时间方面变化,这继而在确定哪一条链路是最先到达的(主链路)、必须完成多少缓冲来提供无中断的故障切换以及节点到节点通信在什么时候处于无中断的状态方面提供了难度。

[0011] 根据本方法的一个实施例,本方法包括:针对每个所接收到的数据包,基于所接收到的数据包的标识符来确定相应的缓冲区索引;基于所确定的缓冲区索引来检验数据是否已经被写入所述相应的缓冲区索引的位置,并且如果数据没有被写入所述相应的缓冲区索

引的位置,则将所接收到的数据包或者与所接收到的数据包相对应的指针写入所确定的缓冲区索引的位置。这使得即便传入的数据包以混乱的顺序到达,也能够对具有特定标识符并且到达传入链路上的数据包进行有效地排序,进入缓冲区中的正确缓冲区索引位置。仅那些仍然还没有被写入缓冲区的数据包被选择用于写入。

[0012] 根据本方法的一个实施例,本方法进一步包括:将在所接收的一数据包被记录之后到达的具有相同标识符的任何所接收到的数据包丢弃,以便恢复存储器。

[0013] 根据本方法的一个实施例,支持同时选择所接收到的数据包并写入所述缓冲区,所接收到的数据包具有不同的标识符并且在不同的链路上被接收。这种同时地写入公用缓冲区有利地增加了数据写入缓冲区的速度,并且因此减少了执行数据流部分的重构所需要的时间。

[0014] 根据本方法的一个实施例,本方法进一步包括:针对每个所接收到的数据包,设置与经确定的缓冲区索引和链路相对应的有效位条目。这提供了一种监视哪些数据包已经在哪一条链路上被接收到的有效方式,这可以例如被用于传入链路的质量监视和无中断的故障切换能力的检测。

[0015] 根据本方法的一个实施例,本方法进一步包括:基于这些有效位条目来监视节点到节点通信的当前冗余,由于复杂的原因,这样做是有利的。这些有效位条目可以进一步被用于监视公用缓冲区是否被充分地填充,使得当一条或多条链路出现故障时,无中断的故障切换能够得以执行。

[0016] 根据本方法的一个实施例,本方法进一步包括:在接收节点处,针对每个所接收到的数据包,记录数据包间到达时间。数据流的转发是以一输出速率来进行的,所述输出速率基于所记录的数据包间到达时间数据确定,这对于最小化已传送的数据流中的抖动/漂移效应的引入是有利的。例如,在以太网传输格式以及去往一节点的数据包到达速率不固定的其他格式或服务(例如,以太网网络上的视频业务)中实现本方法时,这是具有重要意义的。

[0017] 根据本方法的一个实施例,输出速率是逐个数据包来确定的,使得对于特定缓冲区索引的数据包,所述输出速率是基于所接收到的数据包中具有与所述特定缓冲区索引相对应的标识符的那个数据包的所记录的数据包间到达时间来选择,这对于确保这些数据包以确保与始发的数据流具有相同的抖动/漂移特性的方式进行转发是有利的。

[0018] 根据本方法的一个实施例,转发数据流时的输出速率是基于来自所述至少两条链路中的其中一条链路的所记录的数据包间到达时间来选择,这在需要强迫主定时链路来达到转发的数据流的特定的输出速率是有利的。这确保了转发的数据流的发出定时变成与传入的至少两个数据流的发出定时相同,而不是随意的混合。

[0019] 根据本方法的一个实施例,数据流的转发是从已经在至少两条链路上被接收到所述数据流中的第一数据包之后的一段预定时间开始的。

[0020] 根据本方法的一个实施例,本方法可以被用于实现布置在至少两条链路中的一条链路中的中间节点的无中断节点升级,从而使该中间节点能够在不中断业务的情况下进行无中断的拆卸以用于维护。

[0021] 根据本方法的一个实施例,本方法有利地被实现为软件模块,该软件模块用于在被计算机处理器执行时,实施如以上所描述的方法。

[0022] 当对以下详细的说明书、附图和所附的权利要求书进行研究时,本发明的其他目的、特征和优点将变得明显。本领域技术人员意识到本发明的不同特征能够被组合以创建出与在下文中所描述的实施例不相同的实施例。

附图说明

[0023] 通过参照附图对本发明的优选实施例的以下说明性并且非限制性的详细描述,本发明的以上以及额外的目标、特征和优点将被更好地理解,其中,相同的附图标记将被用于相似的元素,其中:

[0024] 图1是根据本发明的一种节点到节点通信系统的示意性的概观;

[0025] 图2至图6是根据本发明的实施例的功能结构的示意图;

[0026] 图7和图8是通过利用本发明的方法对网络中的节点进行无中断的断开的示意图。

[0027] 所有的附图都是示意性的,不一定是按照比例绘出的,并且大体上仅示出了必要的部分,以便阐明本发明,其中,其他部分可以被省略或仅仅是建议的。

具体实施方式

[0028] 现在参照附图对根据本发明的用于节点到节点通信的方法的示例性的实施例进行描述。在这些示例性的实施例中的大多数中,对1+1连接进行了阐述。然而,应当注意的是,所描述的机制也适用于1+1+1连接以及诸如此类的连接。虽然本方法的步骤被显示成一个数字序列,但是这些步骤中的一些步骤可以按照另一种顺序来执行。贯穿该文档,术语“数据包”或“帧”被用作数据粒度,尽管在所有的情况中数据可以不按照自身已分成包的格式被输送。

[0029] 图1是概括了根据本发明的构思的一种节点到节点通信系统10的示意图。这里,该通信系统10包括了使用根据本发明的用于节点到节点通信的方法的一个无中断的1+1以太网连接。该节点到节点通信系统10包括第一节点N0和第二节点N1,第一节点N0和第二节点N1经由两条并行的链路A和B相互连接。每个节点均包括用于接收、处理和发送数据流的模块。然而,为了简化的目的,仅在本文中显示和描述了对理解本发明是必要的这些功能部件。

[0030] 在第一节点N0和第二节点N1之间待传送的数据流11包含一系列的数据包。数据流11中的每个数据包在第一节点N0中均被提供有与该数据包在数据流11中的相对位置相对应的标识符。该标识符可以是在一个预定间隔内选择出的一个序号。然而,根据本方法的实施例,标识符可以被实现为在数据流传送的出口方向上的任何形式的时间戳。

[0031] 进一步地,在提供有标识符n之后,每个有索引的数据包随后在第一节点N0中被复制,使得相同的数据(即,包含针对该数据包的相同的序号n的经复制的数据包)从第一节点N0分别越过第一链路A和第二链路B被传送到第二节点N1,即,从发送节点到接收节点。在第二节点N1处,来自链路A和B的数据流中的数据包的经传送的副本被接收到,并且被选择性地拒绝或缓冲在公用缓冲区(在图1中未示出,参见例如在图2中示意性地图示为表120的公用缓冲区)内,用于该数据流随后在(外向的)链路C上的转发。公用缓冲区120,单个的共享的播放缓冲区,这里使用序号 b_n 来进行索引并且能够由链路A和B同时写入。

