

[19] 中华人民共和国国家知识产权局

[51] Int. Cl.



# [12] 发明专利申请公布说明书

[21] 申请号 200580019242.4

H04J 15/00 (2006.01)

H04L 5/12 (2006.01)

G01R 31/08 (2006.01)

H04L 12/28 (2006.01)

[43] 公开日 2007年5月23日

[11] 公开号 CN 1969492A

[22] 申请日 2005.7.14

[21] 申请号 200580019242.4

[30] 优先权

[32] 2004.7.15 [33] US [31] 10/891,704

[86] 国际申请 PCT/US2005/025011 2005.7.14

[87] 国际公布 WO2006/019925 英 2006.2.23

[85] 进入国家阶段日期 2006.12.12

[71] 申请人 思科技术公司

地址 美国加利福尼亚州

[72] 发明人 让·菲利普·瓦瑟尔

斯特凡诺·B·普雷维蒂

戴维·D·沃德 罗伯特·戈盖恩

[74] 专利代理机构 北京东方亿思知识产权代理有限  
责任公司

代理人 王 怡

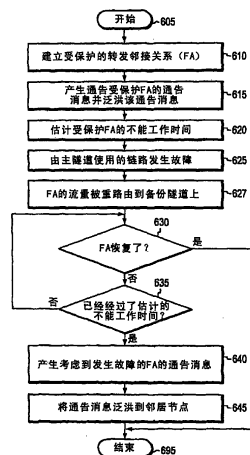
权利要求书 2 页 说明书 13 页 附图 6 页

[54] 发明名称

动态转发邻接关系

[57] 摘要

本发明的技术将受保护的转发邻接关系 (FA) 作为动态实体处理, 使得在 FA 的主隧道已发生故障时其允许与 FA 相关联的备份隧道运送该 FA 的流量一直到预定时间量。 如果已经过了预定时间量, 并且 FA 尚未被恢复 (例如主隧道尚未被重建), 则网络拓扑改变被自动触发, 导致网络在新网络拓扑上收敛。 通过触发网络拓扑改变, 随后可以确定出比与备份隧道相关联的路径更优的路径来运送流量。



1. 一种用于在数据网络中的中间节点处管理转发邻接关系 (FA) 的方法, 其中所述 FA 与主隧道和备份隧道相关联, 所述主隧道跨越所述数据网络中的一条或多条数据链路, 并且被配置为运送所述 FA 的流量, 所述备份隧道跨越所述数据网络中的一条或多条数据链路, 并且被配置为在所述主隧道发生故障时运送所述 FA 的流量, 该方法包括以下步骤:

确定所述备份隧道是否已运送所述 FA 的流量持续了预定时间段; 以及

如果是则触发网络拓扑改变。

2. 如权利要求 1 所述的方法, 还包括以下步骤:

检测所述主隧道跨越的数据链路的故障; 以及  
将去往所述主隧道的数据传送到所述备份隧道。

3. 如权利要求 1 所述的方法, 其中所述预定时间段是基于所述主隧道的故障历史的。

4. 如权利要求 1 所述的方法, 其中所述预定时间段是基于所述主隧道跨越的数据链路的故障历史的。

5. 如权利要求 1 所述的方法, 其中所述触发步骤包括:

在所述中间节点处产生通告消息, 其中所述通告消息指示所述 FA 已发生故障。

6. 如权利要求 6 所述的方法, 其中所述通告消息通过排除与所述 FA 相关联的链路信息元素来指示所述 FA 已发生故障。

7. 如权利要求 6 所述的方法, 其中所述通告消息通过将 FA 相关联的状态指示为有故障的来指示所述 FA 已发生故障。

8. 如权利要求 6 所述的方法, 其中所述通告消息是开放系统路径优先 (OSPF) 链路状态通告 (LSA) 消息。

9. 如权利要求 6 所述的方法, 其中所述通告消息是中间系统到中间系统 (IS-IS) 链路状态分组消息。

10. 如权利要求 6 所述的方法, 还包括以下步骤:

将所述通告消息泛洪到所述中间节点的邻居节点。

11. 一种数据网络中的中间节点，包括：

链路状态数据库，其包含代表转发邻接关系（FA）的条目，其中所述 FA 与主隧道和备份隧道相关联，所述主隧道配置为运送所述 FA 的流量，所述备份隧道配置为在所述主隧道发生故障时运送所述 FA 的流量；以及  
处理器，配置为确定所述备份隧道是否已运送所述 FA 的流量持续了预定时间段，并且如果是则触发网络拓扑改变。

12. 如权利要求 12 所述的中间节点，其中所述处理器被配置为检测与所述主隧道相关联的链路的故障，并且将所述 FA 的流量重路由到所述备份隧道上。

13. 如权利要求 12 所述的中间节点，其中所述处理器被配置为基于所述主隧道的故障历史确定所述预定时间段。

14. 如权利要求 12 所述的中间节点，其中所述处理器被配置为产生通告消息，其中所述通告消息指示所述 FA 已发生故障。

15. 如权利要求 15 所述的中间节点，其中所述通告消息通过排除与所述 FA 相关联的链路信息元素来指示所述 FA 已发生故障。

16. 如权利要求 15 所述的中间节点，其中所述通告消息通过将 FA 相关联的状态指示为有故障的来指示所述 FA 已发生故障。

17. 一种包含具有主隧道和备份隧道的转发邻接关系（FA）的装置，所述装置包括：

用于确定所述备份隧道是否已运送所述 FA 的流量持续了预定时间段的装置；以及

用于如果是则触发网络拓扑改变的装置。

18. 如权利要求 18 所述的装置，还包括：

用于产生通告消息的装置，其中所述通告消息指示所述主隧道已发生故障。

19. 一种计算机可读介质，包括用于以下过程的计算机可执行指令：

确定所述备份隧道是否已运送所述 FA 的流量持续了预定时间段；以及

如果是则触发网络拓扑改变。

## 动态转发邻接关系

### 技术领域

本发明涉及数据联网，更具体地说，涉及在数据网络中维护转发邻接关系。

### 背景技术

数据网络是地理上分布的由通信链路互连的节点和用于在末端站（例如，计算机）之间传送数据的网段的集合。存在从局域网（LAN）到广域网（WAN）的许多类型的网段可用。LAN 一般通过位于大体上同一物理位置（例如，建筑或园区）的专用私有的通信链路连接节点（例如，个人计算机和工作站）。LAN 还可以连接共同位于较近范围内的中间节点，例如路由器。

