



# [12] 发明专利申请公布说明书

[21] 申请号 200480003102.3

[43] 公开日 2007年10月10日

[11] 公开号 CN 101053231A

[22] 申请日 2004.1.22

[21] 申请号 200480003102.3

[30] 优先权

[32] 2003.1.30 [33] US [31] 60/443,573

[32] 2003.12.9 [33] US [31] 10/730,004

[86] 国际申请 PCT/IB2004/000141 2004.1.22

[87] 国际公布 WO2004/068261 英 2004.8.12

[85] 进入国家阶段日期 2005.7.29

[71] 申请人 诺基亚公司

地址 芬兰埃斯波

[72] 发明人 皮特里·伊拉-奥蒂宁

迈克尔·拉特瓦拉 劳里·拉蒂宁

海基·图纳宁 伊尔卡·维斯特曼

伯恩哈德·豪内森

[74] 专利代理机构 中国国际贸易促进委员会专利商  
标事务所  
代理人 董 莘

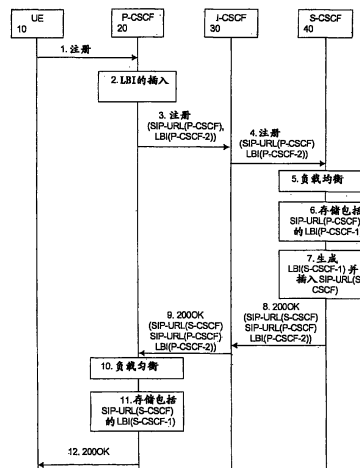
权利要求书 6 页 说明书 20 页 附图 5 页

## [54] 发明名称

负载控制信息的基于消息的传递

## [57] 摘要

本发明涉及一种用于控制分组数据网络中的处理负载的方法和系统，其中负载控制信息被设置在消息的预定字段中。然后，在所述消息的路由路径上检查所述负载控制信息，并响应于检查步骤的结果选择所述分组数据网络的处理资源。由此，可在所述网络处提供负载均衡信息，从而提供有效均衡和负载共享功能，而不会带来繁琐的字符串比较和数据库查询。



1、一种控制分组数据网络中的处理负载的方法，所述方法包括步骤：

- a) 在消息的预定字段（121 到 12n）中设定负载控制信息；
- b) 在所述分组数据网络中路由所述消息；
- c) 在所述消息的路由路径上检查所述控制信息；以及
- d) 响应于所述检查步骤的结果，选择所述分组数据网络的处理资源。