[0032] 现在参照示出本发明的方法的基本的功能结构(如图2中示意性示出的),对本文

中描述的本方法的实施例进行更详细的说明。这里,该基本的功能结构100由传入链路A和B(尽管本方法支持任何合适数据的链路A、B、...、Z)、选择单元110(从性能监视的观点来看,每条传入链路均终止于该选择单元110)以及播放缓冲区120(与参照图1所描述的布置在接收节点N1中的公用的索引缓冲区相对应)。链路A和B上传输的数据中的每个数据包均包含各自的序号,该序号在发送节点中被写入数据流中的数据包内。这里,每个序号的每个数据包针对链路A被称为A(n),并且针对链路B被称为B(n)。

[0033] 播放缓冲区120是编入索引的循环数组。播放缓冲区120进一步包括每一条链路的每个条目的有效位VA和VB、以及播放缓冲区索引 b_n 。选择单元110包括用于基于每个接收到的数据包的标识符(即,数据包的序号)来确定针对该数据包的相应播放缓冲区索引 b_n 的模块。这通过选择单元110从链路A或B检测条目(例如,接收到的数据包或数据包),并且通过将播放缓冲区头条目的数据包的序号的偏移量和接收到的序号进行比较来计算出相应播放缓冲区索引 b_n 来得以实现。进一步地,针对每个条目,如果每一条链路的每一个条目的有效位VA和VB都被清除,那么该数据包被写入播放缓冲区(未示出),并且针对播放缓冲区索引 b_n 的链路的有效位VA或VB被设置。如果对于一条链路的一个特定的条目,另一条链路的针对播放缓冲区索引 b_n 的有效位已经被设置,即另一条链路已经填满条目并且数据包已经被写入播放缓冲区,那么该特定的条目的数据包在设置其有效位之后被丢弃。在图2中,显示了当将数据流转发到链路C时播放缓冲区读取顺序BRO。

[0034] 现在考虑图3中示例的播放缓冲区填充的一个示例。链路A超前于链路B,即,由选择单元110所检测到的来自链路A的数据包的序号大于来自链路B的数据包的序号。来自链路A的数据包一个接一个地被写入播放缓冲区。在 t_1 时刻,A(n+3)到达选择单元110,A(n+3)的数据包被写入播放缓冲区120,并且针对链路A,设置与序号(n+3)相对应的播放缓冲区索引 b_n 的有效位(对于 $b_n=n+3,VA=1$)。在 t_2 时刻,播放缓冲区120已经被写到A(n+4),然而正如能够在针对链路B的有效位VB中看到的,在链路B上仅仅收到了B(n)、B(n+1)和B(n+2),在链路B上,相应的有效位被设置,并且相应的数据包被丢弃,因为它们已经被链路A接收到了。如图3中所示,播放缓冲区写入顺序BW0于是变成了A(n)、A(n+1)、A(n+2)、A(n+3)、A(n+4)。

[0035] 在1+1连接中,差分延迟 Δ 是在不同链路上发送的经复制数据包的两个副本之间测量到的到达时间的差。关键在于播放缓冲区是足够大的,使得不同链路之间的差分延迟 Δ 能够被覆盖。再次参照图3,来自链路B的数据包B(n+3)在 $t_3=t_1+\Delta$ 时刻到达选择单元110,并且由于播放缓冲区已经由链路A写入了,因此在针对与序号(n+3)相对应的播放缓冲区索引 b_n ,设置了播放缓冲区索引的有效位VB之后(在时间 $>t_3$ 时,对于 $b_n=n+3,VB=1$),B(n+3)的数据包将会被丢弃。在前的链路A由于某种原因出现故障,并且不能传送A(n+3)的情况下,播放缓冲区中的缓冲必须被布置成覆盖链路A和链路B之间的差分延迟 Δ ,使得当B(n+3)到达时,B(n+3)能够被用于替代丢失的数据包并且恢复数据流,参见下面的例子。

[0036] 在图4中,着重示出了无中断的故障切换的例子。如图4所示,这里链路A超前于链路B,即,由选择单元110所接收到的来自链路A的数据包的序号大于来自链路B的数据包的序号。这里,链路A出现了故障,即,在A(n+2)之后没有从链路A接收到有效数据包,但是,由于链路A超前于链路B,选择单元110将数据包A(n)、A(n+1)和A(n+2)写入播放缓冲区。如图4所示,链路B将最终赶上并且开始填充播放缓冲区条目,从序号(n+3)开始填充。由此,链路

到链路通信的无中断的故障切换得以实现。如图4中所示,播放缓冲区写入顺序BW0于是变成了A(n)、A(n+1)、A(n+2)、B(n+3)、B(n+4)。

[0037] 现在参照图5,图5图示了同时写入播放缓冲区,链路A和链路B可以同时写到播放缓冲区。也就是说,选择单元110被布置为支持所接收到的数据包或者具有不同标识符(例如,序号)的数据包的同时选择和写入,所接收到的数据包或具有不同标识符的数据包是在不同的链路(这里指链路A和链路B)上接收到的。在该例子中,链路A超前于链路B,但是由于某些以前的故障,存在数据包丢失的情况。当具有序号(n+3)的数据包从链路A被写入时,具有序号(n)的数据包同时地从链路B被写入。此后,根据播放缓冲区写入顺序BW0,随着具有序号(n+4)的数据包由链路A写入,具有序号(n+1)的数据包同时由B写入,依次类推。如图5所示,播放缓冲区写入顺序BW0于是对于某些时间实例具有多个条目。在 t_5 时刻,仅链路B已经传送了具有序号n至(n+2)的数据包,因此对于这些数据包,针对链路A的有效位VA没有被设置,而针对链路B的有效位VB被设置了。正如能够从图5推断出的,将设置有效位VA和有效位VB的第一数据包将是具有序号(n+3)的数据包。

[0038] 根据本方法的一个实施例,在针对不同的链路A、B、...、Z的公用缓冲区中所记录的有效位VA、VB、...、VZ被用于控制数据流向外向链路C中的转发。在该实施例中,数据流的转发是从该数据流中的第一帧或数据包已经在至少两条链路上被接收到之后的一段预定时间开始的。该预定时间被选择成提供一个足够大的缓冲区,该缓冲区覆盖网络所支持的最大差分延迟。

[0039] 根据本方法的一个实施例,在针对不同的链路A、B、...、Z的公用缓冲区中所记录的有效位VA、VB、...、VZ被用于基于有效位条目来监视节点到节点通信的当前冗余。缓冲区监视是通过这样的方式来执行的,当随后从第二节点(参见例如图1中的节点N1)转发数据流时,针对来自播放缓冲区(或播放队列)的每个经转发的数据包,检查针对链路A和B的有效位。如果针对特定序号的数据包,有效位VA和VB均被设置,那么这意味着链路A和B在该数据包从播放缓冲区被转发之前到达了该序号,这因此表明了缓冲的水平足以处理无中断的故障切换。相反地,如果只有一个位被设置,那么缓冲量是不充分的,并且连接因此不能够被认为处于无中断的状态,即处于无中断的故障切换能够被提供的状态。对于多条链路上的节点到节点通信,上文所描述的缓冲区监视能够被用于确定连接的单倍冗余、双倍冗余、三倍冗余,等等。