另一方面，WAN 一般通过长距离通信链路（例如，公共载波电话线路）连接大量在地理上分散的节点。因特网是连接整个世界的不同网络的 WAN 的一个示例，因特网在多个网络上的节点之间提供全球通信。节点一般通过根据预定义协议交换离散的数据帧或分组来在网络上通信，所述协议例如是传输控制协议/因特网协议（TCP/IP）。在此上下文中，协议包括定义节点如何彼此交互的一组规则。

某些中间节点例如路由器通常被配置来在网络中的多个节点之间“路由”数据，例如分组。路由选择一般在开放系统互连参考模型（OSI-RM）的网络层或者第三层（L3）中执行。路由器通常维护转发数据库（FDB），转发数据库一般被配置来存储路由选择信息，路由选择信息包括路由器用来确定数据（例如，数据分组）将被转发向何处以便到达它们的目的地 L3 地址和接口信息。例如，路由器可能具有包含一个或多个条目的路由选择数据库，其中每个条目包含目的地节点的 L3 目的地地址和关于路由器上通过其可到达该目的地节点的接口（例如，端口）的接口

信息。包含与路由选择表中的条目的目的地地址匹配的目的地地址的数据分组被路由器转发到匹配条目指定的接口，以传送到目的地节点。

另外，路由器可以维护链路状态数据库（LSDB），链路状态数据库通常被配置来保存关于数据网络中的数据链路的链路状态信息。在这里，路由器可以利用 LSDB 中的链路状态信息导出网络的拓扑。路由器还可以利用链路状态信息和该路由器接收到的通告消息（下面将描述）中包含的信息来确定网络的拓扑是否已改变。

路由器可以执行一种或多种路由选择协议，路由选择协议使得路由器能够路由分组并且与网络中的其他路由器交换路由选择和链路状态信息。路由器还可以使用这种信息配置（例如，计算）它们的 FDB 和 LSDB，等等。路由选择协议可能包括链路状态协议，例如，中间系统到中间系统（IS-IS）协议或开放最短路径优先（OSPF）协议。路由选择和链路状态信息一般在路由器之间以通告消息的形式被交换。例如，执行 IS-IS 协议的节点利用被称作链路状态分组的通告消息交换路由选择和链路状态信息。类似地，执行 OSPF 的节点利用被称作链路状态通告（LSA）的通告消息交换路由选择和链路状态信息。如这里所使用的，通告消息一般指中间节点用来向网络中的其他中间节点传递路由选择和/或链路状态信息的信息。

获取到通告消息的中间节点可以利用其中包含的信息来更新其 FDB 和 LSDB，并且维护网络拓扑。如果网络拓扑改变了（例如，链路故障），则探测到该改变的中间节点通常更新其 FDB 和 LSDB，并且产生说明该改变的一个或多个通告消息。该中间节点然后将通告消息泛洪到其邻居节点，从而触发“网络收敛”。即，邻居节点获取到通告消息，基于通告消息中包含的信息更新它们的 FDB 和 LSDB，然后将通告消息泛洪到它们的邻居节点，依此类推。该过程可以继续直到网络中的所有中间节点都已更新了它们的 FDB 和 LSDB，并且已在反映出改变后的网络拓扑的新的网络拓扑上收敛为止。

路由器可以通过源和目的地之间的网络利用面向连接的协议以“面向连接”的方式传送数据分组。面向连接的协议通过在源和目的地之间建立

的通常称作连接或电路的预定义路径将数据分组传送过网络。在这里，连接或电路在任何数据被传送之前被建立在源和目的地之间。在建立了连接后，数据通过由该连接定义的路径在源和目的地之间被传送。当不再需要该连接时，连接一般被“拆除”，并且使该连接使用的资源（例如节点、接口、协议等）对其他连接可用。这里所使用的资源指与中间节点相关联的实体。这些实体可以包括中间节点自身、中间节点上的接口（例如，端口）和在中间节点上运行的协议。面向连接的协议的一个示例是多协议标记交换（MPLS）协议。MPLS 提供了一个框架，该框架包含面向连接的链路层实现的各种特征，包括例如服务质量（QoS）、流量工程和基于约束的路由选择（CR）。

一些面向连接的协议利用单向连接，即，在从源到目的地的一个方向上传送数据的连接。例如，路由器 A 和路由器 B 之间的单向连接在从路由器 A 到路由器 B 的一个方向上传送数据。为了在另一个方向上传送数据，即，从路由器 B 到路由器 A，必须建立另一个从路由器 B 到路由器 A 的单向连接。连接可以利用信令协议从端到端被“用信令通知”，所述信令协议例如是资源保留协议（RSVP）。连接的发起该连接的信令的一端通常被称作该连接的“头端”，而该连接的终止信令的一端通常被称作该连接的“尾端”。容留连接的头端的路由器通常被称作头端节点，而容留连接的尾端的路由器通常被称作尾端节点。因此，例如，在路由器 A 容留连接的“头端”而路由器 B 容留连接的尾端的从源到目的地的连接中，路由器 A 是头端节点，而路由器 B 是尾端节点。

为了适应高可用性，某些面向连接的协议包括这样的技术，该技术利用备份隧道保护例如 MPLS 标签交换路径（MPLS LSP）中的各种链路不受故障影响。例如，P. Pan 等的“Fast Reroute Extensions to RSVP-TE for LSP Tunnels”（draft-ietf-mpls-rsvp-fastreroute-04.txt，在因特网工程任务组（IETF）<http://www.ietf.org> 处可获得）描述了 MPLS “快速重路由”（FRR）技术，该技术可以用来将一组流量工程（TE）MPLS LSP 绕过发生故障的“受保护”链路、共享风险链路群组（SRLG）或者节点重路由到预先用信令通知的“备份”隧道上。在一般的布置中，针对一条链路创

建两个 MPLS LSP 隧道（即，主隧道和备份隧道）来保护该链路。该链路正常情况下运送的流量在主隧道上被运送。如果主隧道发生故障，例如，由于主隧道所使用的物理链路发生故障，备份隧道就去往主隧道的数据流量。备份隧道充当了主隧道的“快速重路由”，并且避免了必须采用其他可能的昂贵的措施，例如，拆除主 MPLS TE LSP 并且建立绕过发生故障的网络元件的新的主 MPLS TE LSP。

