



(12) 发明专利申请

(10) 申请公布号 CN 103825766 A

(43) 申请公布日 2014. 05. 28

(21) 申请号 201410073061. 0

(22) 申请日 2014. 02. 28

(71) 申请人 杭州华三通信技术有限公司

地址 310053 浙江省杭州市高新技术产业开发区之江科技工业园六和路 310 号华为杭州生产基地

(72) 发明人 汪江波

(74) 专利代理机构 北京博思佳知识产权代理有限公司 11415

代理人 林祥

(51) Int. Cl.

H04L 12/24 (2006. 01)

H04L 12/26 (2006. 01)

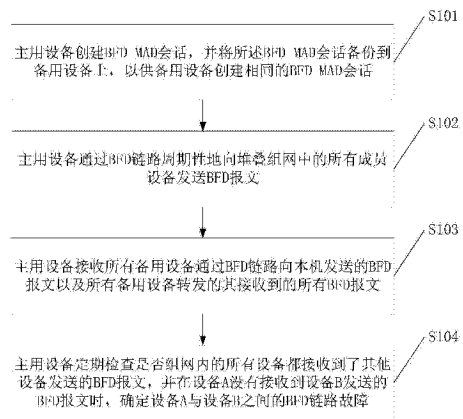
权利要求书2页 说明书5页 附图3页

(54) 发明名称

一种 BFD 链路检测装置和方法

(57) 摘要

本发明提供一种 BFD 链路检测装置和方法, 应用在堆叠组网中的主用设备上。所述方法包括: 创建 BFD MAD 会话, 并将该 BFD MAD 会话备份到备用设备上, 以供备用设备创建相同的 BFD MAD 会话; 通过 BFD 链路周期性地向堆叠组网中的所有备用设备发送 BFD 报文; 接收所有备用设备通过 BFD 链路向本机发送的 BFD 报文以及所有备用设备转发的其接收到的所有 BFD 报文; 定期检查是否组网内的所有设备都接收到了其他设备发送的 BFD 报文, 在成员设备 A 没有接收到成员设备 B 发送的 BFD 报文时, 确定成员设备 A 与成员设备 B 之间的 BFD 链路故障。本发明能够及时地检测出 BFD 链路故障。



1. 一种双向转发检测 BFD 链路检测装置,应用在堆叠组网中主用设备上,其特征在于,所述装置包括:

第一会话单元,用于创建 BFD MAD 会话,并将所述 BFD MAD 会话备份到备用设备上,以供备用设备创建相同的 BFD MAD 会话;

第一发送单元,用于通过 BFD 链路周期性地向堆叠组网中的所有备用设备发送 BFD 报文;

第一接收单元,用于接收所有备用设备通过 BFD 链路向本机发送的 BFD 报文以及所有备用设备转发的其接收到的所有 BFD 报文;

链路检测单元,用于定期检查是否组网内的所有成员设备都接收到了其他成员设备发送的 BFD 报文,在成员设备 A 没有接收到成员设备 B 发送的 BFD 报文时,确定成员设备 A 与成员设备 B 之间的 BFD 链路故障。

2. 根据权利要求 1 所述的装置,其特征在于,

所述 BFD 报文是组播报文,所述组播报文的目的 MAC 地址是预设的组播 MAC;

所述 BFD MAD 会话的源 IP 地址是主用设备的 IP 地址,目的 IP 地址是备用设备自身的 IP 地址;

所述主用设备发送的 BFD 报文的源 IP 地址、目的 IP 地址和备用设备发送的 BFD 报文的源 IP 地址、目的 IP 地址相同。

3. 根据权利要求 1 所述的装置,其特征在于,

所述链路检测单元,进一步用于根据所述备用设备发送的 BFD 报文中携带的成员 ID 检查是否组网内的所有成员设备都接收到了其他成员设备发送的 BFD 报文;

所述成员 ID 携带在 BFD 报文 My Discriminator 字段中预设的偏移值后面。

4. 根据权利要求 1 所述的装置,其特征在于,

所述链路检测单元,进一步用于在成员设备 A 没有接收到成员设备 B 发送的 BFD 报文,但成员设备 B 接收到成员设备 A 发送的 BFD 报文时,确认成员设备 A 到成员设备 B 的 BFD 链路单方向故障。