[0040] 根据本发明的一个实施例,当启动节点到节点通信时,以参照图1和图2所描述的节点到节点的通信系统作为一个示例性的系统,启动条件按照以下方式进行处理:最初链路A和链路B均未建立,并且连接状态被称为是非无中断的状态。随后,链路A或链路B已建立。然而,这两条链路的差分延迟被测量为在所支持的无中断的界限以下,直到链路A和链路B都已建立,并且正如前面提到的,如果从播放队列被转发的数据包已经设置了两个有效位,那么连接状态转换到无中断的状态。

[0041] 在一个实施例中,一旦第一链路出现,该第一链路就开始写入播放缓冲区。当该播放缓冲区达到预定的占用(例如,50%)时,数据包开始被转发。或者,在针对链路A和B已经对头条目的数据进行验证(即,VA和VB已经被设置)之后的一段预定时间,开始数据流的转发。优选地,数据流的转发是受到在下文中的本发明的实施例中描述的预定的发送速率机制的支配来执行的。优选地,对该发送速率机制进行选择,以便维持数据流的最初的发送速

率。当第二链路达成时，第二链路要么超前于现存的链路，要么滞后于现存的链路。如果第二链路超前，则在前的链路将会赶上并且开始填充条目。如果该第二链路滞后，则差分延迟被建立。在一个实施例中，如果差分延迟小于所支持的最大延迟，并且在从第一连接被建立的一段预定的时间范围（例如，10秒）内进入该条件，那么引入自动中断，以迫使连接进入无中断的状态。然而，如果该差分延迟小于所支持的最大延迟，并且在一段预定的时间段（例如，10秒）之后进入该条件，那么需要手动强制执行无中断的操作，以触发中断并且使连接进入无中断的状态。

[0042] 在一个实施例中，在已经引入中断之后，两条链路都是可用的。在这种情况下，对各自链路的序号进行比较，并且选择具有最小的序号的链路作为开始写入播放缓冲区的链路，并且因此建立播放缓冲区头条目。在另一个实施例中，采用随机链路选择机制。

[0043] 要注意的是，通过在将数据包转发到播放缓冲链路中之前等待缓冲区占用达到50%，会引入一定的延迟。如果差分延迟小于50%的缓冲区占用对应的业务量，则该延迟可能是大于提供无中断的故障切换所需要的延迟。对于延迟敏感的应用，因此总是推荐在连接建立之后强行中断，使得缓冲量能够得以最佳化，并且因此延迟能够得以最优化。

[0044] 正如上面所提到的，为了减小节点到节点通信中的抖动/漂移，可以采用不同的发送速率机制，这将在本文下方针对本发明的方法的实施例进行描述。

[0045] 根据本方法的一个实施例，为了减小节点到节点通信中的抖动/漂移，对播放缓冲区的数据包发送速率进行控制，以便维持在与数据包到达接收节点的速率相同的播放缓冲区发送速率。现在参照图6，本方法进一步包括：在接收节点N1处并且针对每个接收到的数据包，记录数据包间的到达时间。每当数据包到达链路A或B，选择单元110中的循环计数器60将被启动，并且当下一个数据包到达时，该循环计数器的值被记录并且与该到达的数据包相关联。也就是说，在接收节点N1处，每条链路上到达的数据包之间的循环次数被记录。进一步地，基于所记录的数据包间的到达时间数据，转发数据流时的输出速率是例如通过针对具有某一序号的首先接收到的数据包，选择数据包间的到达时间数据来确定的，参见下面的例子。所确定的时间被称为发送延迟d。一旦一个数据包变成了队列的头（如图6中的头ptr所表示）并且准备被发出，在该数据包被发出之前，在由发送延迟d所指示的循环次数内保持该数据包。该机制确保了数据包以与它们到达时的速率相同的速率被发出。发送延迟序列 T_a 能够被用来描述一条链路的数据包的发送延迟特性：

[0046] $T_{a\text{链路}} = [d_{\text{链路}(n)}, d_{\text{链路}(n+1)}, \dots, d_{\text{链路}(N)}]$ 。

[0047] 根据本方法的实施例，输出速率是逐个数据包来确定的，使得对于特定缓冲区索引 b_n 的数据包，输出速率是基于接收到的数据包中具有与该特定缓冲区索引相对应的标识符的那个数据包的发送延迟d来选择的。考虑三条传入链路A、B、D（未示出）。对于具有序号n+3的数据包，至多确定三个发送延迟 $d_{A(n+3)}$ 、 $d_{B(n+3)}$ 、 $d_{D(n+3)}$ ，并且在该数据包被发送到外向链路C上时， $d_{B(n+3)}$ 被选为针对该数据包的发送延迟。对于具有序号n+4的下一个数据包，至多确定三个发送延迟 $d_{A(n+4)}$ 、 $d_{B(n+4)}$ 、 $d_{D(n+4)}$ ，并且在该数据包被发送到外向链路C上时， $d_{D(n+4)}$ 被选为针对该数据包的发送延迟。对于具有序号n+5的下一个数据包，至多确定三个发送延迟 $d_{A(n+5)}$ 、 $d_{B(n+5)}$ 、 $d_{D(n+5)}$ ，并且这里的 $d_{A(n+5)}$ 被选择。针对这三个数据包所导致的播放缓冲区的发送延迟序列因此对应于：

[0048] $T_{aC} = [..d_{B(n+3)}, d_{D(n+4)}, d_{A(n+5)} \dots]$ 。

[0049] 根据本方法的一个实施例,转发来自接收节点的数据流时的输出速率是基于来自这些传入链路中的其中一条的所记录的数据包间到达时间来选择的。考虑几乎同步的两条链路A和B,即这两条链路A和B上到达的数据包的序号彼此非常接近,但是在到达率方面,这些链路包含一定的可变性。如果在逐个数据包的基础上确定发送延迟,那么这种情况的结果可以是在这两条链路之间高度交叉的一个播放缓冲区。因此,链路C的发送速率特性可以很明显地不同于各自的链路A和B上到达的数据流中的任一个数据包的发送速率。例如,来自链路A的数据包到达选择单元110,产生一个确定的发送延迟序列: $T_{aA}=[15,2,15,2,15]$ 。进一步地,来自链路B的数据包到达,产生一个确定的发送延迟序列: $T_{aB}=[2,15,2,15,2]$ 。让我们假设这些链路都是几乎同步的,并且我们得到一个播放缓冲链路,在该播放缓冲链路中,每隔一个数据包是从交替的链路中挑选出的。针对该播放缓冲区的发送延迟序列于是变成了 $T_{aB}=[15,15,15,15,15]$,这产生了针对链路C中转发的数据流的发送速率,该发送速率非常不同于链路A或B中的数据流中的数据包的到达速率。这可能是不期望的。这里,这是通过允许这些链路中的其中一条链路称为主定时链路来解决的。例如,滞后的链路被选为主定时链路。该主定时链路然后重写有效播放缓冲区条目的发送延迟。由于缓冲,这将允许播放缓冲链路得到与该滞后的链路相同的延迟序列。发送延迟序列于是变成了 $T_{aC}=[15,2,15,2,15]$ 或 $[2,15,2,15,2]$,这导致了数据流在链路C中被转发的发送速率与链路A或B中的数据流中的数据包的到达速率对应相同。将哪条链路作为主定时链路可以基于例如下面的标准来进行选择:出现的第一条链路、超前于其他链路的链路、基于链路稳定性的历史信息的链路、或者是基于上述标准的任意组合的链路。