MPLS TE LSP 可以在通告消息中被“声明”为“链路”，以便使 MPLS TE LSP 看起来是到网络中的其他中间节点的单条“物理链路”，即使其实际上可能跨越网络中的一条或多条物理链路也是如此。以这种方式通告的 MPLS TE LSP 通常被称作“转发邻接关系”（FA）。FA 在 K. Kompella 等的“LSP Hierarchy with Generalized MPLS TE”（draft-ietf-mpls-lsp-hierarchy-08.txt，从 IETF 可获得）中有所描述。与其他 MPLS TE LSP 一样，与 FA 相关联的 MPLS TE LSP 可以通过采用例如上述 MPLS FRR 方案之类的保护方案而被保护。在这里，主隧道被作为 FA 建立，其中主隧道被配置为在正常情况下运送流量，而备份隧道被配置为在主隧道发生故障时为 FA 运送流量。

受保护的 FA 的一个优点是从主隧道“切换”到备份隧道所花费的时间一般是数十毫秒数量级，并且对于网络来说通常是无缝的。该切换时间比中间节点必须在考虑到故障的新网络拓扑上收敛的情况下快得多。此外，由于切换是无缝的，所以网络拓扑改变未被触发，从而避免了网络需要在新网络拓扑上收敛，这种收敛可能消耗宝贵的时间并且可能导致网络停歇。换言之，由于主隧道被配置为 FA，并且利用诸如 FRR 之类的本地修复机制得到保护，所以对于网络中的其他节点网络拓扑保持不变。实际上，FA MPLS TE LSP 被本地重路由到备份隧道上，并且在通告消息中仍被声明为物理链路。

受保护 FA 的一个缺点是它们不是动态的。即，数据可能在发生故障的 FA 的备份隧道上被无限期地运送，直到例如该故障不再存在并且 FA 的主隧道被重新建立（恢复）为止。如果发生故障的 FA 的备份隧道沿着不是最优的路径，则网络效率和性能可能受到影响并持续不可接受的时间

段。

## 发明内容

本发明通过采用下述措施克服了与现有技术相关联的缺点：将受保护的转发邻接关系（FA）作为数据网络中的动态实体处理，使得在 FA 的主隧道已发生故障时其允许与 FA 相关联的备份隧道运送该 FA 的流量一直到预定时间量。如果已经过了预定时间量，并且 FA 尚未被恢复（例如，FA 的主隧道尚未被重建），则网络拓扑改变被自动触发，导致网络在新网络拓扑上收敛。通过触发网络拓扑改变，随后可以确定出比与备份隧道相关联的路径更优的路径来运送流量。

在所示出的本发明的实施例中，中间节点基于与 FA 相关联的主隧道不能工作的时间的历史，预先确定备份隧道可以运送发生故障的 FA 的流量的时间量。在与 FA 相关联的主隧道发生故障后，中间节点监控 FA 并且确定 FA 是否已恢复（例如，FA 的主隧道已被重建）。如果 FA 在预定时间量内尚未恢复，则中间节点通过产生将 FA 报告为发生故障的链路的通告消息并泛洪该通告消息从而触发网络拓扑改变。这又导致网络中的中间节点在新网络拓扑上收敛。

有利的是，通过在 FA 的备份隧道开始运送 FA 的流量一定时间后自动触发拓扑改变，本发明的技术使网络在可能包括比 FA 的备份隧道所使用的路径更优的路径的新拓扑上收敛。

## 附图说明

结合附图参考下面的描述，可以更好地理解本发明的上述和其他优点，在附图中类似的标号表示相同或者功能相似的元件：

图 1 是益于与本发明一起使用的数据网络的高层示意框图；

图 2 是益于与本发明一起使用的中间节点的高层示意框图；

图 3 是可以与本发明一起使用的监管程序引擎的部分示意框图；

图 4 是益于与本发明一起使用的线路卡的部分示意框图；

图 5 是益于与本发明一起使用的中间系统到开放系统最短路径优先



(OSPF) 协议链路状态通告 (LSA) 的部分示意框图; 以及

图 6 是根据本发明技术的可以用于配置中间节点来维护受保护的转发邻接关系 (FA) 的步骤序列的流程图。

### 具体实施方式

图 1 是益于与本发明一起使用的数据网络 100 的示意框图。数据网络 100 包括连接到多个网络实体 (例如, 端节点 108 和中间节点 200) 来形成计算机节点的互连网络的通信 (数据) 链路 104 的集合。这些网络互连的节点通过根据预定协议集合交换数据分组来进行通信, 所述协议例如是传输控制协议/因特网协议 (TCP/IP)。这里所使用的协议是描述如何在数据网络中的两个实体中间传送数据的形式规则。

图 2 是中间节点 200 的高层部分示意框图, 所示出的中间节点 200 是路由器。可以与本发明一起使用的合适的中间节点包括可从 Cisco Systems Incorporated, San Jose, CA 获得的 Cisco 7200、7600 和 12000 系列路由器。中间节点 200 包括通过底板 220 互连的一个或多个线路卡 400 和一个监管程序引擎卡 300。节点 200 被配置为执行各种传统的第 2 层 (L2) 和第 3 层 (L3) 交换和路由选择功能, 包括根据本发明的技术维护转发邻接关系 (FA)。这里所使用的 L2 和 L3 分别指开放系统互连参考模型 (OSI-RM) 的数据链路层和网络层。节点 200 还被配置为支持各种协议, 所述协议可能包括开放最短路径优先 (OSPF)、中间系统到中间系统 (IS-IS)、多协议标签交换 (MPLS)、TCP/IP、IP 快速重路由 (FRR)、MPLS FRR、以太网、异步传输模式 (ATM) 和帧中继 (FR)。

底板 220 包括点到点互连总线, 该总线互连各个板卡并且允许数据和信号从一个板卡被传送到另一个板卡。线路卡 400 将中间节点 200 与网络 100 连接 (用接口连接)。线路卡 400 利用诸如 ATM 和以太网之类的各种协议经由端口 215 向网络传送数据分组和从网络获取数据分组。在功能上, 线路卡 400 经由端口 215 获取来自网络 100 的数据分组并且将数据分组转发到数据总线 220, 以及将从数据总线 220 接收到的数据分组经由端口 215 发送到网络 100。端口 215 可以包括例如 ATM、以太网、快速以太

网（FE）、G 比特以太网（GE）和 FR 端口。

监管程序引擎 300 包括这样的逻辑装置，该逻辑装置除了其他以外还被配置为管理节点 200、维护其分配给线路卡 400 的集中转发数据库（FDB）、维护链路状态数据库（LSDB），并且执行各种协议，例如，OSPF、IS-IS、MPLS、MPLS FRR、IP 和 IP FRR。此外，引擎 300 还执行其他功能，包括结合了本发明的技术的多个方面的功能。图 3 是益于与本发明一起使用的监管程序引擎的高层部分示意框图。监管程序引擎 300 包括处理器 320、系统控制器 330、接口逻辑装置 360 和存储器 340。接口逻辑装置 360 被耦合到底板 220，并且被配置为在底板 220 和处理器 320 之间传送数据。