5. 根据权利要求 1 所述的装置,其特征在于,

所述第一接收单元从主用设备和备用设备之间的内部通道接收所有备用设备转发的其接收到的所有 BFD 报文。

6. 一种 BFD 链路检测方法,应用在堆叠组网中主用设备上,其特征在于,所述方法包括:

创建 BFD MAD 会话,并将所述 BFD MAD 会话备份到备用设备上,以供备用设备创建相同的 BFD MAD 会话;

通过 BFD 链路周期性地向堆叠组网中的所有备用设备发送 BFD 报文;

接收所有备用设备通过 BFD 链路向本机发送的 BFD 报文以及所有备用设备转发的其接收到的所有 BFD 报文;

定期检查是否组网内的所有成员设备都接收到了其他成员设备发送的 BFD 报文,在成员设备 A 没有接收到成员设备 B 发送的 BFD 报文时,确定成员设备 A 与成员设备 B 之间的 BFD 链路故障。

7. 根据权利要求 6 所述的方法,其特征在于,

所述 BFD 报文是组播报文,所述组播报文的目的 MAC 地址是预设的组播 MAC;

所述 BFD MAD 会话的源 IP 地址是主用设备的 IP 地址,目的 IP 地址是备用设备自身的 IP 地址;

所述主用设备发送的 BFD 报文的源 IP 地址、目的 IP 地址和备用设备发送的 BFD 报文的源 IP 地址、目的 IP 地址相同。

8. 根据权利要求 6 所述的方法,其特征在于,

所述定期检查是否组网内的所有成员设备都接收到了其他成员设备发送的 BFD 报文的过程包括:根据所述备用设备发送的 BFD 报文中携带的成员 ID 检查是否组网内的所有成员设备都接收到了其他成员设备发送的 BFD 报文;

所述成员 ID 携带在 BFD 报文 My Discriminator 字段中预设的偏移值后面。

9. 根据权利要求 6 所述的方法,其特征在于,所述方法还包括:

在成员设备 A 没有接收到成员设备 B 发送的 BFD 报文,但成员设备 B 接收到成员设备 A 发送的 BFD 报文时,确认成员设备 A 到成员设备 B 的 BFD 链路单方向故障。

10. 根据权利要求 6 所述的方法,其特征在于,

所述接收所有备用设备转发的其接收到的所有 BFD 报文的过程包括:从主用设备和备用设备之间的内部通道接收所有备用设备转发的其接收到的所有 BFD 报文。

## 一种 BFD 链路检测装置和方法

### 技术领域

[0001] 本发明涉及通信技术领域,尤其涉及一种 BFD 链路检测装置和方法。

### 背景技术

[0002] 堆叠技术是一种软件虚拟化技术。它是将多台设备连接在一起,进行必要的配置以后,虚拟化成一台设备。使用这种虚拟化技术可以集合多台设备的硬件资源和软件处理能力,实现多台设备的协同工作、统一管理和不间断维护。

[0003] 以 IRF(Intelligent Resilient Framework,智能弹性架构)堆叠技术为例。在 IRF 组网环境中,所有单台设备称为成员设备,成员设备按照功能不同,分为主用设备(Active)和备用设备(Standby)两种角色。主用设备负责管理整个 IRF 组网中的成员设备,由角色选举产生,一个 IRF 组网中只能有一台主用设备。备用设备是隶属于主用设备的成员设备,作为主用设备的备份设备运行。在正常的情况下,IRF 组网中只有一台主用设备,其他设备均是备用设备。各个成员设备之间使用堆叠电缆连接形成 IRF 链路。如果成员设备之间的 IRF 链路发生故障,网络中就会分裂出两台或者两台以上的配置完全相同的主用设备,这就可能会引起网络故障,比如:IP 地址冲突、二层协议计算错误等。

### 发明内容

[0004] 有鉴于此,本发明提供一种 BFD 链路检测装置和方法,应用在堆叠组网中主用设备上。

[0005] 具体地,所述装置包括:

[0006] 第一会话单元,用于创建 BFD MAD 会话,并将所述 BFD MAD 会话备份到备用设备上,以供备用设备创建相同的 BFD MAD 会话;