[0050] 进一步有利地,本发明的方法能够被用于实现网络中的中间节点的无中断节点升级。现在参照图7,其示出了当本方法被用于在无中断方式中断网络中的一个节点时(例如,为了升级该节点),根据本发明的节点到节点通信系统700的一个实施例。在正常业务期间,第一节点N0经由中间节点NI被连接到第二节点N1。现在,节点NI需要在不中断业务的情况下临时地从网络中断开。这通过下面的方式来实现:从第一节点N0增加一条冗余链路NS到第二节点N1,从而对节点NI进行旁路,以便进行升级。正如前面所描述的,数据流11在第一节点N0被复制,并且现在经由该中间节点NI和该冗余链路NS被发送到第二节点N1。一旦连接被确认是无中断的,中间节点NI上的原始连接就能够被拆除,并且中间节点NI就能够被升级了。一旦被升级完,NI上的原始连接就被恢复,并且一旦被确认是无中断的,该冗余的旁路连接NS就能够被移除。该机制提供了在不中断业务的情况下拆除需要维护的节点的能力。

[0051] 如图8所示,本方法和系统可以有利地被用于多播系统中连接的无中断的断开。首先参见图8a,多播节点NM的分支遍布多播节点到节点通信网络,并且这里被连接到第一节点N0、第二节点N1和第三节点N2。为了提供该多播节点NM的无中断的连接,例如为了升级该节点,现在参见图8b,一条冗余链路NS被增加在该第一节点和该第三节点之间。随后,多播从第一节点N0到第二节点N1并且从第一节点N0到第三节点N2(在冗余链路NS)被建立。一旦两条连接都被确认是无中断的,则该多播节点NM能够被拆除并且被升级。一旦被升级完,NM上的原始连接被恢复,并且一旦被确认是无中断的,该冗余的旁路连接NS就能够被移除,并且来自第一节点N0的多播信道被拆除。

[0052] 通过适于执行本发明的本方法的软件模块(未示出),本发明的方法的实施例优选

地被实施在节点到节点通信中。该软件模块可以被集成在一个包括适当的处理模块和存储器模块的节点内,或者可以被实施在一个包括适当的处理模块和存储器模块外部设备中,并且该外部设备被布置成与现有节点互连。

[0053] 本领域技术人员意识到本发明决不被局限于以上描述的实施例。相反地,在所附的权利要求书的范围内,很多修改和变型都是可能的。进一步地,必须要注意的是,本发明的方法适用于节点到节点通信和任何适当网络中的无中断的节点升级,同样适用在同步网络中。