存储器 340 包括可由系统控制器 330 寻址的随机访问存储器（RAM）位置，用于存储例如数据结构和软件程序。具体而言，存储器 340 是计算机可读介质，包括配置来实现 128 兆字节（Mb）随机访问存储器的动态随机访问存储器（DRAM）器件。存储器 340 包含由处理器 320 使用的各种软件和数据结构，包括实现本发明的软件和数据结构。本领域技术人员知道其他计算机可读介质（例如，盘存储设备和闪存存储器器件）也可以被用来存储实现本发明多个方面的计算机可执行指令。此外，本领域技术人员知道还可以产生电磁信号来通过例如无线数据链路或者诸如因特网之类数据网络运送实现本发明多个方面的计算机可执行指令。

存储器 340 包含操作系统 342、LSDB 344、FDB 346 和路由选择过程 348。LSDB 344 保存涉及网络中的链路（例如，物理数据链路和 FA）的信息，这些信息可以用来导出网络 100 的拓扑。FDB 344 包含传统的转发信息，例如，网络中的节点的 L2 和 L3 地址，以及标识通过其可以到达与 FDB 344 中包含的地址相关联的节点的接口（例如，端口 215）的接口标识符（ID）。操作系统 342 包含计算机可执行指令，这些指令通过例如调用在监管程序引擎 300 上执行的软件过程支持的操作来在功能上组织中间节点 200。这些过程包括路由选择过程 348，该过程被配置为实现由中间节点 200 支持的各种路由选择和交换协议，以及本发明的各个方面。

系统控制器 330 耦合到处理器 320 和存储器 340，并且包括配置为使

处理器 320 能够访问（例如，读、写）存储器 340 中包含的存储器位置的电路。处理器 320 是传统的中央处理单元（CPU），其被配置为执行存储器 340 中包含的指令来除了其他以外还维护 LSDB 344 和 FDB 346。具体而言，处理器 320 执行用于获取关于与网络 100 中包含的各个中间节点 200 相关联的链路和路由的信息的指令，并且利用这种信息来维护 LSDB 344 和 FDB 346。此外，处理器 320 执行指令来产生包含对中间节点 200 已知的链路和路由信息的通告消息，并且将这些通告消息分发给网络中可以处理这种信息以相应地维护它们的 LSDB 和 FDB 的其他中间节点 200。

数据（分组）经由线路卡 400 被传送给网络 100 和从网络 100 传送出。图 4 是益于与本发明一起使用的示例性线路卡 400 的高层部分示意框图。线路卡 400 包括网络接口逻辑装置 420、编码地址识别逻辑装置（EARL）440、底板接口逻辑装置 460 和输出排队逻辑装置 450。此外，线路卡 400 还可以包含耦合到网络 100 的一个或多个端口 215。

网络接口逻辑装置 420 将线路卡 400 接口到网络 100，并且使得线路卡 400 能够经由端口 215 传送给往和来自网络 100 的数据。就此而言，逻辑装置 420 包括可以结合信号、电气和机械特性的传统接口电路，以及交换电路，这些电路是将线路卡 400 与网络的物理介质和在该介质上运行的协议相接口所需的。

底板接口逻辑装置 460 包含这样的电路，该电路用于将线路卡 400 接口到底板 220，并且使得线路卡 400 能够传送数据到耦合到底板 220 的其他板卡和获取来自所述其他板卡的数据。输出排队逻辑装置 450 包含配置为对经由端口 215 将数据分组传送到网络 100 上进行控制的电路，例如，输出队列和调度控制逻辑电路。EARL 440 说明性地被实现在专用集成电路（ASIC）中，该专用集成电路包括被配置为除了其他以外还获取和处理数据分组的电路，这包括利用例如 EARL 440 中包含的线路卡转发数据库（LCFDB）442 执行分组的转发判决。LCFDB 442 包含诸如目的地地址和端口之类的信息，这种信息使 EARL 440 能够确定由该 EARL 440 处理的分组的目的地。

在工作中，数据分组经由端口 215 被网络接口 420 从网络 100 获取，

并且被传送到其中分组被处理的 EARL 440。这种处理可以包括利用 LCFDB 442 来确定每个分组的目的地，例如，耦合到底板 220 的另一个板卡或者线路卡 400 上的端口 215。在确定了分组的目的地后，如果目的地是另一个板卡，则 EARL 440 引导底板接口 460 经由底板 220 将该分组传送到该目的地，如果目的地是线路卡 400 上的端口 215 的话，则传送到输出排队逻辑装置 450。去往监管程序引擎 300 的数据分组被接口逻辑装置 360 从底板 220 获取，并且被置于分组缓冲区 350 中，分组被保存在分组缓冲区 350 中以由处理器 320 进一步处理。

作为示例，中间节点 200 被配置为执行 OSPF 协议，并且利用被称作链路状态通告（LSA）的通告消息周期性地交换（通告）链路状态信息。可以用来配置中间节点 200 的一个 OSPF 版本在 J. Moy 的请求注释（RFC）2328 “OSPF Version 2” 中有所描述，该文献可从因特网工程任务组（IETF）<http://www.ietf.org> 处获得，并且通过引用整体结合进来，如同在此全面阐述了一样。应当理解，其他公知的协议，例如，IS-IS 协议，也可以利用本发明的技术。

LSA 是描述中间节点的本地状态的通告消息，所述本地状态包括例如中间节点的接口、物理数据链路和 FA 的链路状态。链路状态被泛洪到与该中间节点相关联的整个路由选择域，从而形成例如中间节点的 LSDB 344 中包含的信息的基础。图 5 是益于与本发明一起使用的 LSA 500 的示意框图。LSA 500 包含头部字段 510、链路信息头部 530 和一个或多个链路信息字段 550。头部字段 510 包含与 LSA 相关联的各种信息，这些信息包括 LSA 的“年龄”、各种选项、链路状态标识符（ID）、通告路由器的标识、LSA 的序号、在该 LSA 中出现的链路信息字段的数目（“链路数目”）等等。链路信息头部 530 包括路由器类型字段和链路数目字段，路由器类型字段包含指示发布该 LSA 500 的中间节点的类型（例如，边沿路由器、边界路由器）的值，链路数目字段保存指示该 LSA 500 中包含的链路信息字段的数目的值。