[0007] 第一发送单元,用于通过 BFD 链路周期性地向堆叠组网中的所有备用设备发送 BFD 报文;

[0008] 第一接收单元,用于接收所有备用设备通过 BFD 链路向本机发送的 BFD 报文以及所有备用设备转发的其接收到的所有 BFD 报文;

[0009] 链路检测单元,用于定期检查是否组网内的所有成员设备都接收到了其他成员设备发送的 BFD 报文,在成员设备 A 没有接收到成员设备 B 发送的 BFD 报文时,确定成员设备 A 与成员设备 B 之间的 BFD 链路故障。

[0010] 所述方法包括:

[0011] 创建 BFD MAD 会话,并将所述 BFD MAD 会话备份到备用设备上,以供备用设备创建相同的 BFD MAD 会话;

[0012] 通过 BFD 链路周期性地向堆叠组网中的所有备用设备发送 BFD 报文;

[0013] 接收所有备用设备通过 BFD 链路向本机发送的 BFD 报文以及所有备用设备转发的其接收到的所有 BFD 报文;

[0014] 定期检查是否组网内的所有成员设备都接收到了其他成员设备发送的 BFD 报文,

在成员设备 A 没有接收到成员设备 B 发送的 BFD 报文时,确定成员设备 A 与成员设备 B 之间的 BFD 链路故障。

[0015] 由以上描述可以看出,本发明堆叠组网内各个成员设备之间发送相互发送 BFD 报文,然后主用设备通过各个成员设备对 BFD 报文的接收情况来检测 BFD 链路状况,从而能够及时地检测出 BFD 链路故障。

#### 附图说明

[0016] 图 1 是本发明一种实施方式 IRF 组网图;

[0017] 图 2 是本发明一种实施方式中 BFD 链路检测装置的逻辑结构图;

[0018] 图 3 是本发明一种实施方式中 BFD 链路检测方法流程示意图。

[0019] CZ1410017 具体实施方式

[0020] 针对堆叠链路故障引发网络故障的问题,通常使用 MAD (Multi-Active Detection,多 Active 检测)检测和处 理机制。具体地,仍以 IRF 技术为例,在 IRF 分裂的时候,MAD 可以检测出 IRF 网络中同时存在多个主用设备,并通过保留一个主用设备,禁用其他主用设备的方法,降低 IRF 分裂对业务产生的影响。其中,MAD 检测方式主要有 :BFD MAD 检测。BFD (Bidirectional Forwarding Detection,双向转发检测)是一个通用的、标准化的、介质无关和协议无关的快速故障检测机制。请参考图 1 所示的 IRF 组网环境,要 使能 BFD MAD 检测功能,需要预先在每一个成员设备上配置一个检测接口,将这些成员设备之间的检测接口连接起来,形成 BFD 链路。同时为每一个成员设备配置一个 MAD IP 地址,并将其划入同一个 VLAN (Virtual Local Area Network,虚拟局域网)。

[0021] 当 IRF 正常运行时,主用设备上的 MAD IP 地址生效,备用设备上的 MAD IP 地址不生效,只有主用设备在单方向地从 BFD 链路发送 BFD MAD 报文,无法建立 BFD 会话。而当 IRF 链路故障分裂后形成多个主用设备后,不同主用设备上配置的 MAD IP 地址均会生效,BFD 会话被激活,从而检测到多个主用设备冲突。在这个过程中,如果 BFD 链路发生故障,那么即便是 IRF 链路故障,也无法建立 BFD 会话。

[0022] 针对上述问题,本发明提供一种 BFD 链路检测装置和方法,通过各个成员设备周期性地发送 BFD 报文来检测 BFD 链路状态。如果成员设备 A 接收到了成员设备 B 发送的 BFD 报文,则说明成员设备 A 和成员设备 B 之间的 BFD 链路状态正常。如果成员设备 A 没有接收到成员设备 B 发送的 BFD 报文,则说明成员设备 A 和成员设备 B 之间的 BFD 链路状态异常。