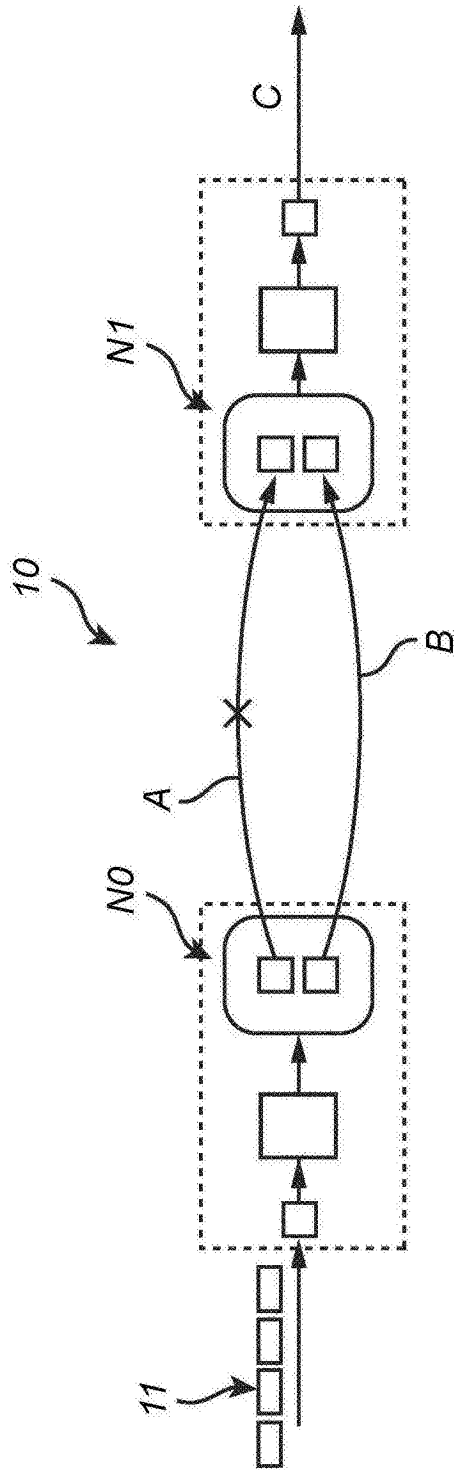


图1

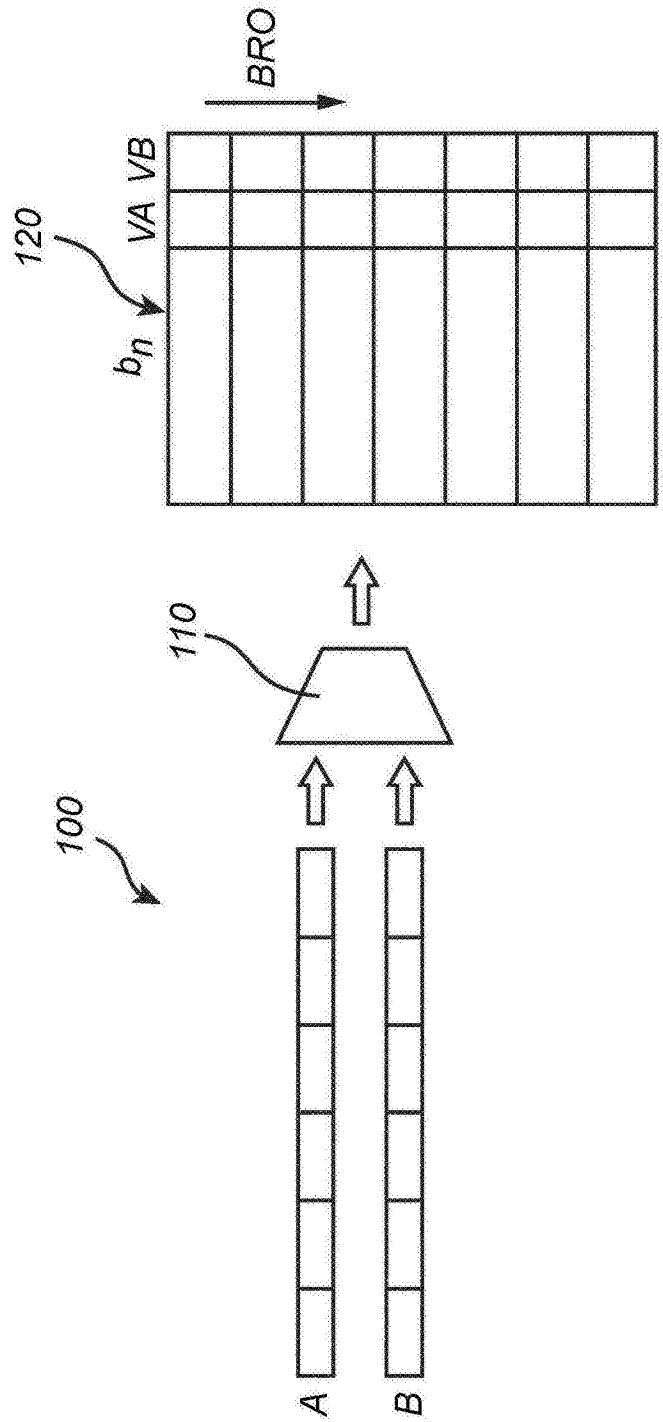


图2