每个链路信息字段 550 提供关于由中间节点 200 通告的链路的各种信息，具体而言，字段 550 包含链路 ID、链路数据、链路类型、服务类型

(TOS) 度量数、一个度量和零或多个 TOS 度量 (未示出)。链路 ID 是标识连接到中间节点的链路的值。链路类型是标识链路的类型的值 (例如, 到另一个中间节点的点到点连接)。链路数据是表示取决于链路类型值的多种信息的值。TOS 度量数是表示链路信息字段中包含的 TOS 度量的数目的值。如果在 LSA 500 中没有 TOS 字段, 则该值为零。一般来说, TOS 度量 (未示出) 遵从度量, 并且可以包括表示与链路相关联的各种 TOS 信息的值。度量是表示与链路相关联的代价的值。

本发明涉及通过将 FA 作为动态实体处理来维护 FA, 因为本发明的技术允许与 FA 相关联的备份隧道被用来运送该 FA 的流量直至预定时间量。此后, 如果 FA 未恢复, 则网络拓扑改变被自动触发, 从而导致网络在新的网络拓扑上收敛。通过触发网络拓扑改变, 随后可以确定出比与备份隧道相关联的路径更优的路径来运送流量。

图 6 是根据本发明的技术可以用来配置中间节点 200 以维护受保护的 FA 的步骤序列的流程图。该序列开始于步骤 605, 然后前进到步骤 610, 在该步骤中受保护的 FA 被建立。作为示例, 受保护的 FA 是根据在 K. Kompella 等的 “LSP Hierarchy with Generalized MPLS TE” (draft-ietf-mpls-lsp-hierarchy-08.txt)、P. Pan 等的 “Fast Reroute Extension to RSVP-TE for LSP Tunnels” (draft-ietf-mpls-rsvp-fastreroute-04.txt) 中描述的技术建立的, 这两个文献都能够从 IETF 获得, 并且通过引用整体被结合进来, 如同在此全面阐述了一样。假设中间节点 200c 建立了到中间节点 200d 的受保护 FA, 该受保护 FA 的主隧道 106a (图 1) 从节点 200c 到节点 200d 跨越链路 104d, 并且其备份隧道 106b 沿下述路径, 该路径顺序包括节点 200c、200b、200a、200e、200f、200g 和 200d, 跨越链路 104c、104b、104f、104g、104h 和 104i。另外, 假设与每条链路 104 相关联的代价相同。

在步骤 615 中, 中间节点 200c 产生通告消息, 并且将该通告消息泛洪到其邻居节点 200b 和 200d, 以通告该受保护的 FA。作为示例, 中间节点 200c 产生包含链路信息字段 550 的 LSA 500, 并且将 LSA 500 泛洪到其邻居节点 200b 和 200d, 其中链路信息字段 550 标识该受保护的 FA。

在步骤 620 中，中间节点 200c 估计受保护 FA 的“不能工作时间”。作为示例，中间节点 200c 基于故障历史来估计该不能工作时间，故障历史包括例如被受保护 FA 的主隧道 106a 使用的链路 104d 的故障持续期间。具体而言，节点 200c 确定对于链路 104d 上发生故障的某一百分比（例如，90%），故障的持续期间是例如“X”毫秒（ms）。中间节点 200c 然后将 X 乘以一个因子（例如，1.2）来计算出链路 104d 若发生故障的预期不能工作时间。例如，如果在链路 104d 上发生故障的百分比为 90%，链路处于“不能工作”状态（即，不可用）持续 20ms，并且因子为 1.2，则链路 104d 若发生故障的预期不能工作时间为  $1.2 \times 20\text{ms}$ ，或者说 24ms。

应当注意到，用于确定估计的不能工作时间的上述技术仅是要说明可以估计 FA 的不能工作时间的一种方法，而不是要限制本发明。用于估计 FA 的不能工作时间的其他技术也益于与本发明的技术一起使用。例如，对主隧道跨越的链路和/或由主隧道使用的节点的历史不能工作时间进行平均的技术，或者采用一些其他统计方法来估计链路的不能工作时间的技术也可以与本发明一起使用。

还应当注意，为了说明目的，主隧道 106a 是单跳隧道。然而，这并不是要限制本发明。相反，本发明的技术可以与可以跨越（利用）多于一条链路和/或多于一个节点的多跳隧道一起使用。在这里，用于确定 FA 的主隧道的估计的不能工作时间的技术可以基于由受保护 FA 的主隧道使用的物理链路和节点的某些组合的估计的不能工作时间。

在步骤 625 中，受保护 FA 的主隧道 106a 发生故障（例如，链路 104d 发生故障），并且该受保护 FA 的流量被重路由到其备份隧道 106b 上（步骤 627）。注意，数据流量现在沿备份隧道 106b 上的路由到中间节点 200d，该路由与主隧道 106a 被用来运送流量时其所沿的路由相比不是最优的。即，流量现在在备份隧道 106b 上顺序经由节点 200b、200a、200e、200f 和 200g 并且经由链路 104c、104b、104f、104g、104h 和 104i 从节点 200c 到达节点 200d，而不是在发生故障前经由链路 104d 从节点 200c 到节点 200d。

在步骤 630 中，中间节点 200c 执行检查来确定 FA 是否已恢复（例

如，链路 104d 不再处于故障中，并且 FA 的主隧道 106a 已通过链路 104d 被重建）。如果已恢复，则该序列前进到步骤 695，在该步骤中该序列结束。否则，该序列前进到步骤 635，在该步骤中，执行检查来确定发生故障的主隧道 106a 是否已持续一段时间“不能工作”（即，不可用），其中该时间段与隧道 106a 的估计的不能工作时间匹配（即，从隧道 106a 发生故障起已经过的时间与隧道 106a 的估计的不能工作时间匹配）。如果不是，则序列返回到步骤 630。

否则，该序列前进到步骤 640，在该步骤中中间节点 200c 产生指示 FA 已发生故障的通告消息。作为示例，中间节点 200c 产生其中省略了与受保护 FA 相关联的链路信息元素 550 的 LSA 500。或者，中间节点 200c 可以产生将与 FA 相关联的状态指示为不可用或者发生故障的通告消息。

在步骤 645 中，中间节点 200c 将通告消息 500 泛洪到其邻居节点 200b 和 200d。注意，泛洪（转发）该通告消息 500 触发了网络拓扑改变，并且导致中间节点 200 例如重新生成它们的 FDB 346 和 LSDB 344 来在考虑到发生故障的 FA 的新网络拓扑上收敛。在步骤 695 中，该序列结束。