[0023] 下面以软件实现为例,详细描述本发明具体实现。所述 BFD 链路检测装置,应用在 IRF 组网中主用设备上。作为本发明装置的运行载体,所述主用设备通常至少包括有 CPU、内存以及非易失性存储器,当然还可能包括各种转发芯片和 I/O 接口等硬件。请参考图 2,所述 BFD 链路检测装置包括有 :第一会话单元、第一发送单元、第一接收单元以及链路检测单元。请参考图 3,在一个示例性的实施方案中,该装置在运行过程中执行如下步骤 :

[0024] 步骤 101,第一会话单元创建 BFD MAD 会话,并将所述 BFD MAD 会话备份到备用设备上,以供备用设备创建相同的 BFD MAD 会话。

[0025] 在 IRF 组网中,要使能 BFD MAD 对 IRF 分裂的检测功能,通常需要管理人员事先在各个成员设备上为其配置 MAD IP 地址,所述 MAD IP 地址只有在本设备是主用设备的时候

才会生效。也就是说在主用设备上,BFD MAD 会话被激活,主用设备会周期性地向其他备用设备发送 BFD 报文,所述 BFD 报文的源 IP 地址是主用设备的 MAD IP 地址,目的 IP 地址是本机 BFD 链路对端的成员设备的 MAD IP 地址,也就是所述 BFD 报文接收成员设备的 MAD IP 地址。对于备用设备而言,由于 MAD IP 地址没有生效,所以其上没有 BFD MAD 会话,在接收到主用设备发送的 BFD 报文时,也不会响应。所以,在 IRF 未分裂时,主用设备和备用设备之间不会形成 BFD MAD 会话。

[0026] 本发明中,为了不影响对 IRF 分裂的检测功能。主用设备的第一会话创建单元在创建完毕 BFD MAD 会话后,将所述 BFD MAD 会话备份到各个备用设备上,以供备用设备创建和主用设备相同的 BFD MAD 会话。也就是说,备用设备上创建的用于检测 BFD 链路的 BFD 报文的源 IP 地址是主用设备的 MAD IP 地址,目的 IP 地址是本机自身的 MAD IP 地址。即,主用设备发送的 BFD 报文的源 IP 地址、目的 IP 地址和备用设备发送的 BFD 报文的源 IP 地址、目的 IP 地址相同。因此,本发明中用于检测 BFD 链路的 BFD MAD 会话不会和用于 IRF 链路检测的 BFD MAD 会话产生冲突。

[0027] 步骤 102,第一发送单元通过 BFD 链路周期性地向堆叠组网中的所有成员设备发送 BFD 报文。

[0028] 步骤 103,第一接收单元接收所有备用设备通过 BFD 链路向本机发送的 BFD 报文以及所有备用设备转发的其接收到的所有 BFD 报文。

[0029] 为实现本发明目的,通过步骤 101 在主用设备和备用设备上建立了 BFD MAD 会话。主用设备和备用设备都会通过 BFD 链路周期性地向组网内的其他成员设备发送 BFD 报文,同时也会接收其他成员设备发送给自己的 BFD 报文。其中,备用设备还需要将自己接收到的 BFD 报文转发给主用设备,以供主用设备进行汇总。具体地,所述备用设备通过检查接收到的报文的 UDP 目的端口号是否为 BFD 报文预设的端口号来识别 BFD 报文,然后通过主用设备和备用设备之间的内部通道将本机接收到的 BFD 报文发送给主用设备。以此来防止 BFD 链路故障而导致的报文转发失败。所述内部通道是堆叠组网中各个成员设备之间进行交互、备份的通道。比如:各个成员设备之间的堆叠链路。在成员设备是主控板的时候,所述内部通道也可以是背板。本发明对此不作特殊限制。

[0030] 为了确保 IRF 组网中的所有成员设备都能够收到其他成员设备发送的 BFD 报文,在本发明的一种优选的示例性实施方式中,所述 BFD 报文是组播报文,所述组播报文的目的 MAC 地址是组播 MAC。所述组播 MAC 可以采用现有的组播 MAC。优选地,为了确保组播 MAC 的可用性,所述组播报文的目的 MAC 采用预设的组播 MAC。所述 BFD 报文的发送周期可以由管理人员根据 IRF 组网的情况进行设置,比如,设置每秒钟发送一次 BFD 报文,本发明对此不作特殊限制。