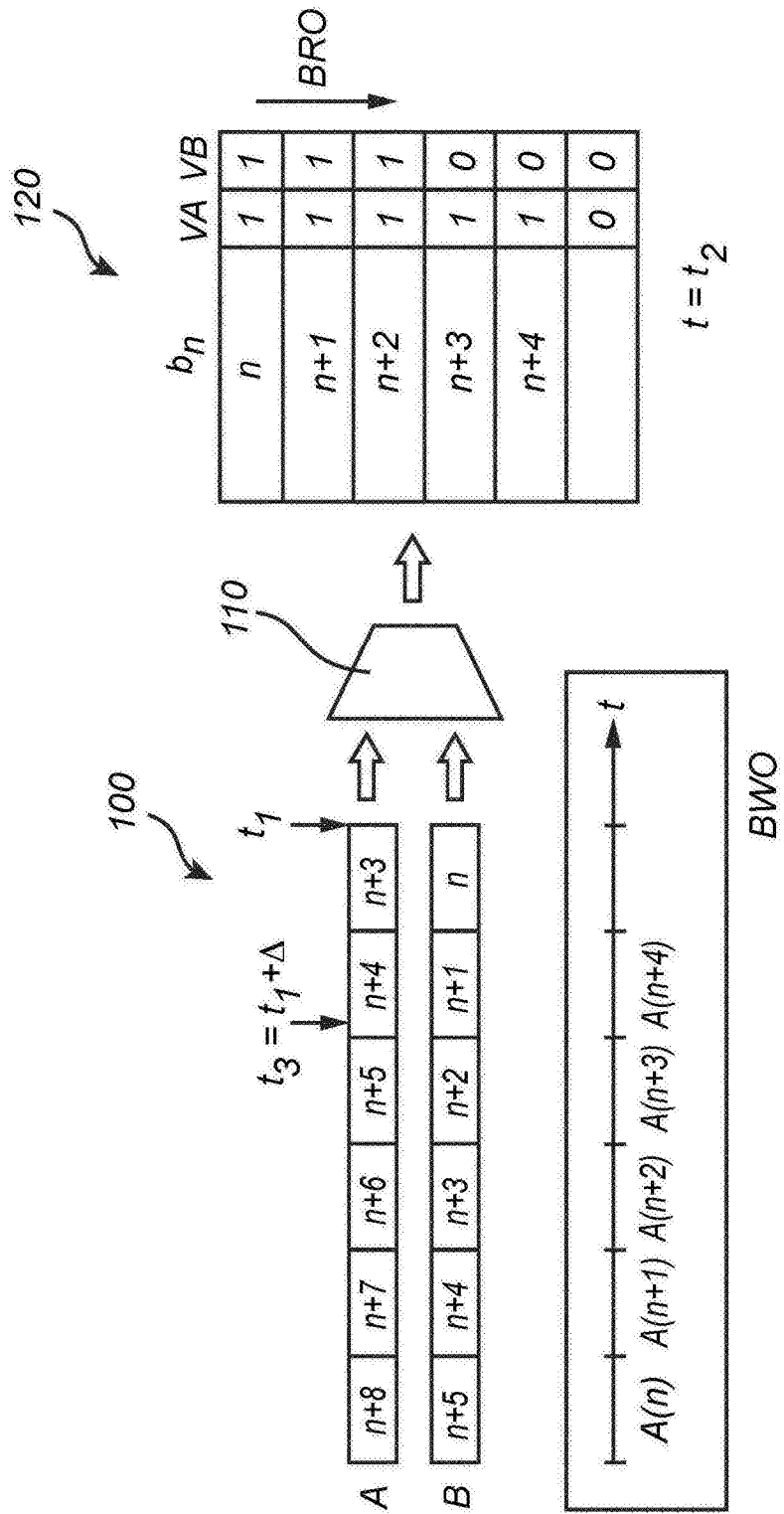


图3

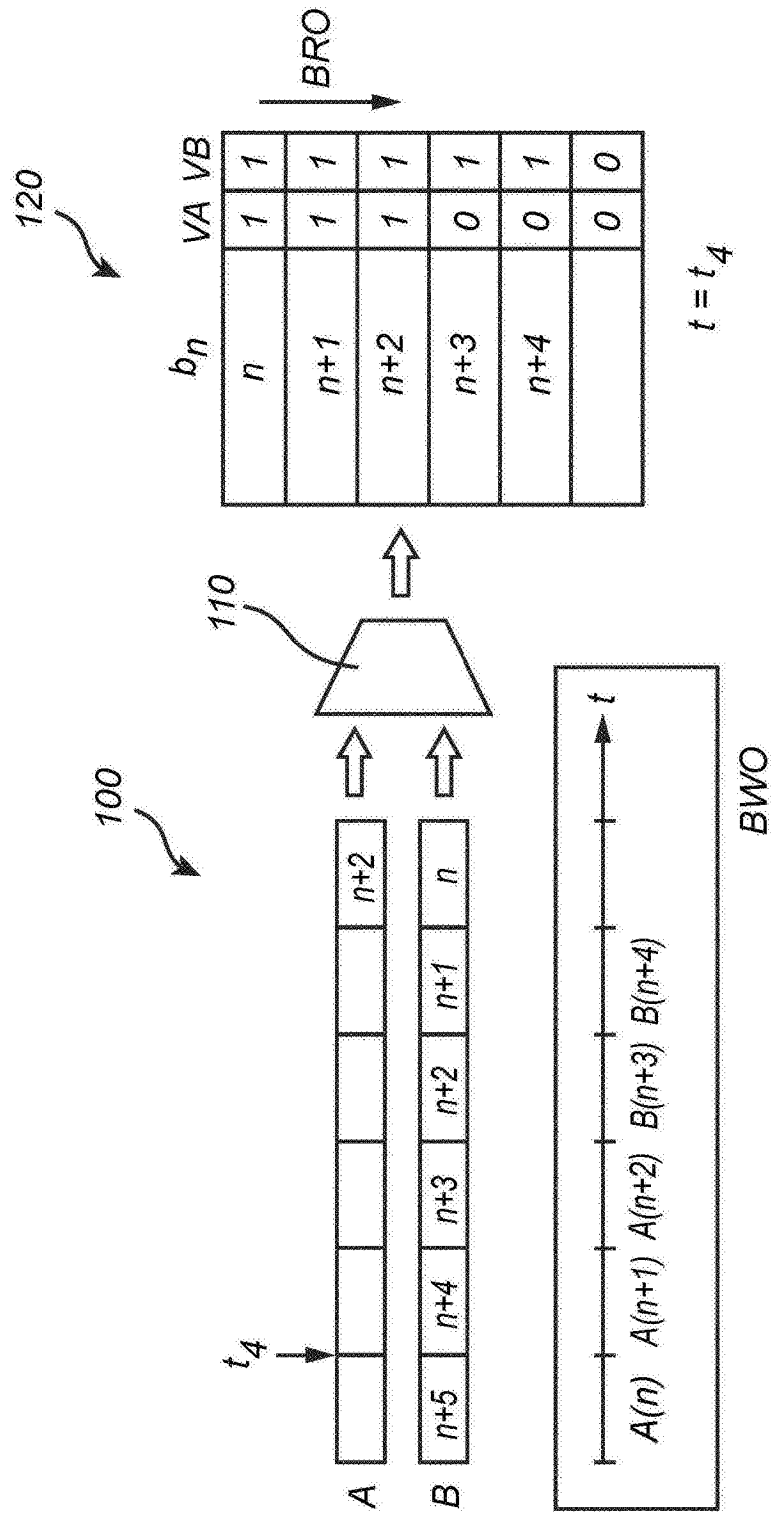


图4

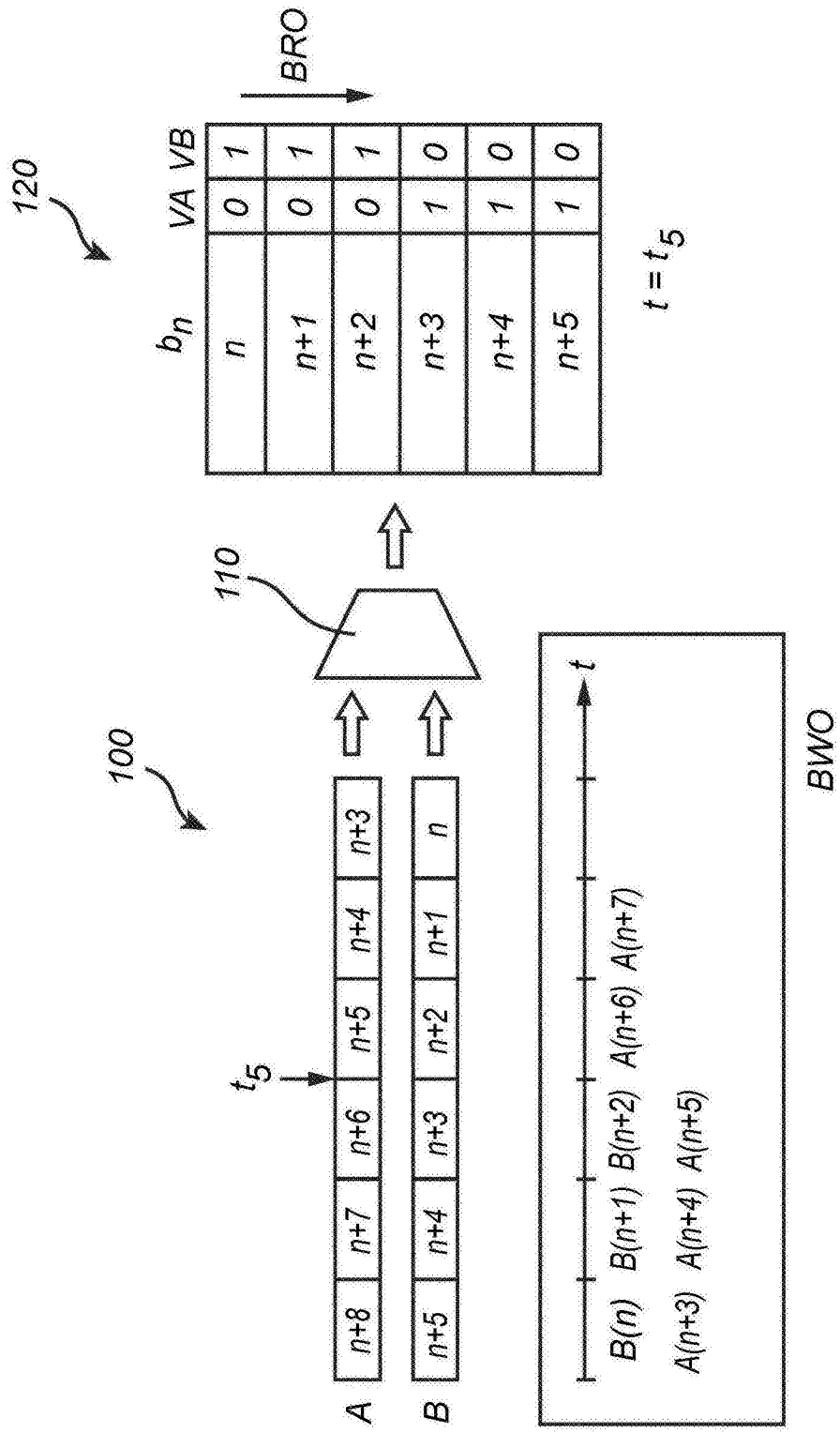