应当注意，上述实施例将本发明的技术描述为与沿一个方向流动（例如，从节点 200c 到节点 200d）的 FA 流量一起使用。这并不是要限制本发明。相反，根据 FA 的操作，通常可以建立第二组主隧道和备份隧道来处理沿另一个方向流动（例如，从节点 200d 到节点 200c）的 FA 流量，本发明的技术也可以被采用来与这另一组隧道一起使用。

总而言之，本发明的技术是对现有技术的改进，这在于其通过允许与受保护 FA 相关联的备份隧道被使用持续多至预定时间量，从而将该受保护 FA 作为动态实体处理。如果已经过了该时间量并且 FA 尚未恢复（例如，与 FA 相关联的主隧道尚未被重建），则考虑到该发生故障的 FA 的通告消息被发布（泛洪），以触发网络拓扑改变。触发拓扑改变导致网络中的节点在可能包含比备份隧道更优的路由的新拓扑上收敛，从而提高了网络效率。

前面已针对本发明的特定实施例进行了描述。很清楚，在保留本发明

---

的一些或全部优点的情况下，可以对所述实施例作出其他改变和修改。因此，所附权利要求书的目的是覆盖所有这种改变和修改，只要它们落入本发明的真实精神和范围内。



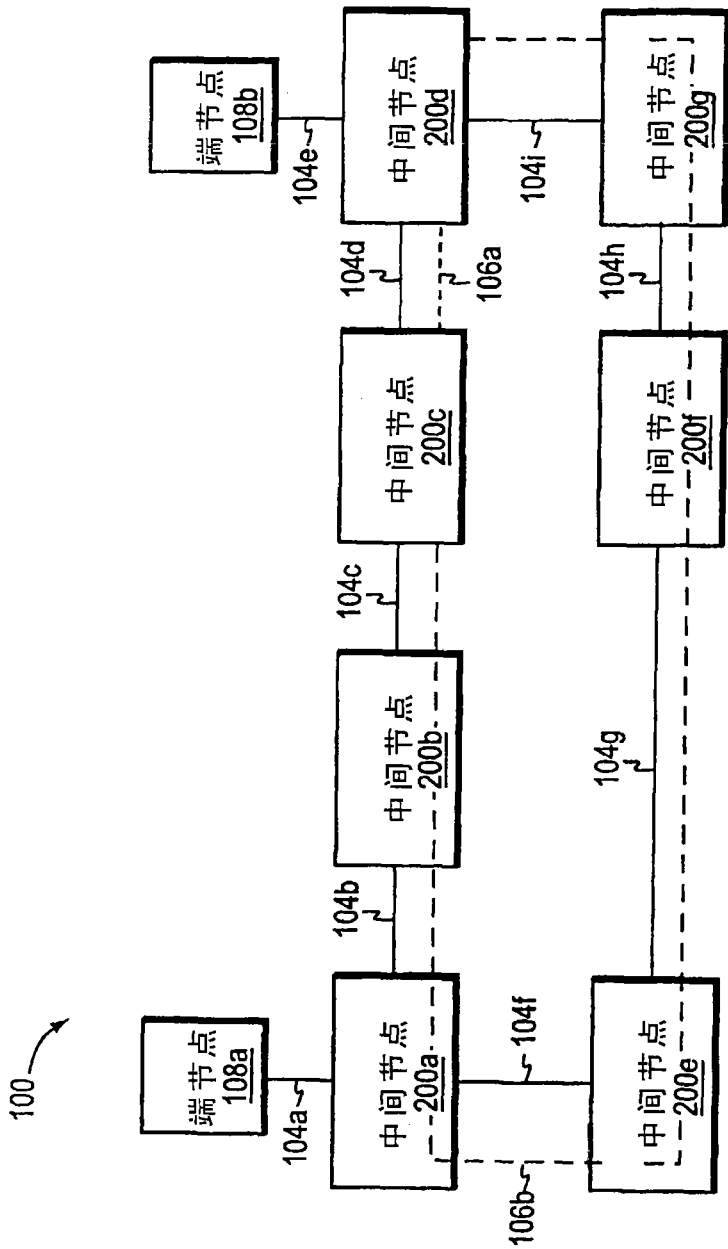