[0031] 步骤 104,主用设备的链路检测单元定期检查是否组网内的所有设备都接收到了其他设备发送的 BFD 报文,并在设备 A 没有接收到设备 B 发送的 BFD 报文时,确定设备 A 与设备 B 之间的 BFD 链路故障。

[0032] 本步骤中,主用设备除了从 BFD 链路接收备用设备发送的 BFD 报文外,还会从堆叠链路接收备用设备转发的其接收到的 BFD 报文。主用设备的链路检测单元汇总组网内所有成员设备接收到 BFD 报文的情况,并定期进行检查。主用设备的链路检测单元可以通过维护一个全网成员设备 BFD 报文接收情况表,然后定期查表来检查是否组网内的所有设备都

接收到了其他设备发送的 BFD 报文。具体地,主用设备在接收到 BFD 报文后,先判断是从哪种链路上接收到的 BFD 报文。如果是从 BFD 链路上接收到所述 BFD 报文,则确认所述 BFD 报文是备用设备发送给自己的用于检测 BFD 链路的检测报文,进而根据 BFD 报文中携带的成员 ID 判断是哪一台备用设备发送给自己的,然后更新所述 BFD 报文接收情况表。请参考表 1,假定 IRF 组网中有三台成员设备,其成员 ID 分别为 1、2 和 3。表 1 中的行表示发送 BFD 报文的成员 ID,列表示接收 BFD 报文的成员 ID。成员设备 1 是主用设备,成员设备 2 和 3 是备用设备。Normal 表示 BFD 报文接收正常,Failure 表示 BFD 报文接收失败。如果主用设备 1 从 BFD 链路上接收到的 BFD 报文中携带的成员 ID 是 2,则说明这个 BFD 报文是成员设备 2 发送过来的,于是更新表 1 中第一行第二列状态为 Normal。

[0033] 如果主用设备是从堆叠链路上接收到 BFD 报文,则根据接收到所述 BFD 报文的堆叠端口可以得知是哪台备用设备转发给自己的,然后根据 BFD 报文中携带的成员 ID 得知是哪台成员设备发出的。举例来说,如果主用设备 1 从堆叠链路上接收到一个 BFD 报文,根据堆叠端口得知是备用设备 2 转发给自己的,然后根据成员 ID 得知是备用设备 3 发出的,就可以确定所述 BFD 报文是备用设备 3 发送给备用设备 2 的,进而更新表 1 中第二行第三列的状态是 Normal。

[0034] 链路检测单元的检查周期通常也是由管理人员配置,所述检测周期要大于成员设备发送 BFD 报文的周期,以成员设备发送 BFD 报文的周期是 1 秒钟为例,链路检测单元的检查周期可以设置为 5 秒。

[0035] 本发明中,所述成员 ID 可以添加在 BFD 报文中预设的字段里面。优选地,使用 BFD 报文的 My Discriminator 字段。主用设备发送的 BFD 报文中的 My Discriminator 字段是一个固定的值,为了和主用设备区分,备用设备可以在 My Discriminator 字段中填充特殊值来标识其成员 ID,比如:0xbf00000 (一个固定的偏移值)+备用设备的成员 ID。当然,本领域技术人员也可以采用其他的方式来标记成员 ID,本发明对此不作特殊限制。

[0036]

	1	2	3
--	---	---	---

[0037]

1	N/A	Normal	Failure
2	Normal	N/A	Normal
3	Failure	Failure	N/A

[0038] 表 1

[0039] 进一步地,本发明设置 BFD 链路状态的双向检测。如表 1 所示,成员设备 2 能够收到成员设备 3 发送的 BFD 报文,但是成员设备 3 无法收到成员设备 2 发送的 BFD 报文,此时,所述链路检测单元确定成员设备 2 到成员设备 3 的 BFD 链路方向故障,也就是成员设备 2 和成员设备 3 之间的 BFD 链路单方向故障。

[0040] 更进一步地,所述链路检测单元还用于在检查完毕组网内各个成员设备之间 BFD 链路的状态后,输出日志信息,并在有 BFD 链路故障时,上报提示信息,以便于用户能够及时地进行 BFD 链路故障的排查。