图5

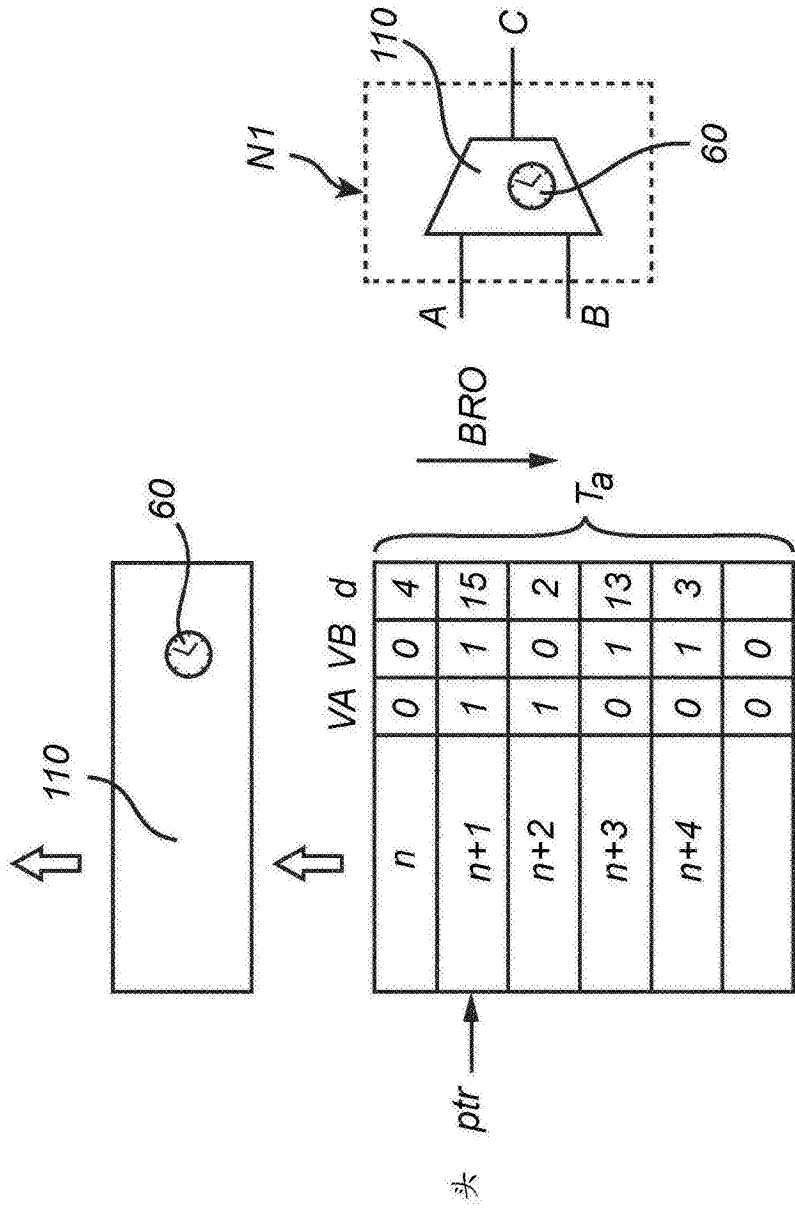


图6

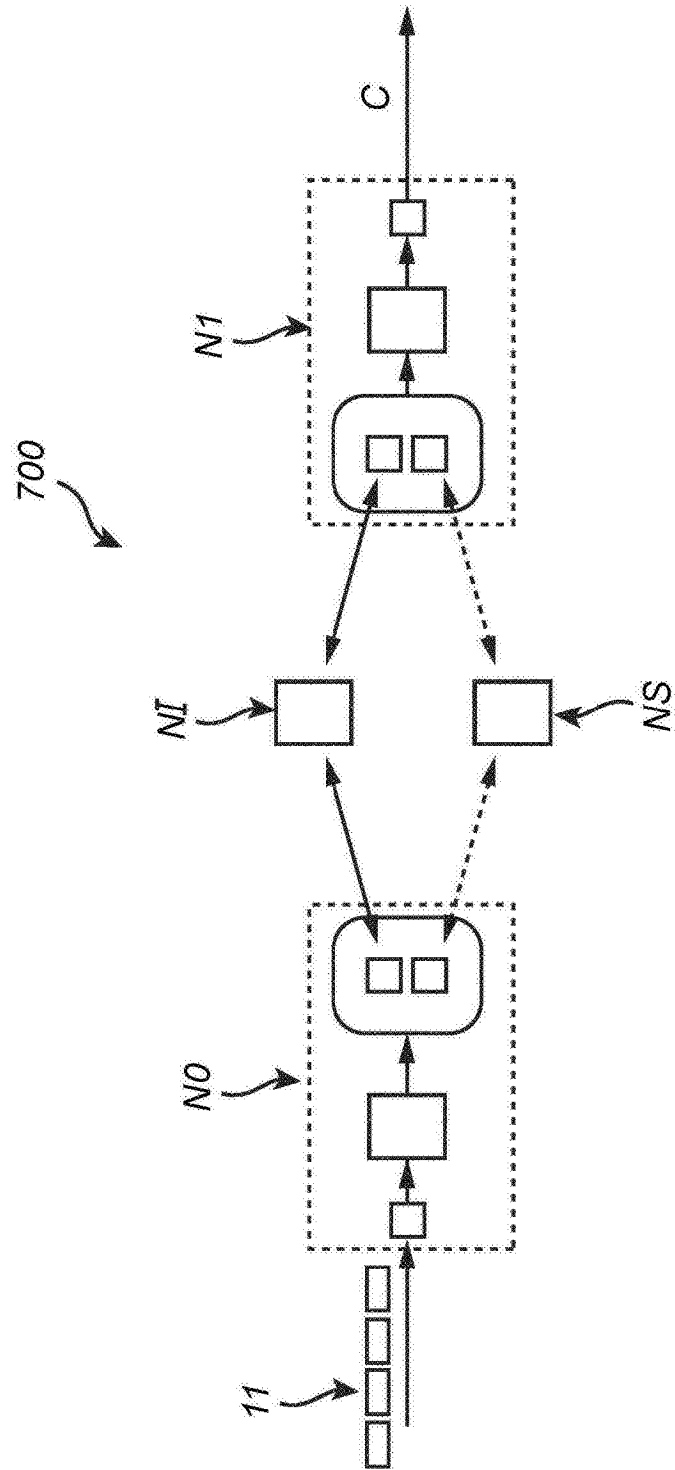