图1

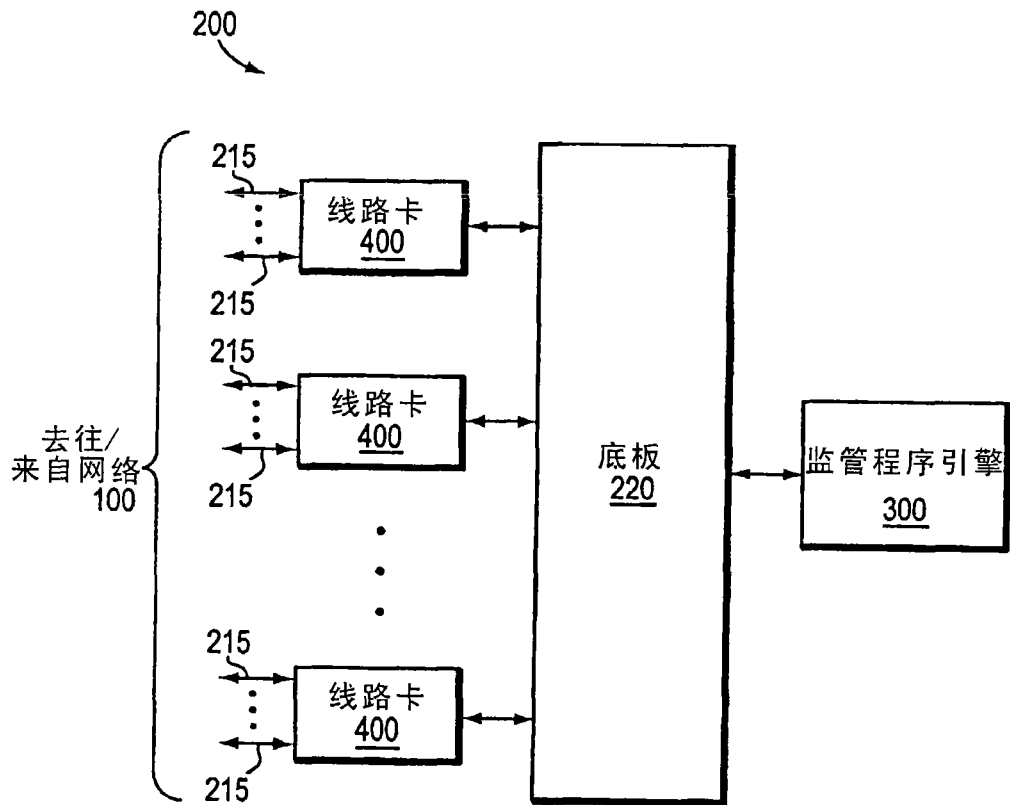


图2

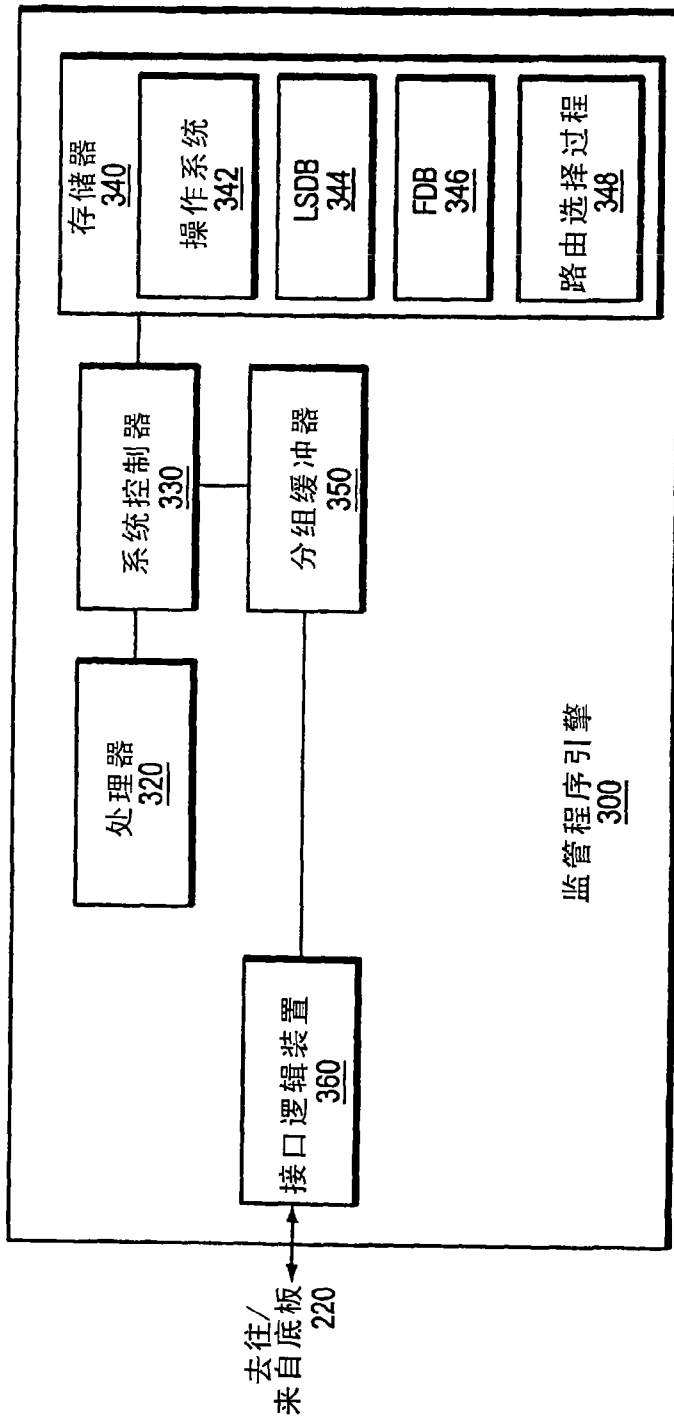


图3

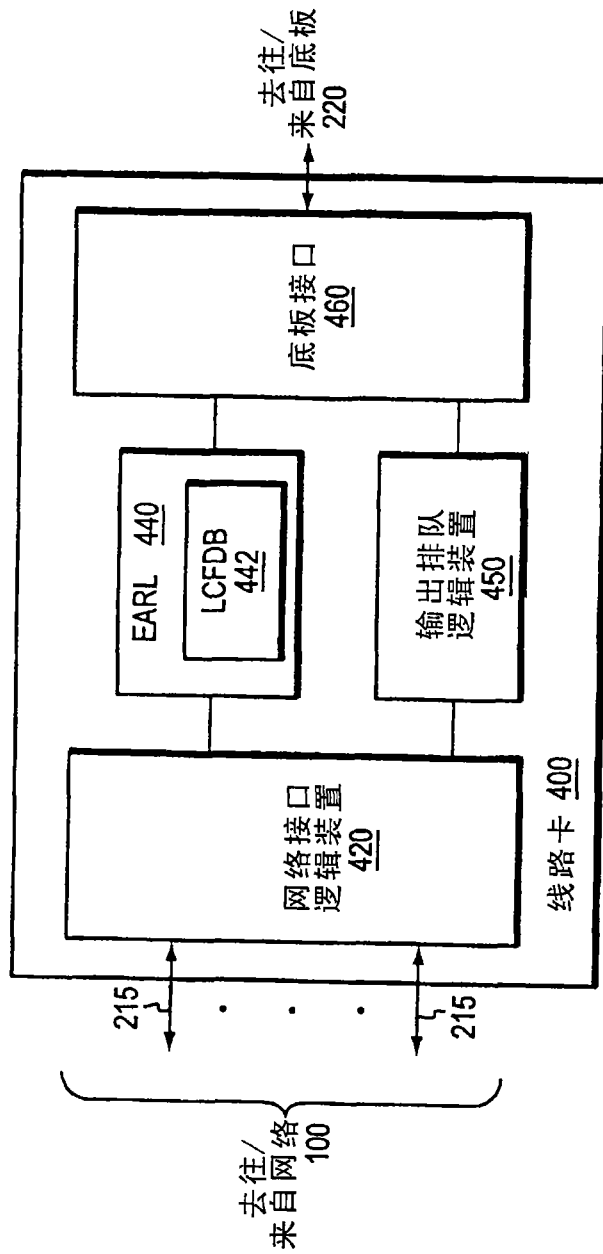


图4

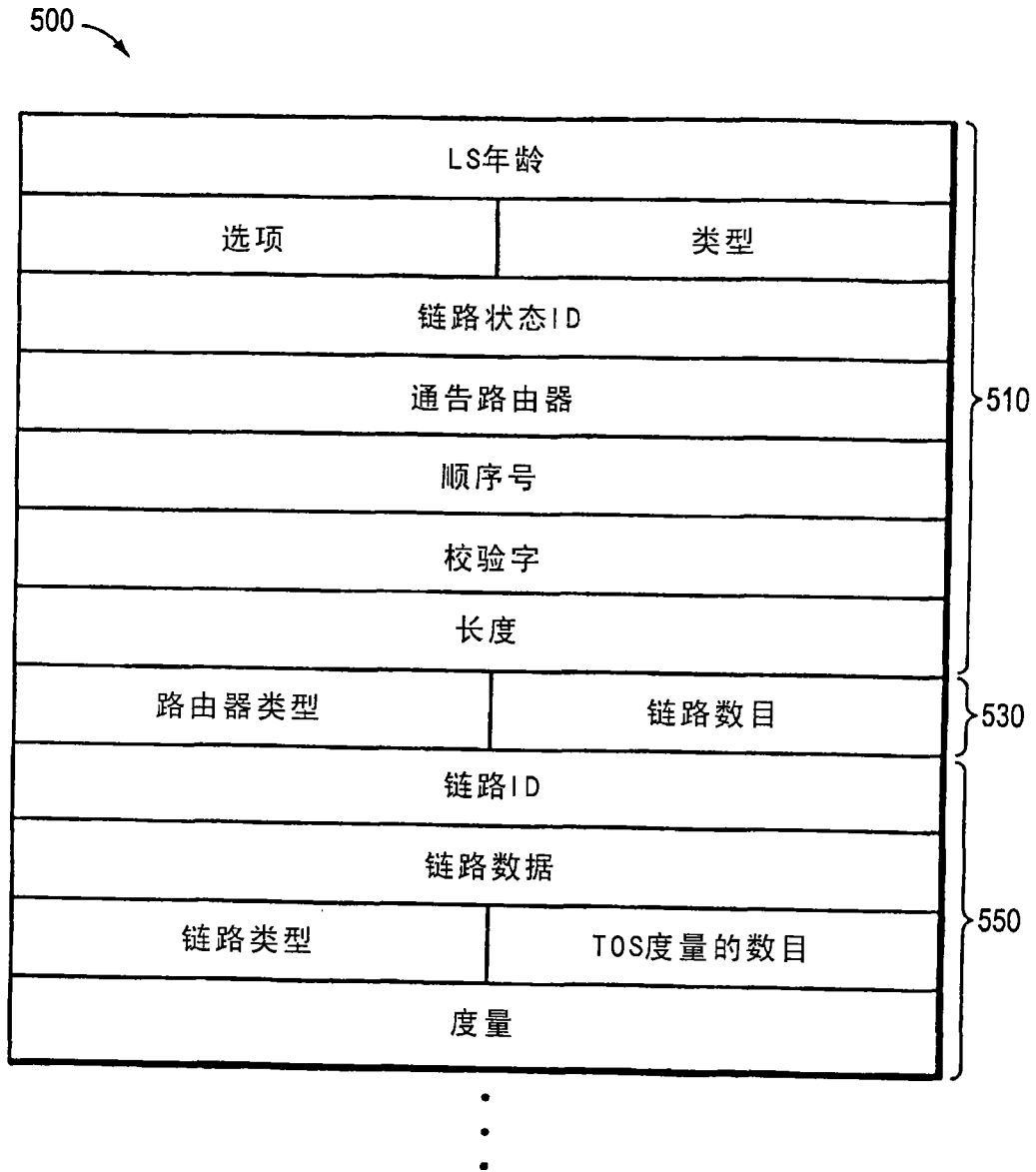


图5

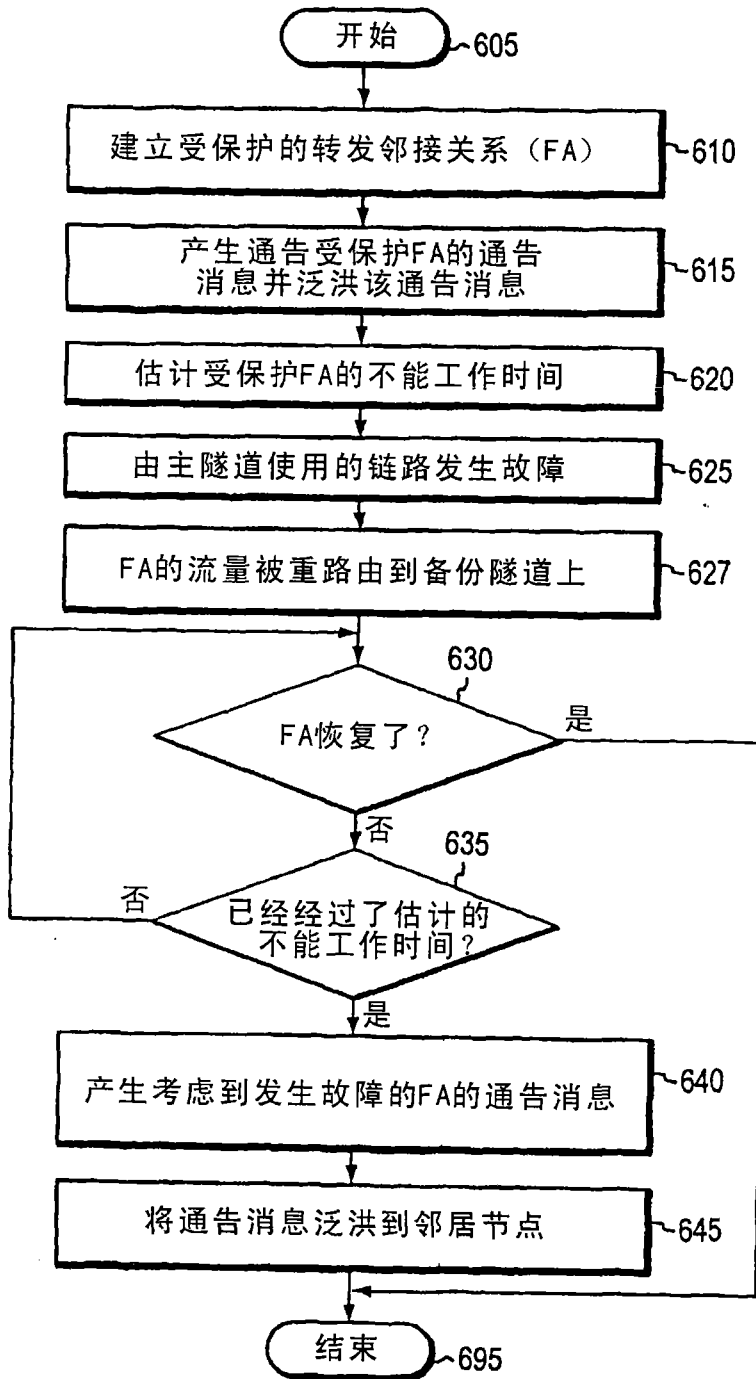


图6