[0041] 由以上描述可以看出,本发明堆叠组网内各个成员设备之间发送相互发送 BFD 报文,然后主用设备通过各个成员设备对 BFD 报文的接收情况来检测 BFD 链路状况,从而能够及时地检测出 BFD 链路故障。

[0042] 以上所述仅为本发明的较佳实施例而已,并不用以限制本发明,凡在本发明的精神和原则之内,所做的任何修改、等同替换、改进等,均应包含在本发明保护的范围之内。



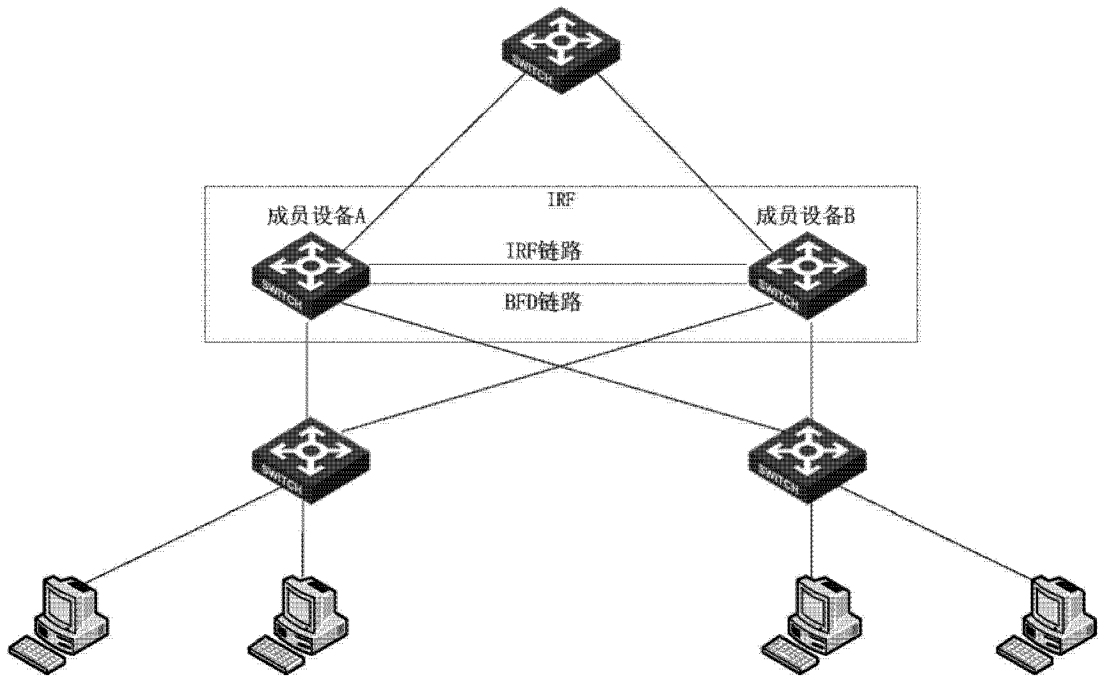


图 1

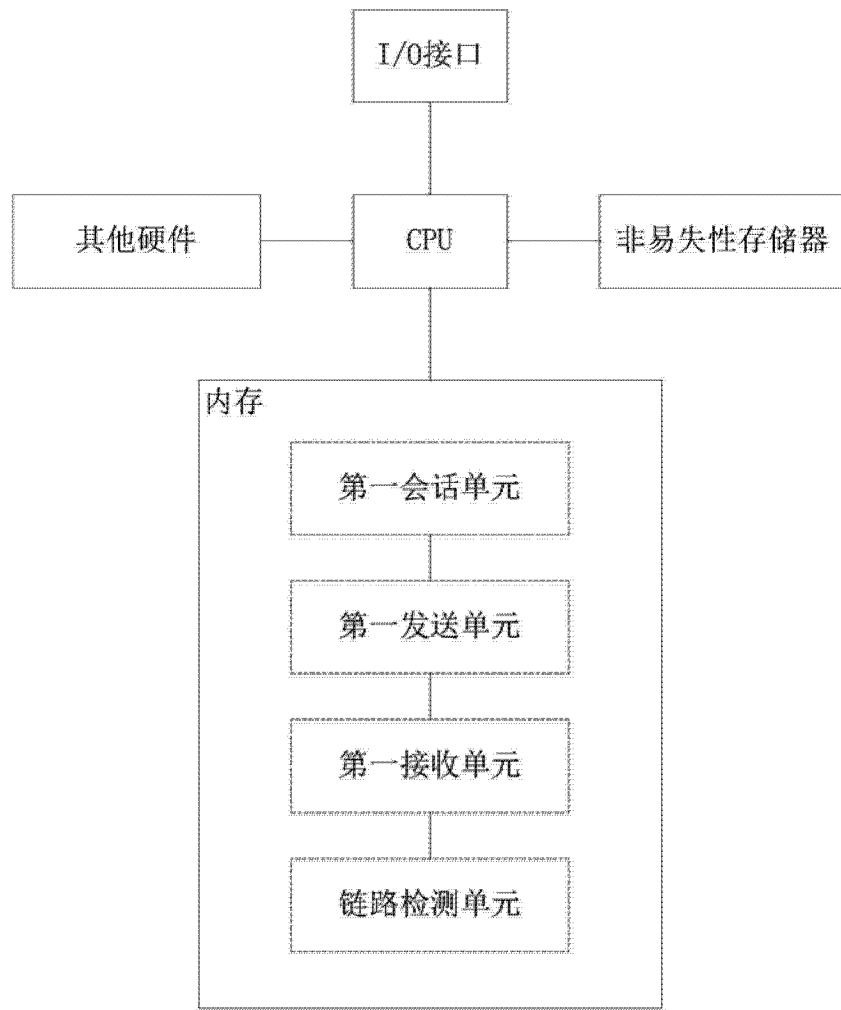


图 2

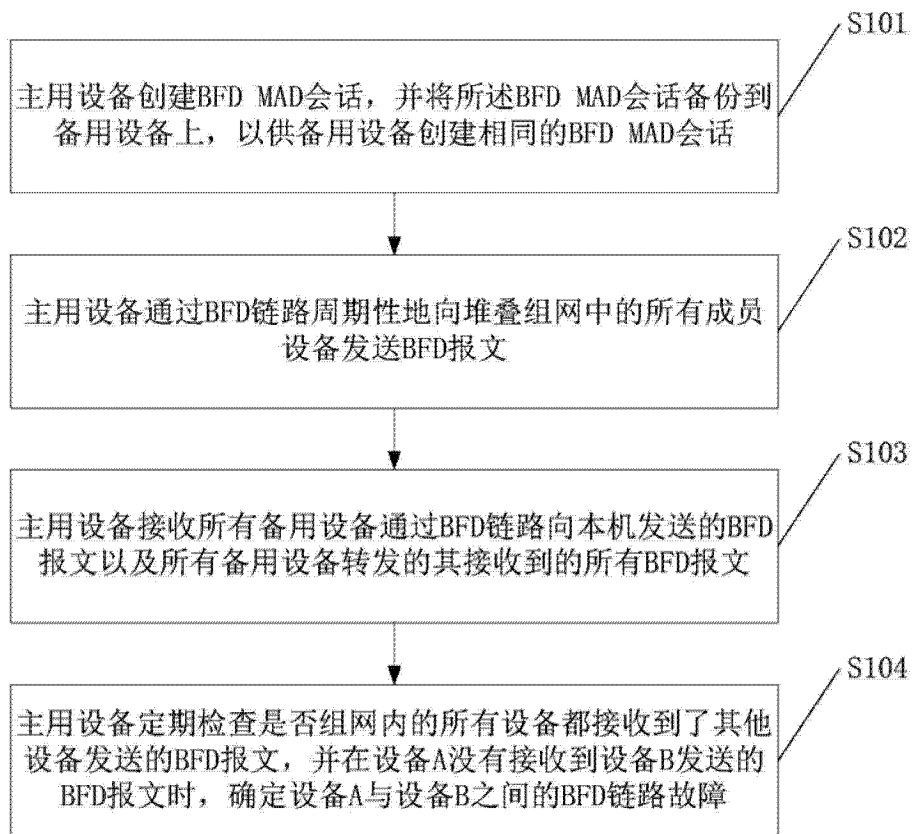


图 3