图7

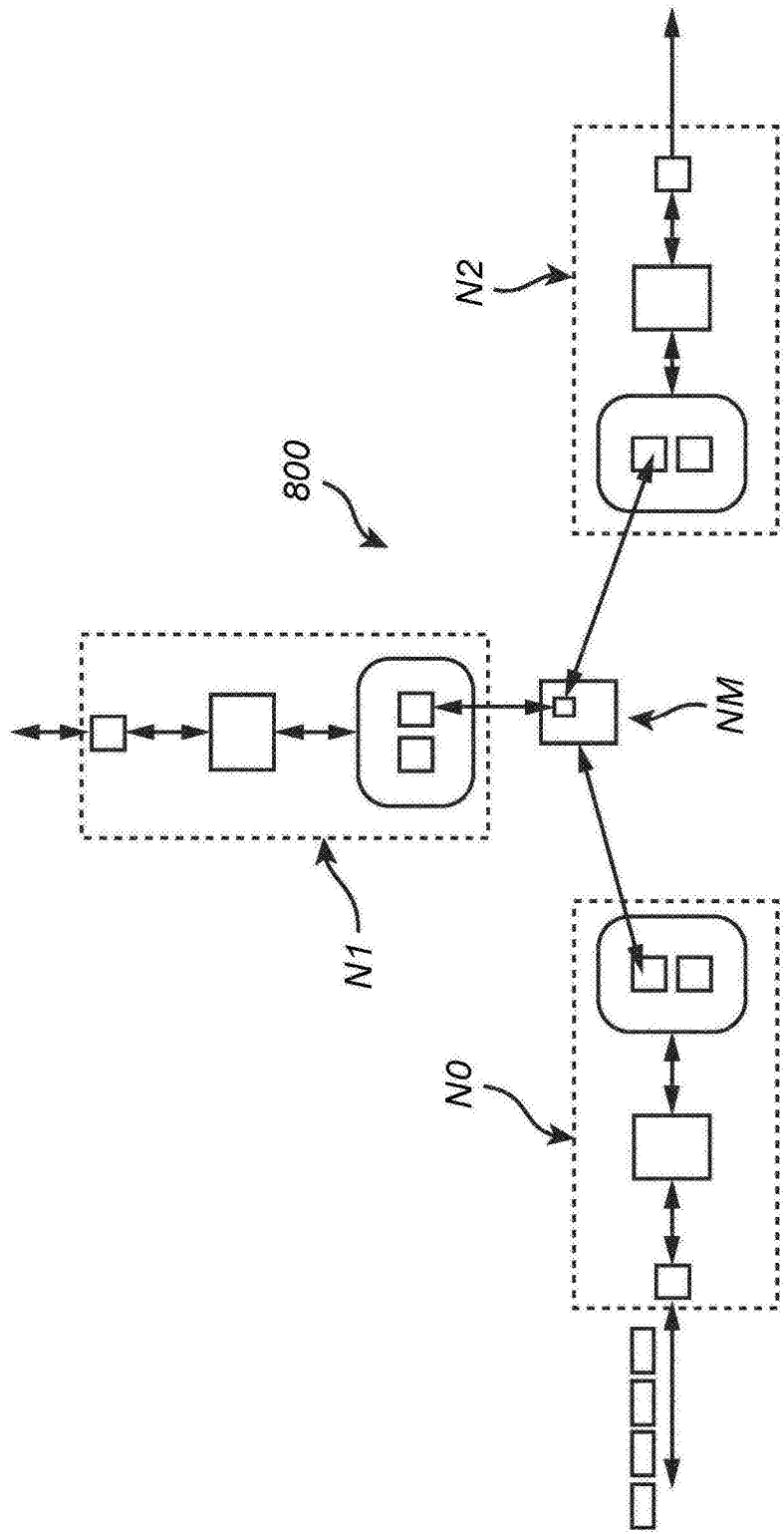


图8a

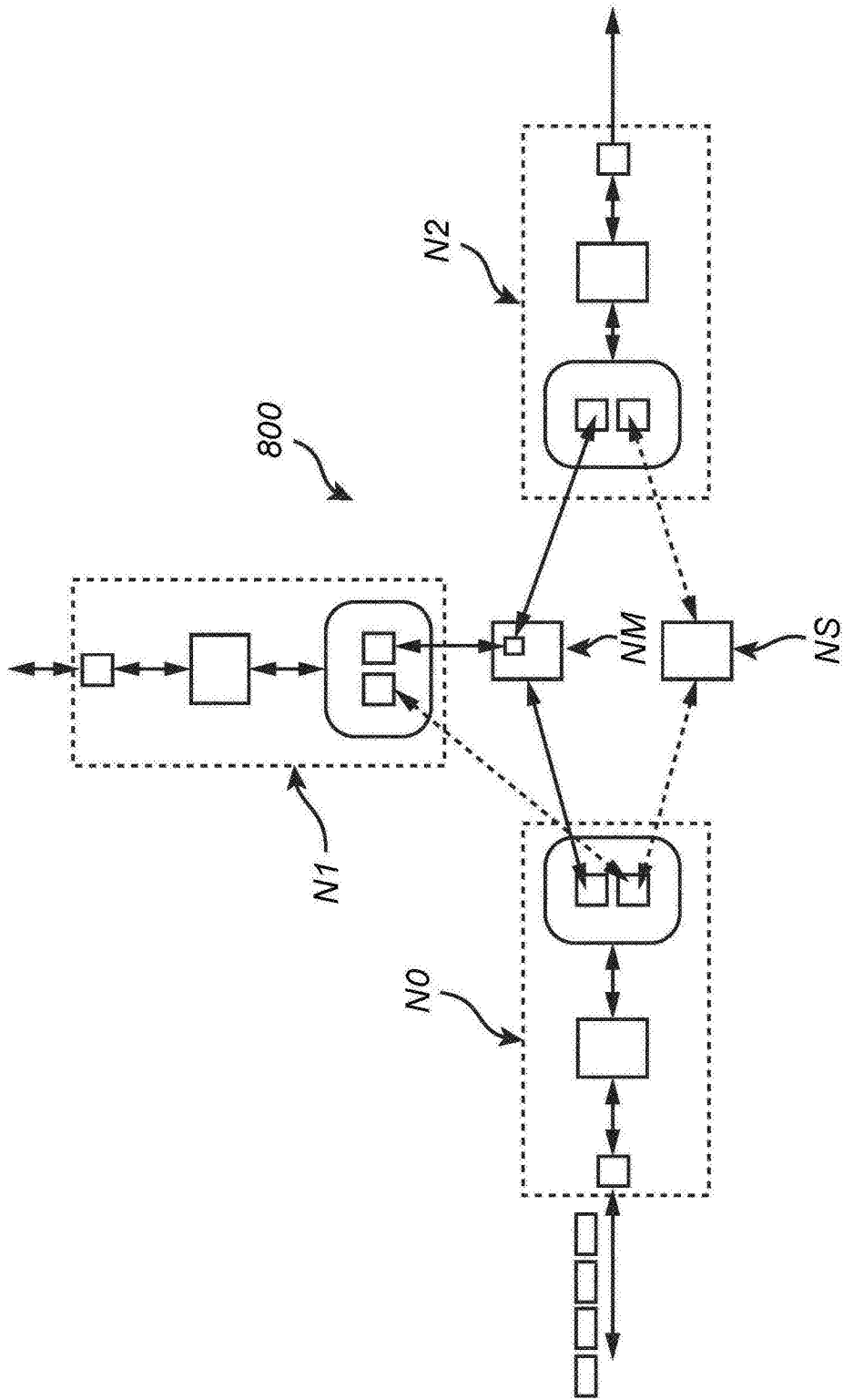


图8b