2、根据权利要求 1 的方法，其中所述预定字段是地址标题（100）的用户部分（120）的子字段（121 到 12n）。

3、根据权利要求 1 的方法，其中所述预定字段是 SIP 消息的通路分支。

4、根据权利要求 3 的方法，还包括将所述负载均衡信息从另一预定字段拷贝到所述预定字段的步骤。

5、根据权利要求 2 的方法，其中所述地址标题是 SIP 路由标题的 URI（100）。

6、根据权利要求 2 或 5 的方法，还包括步骤：在所述用户部分（120）中提供多个子字段（121 到 12n），以传递不同类型的所述负载控制信息。

7、根据权利要求 6 的方法，其中所述用户部分（120）被语法分析并分为所述子字段（121 到 12n）。

8、根据权利要求 6 的方法，其中所述子字段（121 到 12n）的结构、顺序和使用中的至少一个是预定的。

9、根据权利要求 6 到 8 中任何一个的方法，其中所述子字段（121 到 12n）由预定比特串、字符、字符串分隔。

10、根据权利要求 1 的方法，其中虚拟地址由多个处理器节点共享。

11、根据权利要求 10 的方法，其中所述处理器节点具有基于 IP 的蜂窝网络的呼叫状态控制功能。

12、根据权利要求 2 到 11 中任何一个的方法，其中所述负载控制信息包括第一端口号，所述第一端口号指示用于接收请求消息的第一端口。

13、根据权利要求 2 到 12 中任何一个的方法，其中所述负载控制信息包括第二端口号，所述第二端口号指示用于接收响应消息的第二端口。

14、根据权利要求 1 的方法，其中所述负载控制信息包括第一信息，所述第一信息指示所述消息的会话是否已经存在。

15、根据权利要求 14 的方法，其中所述负载控制信息包括第二信息，所述第二信息指示所述现有会话的识别。

16、根据权利要求 14 或 15 的方法，其中所述负载控制信息被存储在 SIP 消息的路由标题字段、通路标题字段或接触标题字段中。

17、根据权利要求 14 到 16 中任何一个的方法，其中所述负载控制信息是对于其它网络而言没有意义的隐藏信息。

18、根据权利要求 14 到 17 中任何一个的方法，其中所述负载控制信息被设定为标题字段的主机名称的一部分。

19、根据权利要求 14 到 17 中任何一个的方法，其中所述负载控制信息被设定为标题字段的参数。

20、根据权利要求 14 到 17 中任何一个的方法，其中所述负载控制信息被设定为标题字段的端口号。

21、根据权利要求 20 的方法，其中所述端口号用于区分第一消息与后续消息。

22、根据权利要求 14 到 17 中任何一个的方法，其中所述负载控制信息被设置为标题字段的扩展标题字段。

23、根据权利要求 14 到 17 中任何一个的方法，其中所述负载控制信息被设定在所述消息的有效负荷部分中。

24、根据权利要求 15 的方法，还包括步骤：响应于检测所述第一信息，提取所述第二信息，并将所述第二信息用于所述选择步骤。

25、一种分配分组交换网络中的负载控制信息的方法，所述方法包括步骤：

- a) 在第一网络单元 (20) 中生成第一负载控制信息；
- b) 将所述第一负载控制信息设定到消息的预定字段中；
- c) 将所述消息路由到第二网络单元 (40)；

- d) 将所述第一负载控制信息存储在所述第二网络单元中;
- e) 在所述第二网络单元中生成第二负载控制信息;
- f) 将所述第二负载控制信息设定到消息的预定字段中;
- g) 将所述消息路由到所述第一网络单元; 以及
- h) 将所述第二负载控制信息存储在所述第一网络单元中。

26、一种控制分组数据网络中的处理负载的网络设备, 所述设备(20、40)包括:

- a) 检查装置, 用于检查设置在消息的预定字段(121到12n)中的负载控制信息;
- b) 选择装置, 用于响应于所述检查装置, 选择用于所述消息的处理资源。

27、根据权利要求26的网络设备, 其中所述网络设备(20、40)包括基于IP的蜂窝网络的呼叫状态控制功能。

28、根据权利要求26或27的网络设备, 其中所述选择装置被设置为选择所述消息被分配给的预定处理器节点。

29、根据权利要求26或27的网络设备, 其中所述选择装置被设置为启动新会话的创建。

30、根据权利要求29的网络设备, 其中所述负载控制信息包括第一信息, 所述第一信息指示所述消息的会话是否已经存在。

31、根据权利要求30的网络设备, 其中所述负载控制信息包括第二信息, 所述第二信息用于识别所述现有会话。

32、一种用于将消息传送到分组数据网络的设备, 所述设备(10、

20、40) 被设置为将负载控制信息设定到所述消息的预定字段(121到12n)中,以选择所述分组数据网络的处理资源。

33、根据权利要求32的设备,其中所述设备包括基于IP的蜂窝网络的呼叫状态控制功能。

34、根据权利要求33的设备,其中所述呼叫状态控制功能是服务呼叫状态控制功能或代理呼叫状态控制功能。

35、根据权利要求32到34中任何一个的设备,其中所述设备被设置为在所述消息的标题地址(100)的用户部分(120)中,设定所述负载控制信息。

36、根据权利要求35的设备,其中所述标题地址是SIP URI(100)。

37、根据权利要求32到34中任何一个的设备,其中所述设备被设置为在主机名称、标题参数、端口号、所述消息的标题部分的扩展标题字段中或者在所述消息的有效负荷部分中,设定所述负载控制信息。

38、根据权利要求37的设备,其中所述负载控制信息包括第一信息,所述第一信息指示所述消息的会话是否已经存在。

39、根据权利要求38的设备,其中所述负载控制信息包括第二信息,所述第二信息指示所述现有会话。

40、一种控制分组数据网络中的处理负载的系统,所述系统包括:  
a) 第一网络单元(10、20、40),用于在将被在所述分组数据

网络中路由的消息的预定字段(121到12n)中设定负载控制信息;以及

b) 第二网络单元(20、40), 用于在所述消息的路由路径上检查所述负载控制信息; 以及响应于所述检查步骤的结果, 选择所述分组数据网络的处理资源。

41、一种分配分组交换网络中的负载控制信息的系统, 所述系统包括:

a) 第一网络单元(20), 用于生成第一负载控制信息, 并将所述第一负载控制信息设置到消息的预定字段中; 以及

b) 第二网络单元(40), 用于接收所述消息, 存储所述第一负载控制信息, 生成第二负载控制信息, 将所述第二负载控制信息设定到第二消息的预定字段中, 以及将所述第二负载控制信息路由到所述第一网络单元;

c) 其中所述第一网络单元(20)用于存储所述第二负载控制信息。

42、根据权利要求40或41的系统, 其中所述第一和第二网络设备包括呼叫状态控制功能。

## 负载控制信息的基于消息的传递

### 技术领域

本发明涉及一种用于控制分组数据网络内的处理负载的方法和系统，例如设置在分组数据网络之上的互联网协议（IP）多媒体子系统（IMS），以向第三代移动设备提供语音和多媒体业务。

### 背景技术

假定在未来几乎所有固定和移动通信网络都将基于互联网技术。尤其，组合若干通信类型和模式的业务将会主宰未来的网络。语音本身尽管很重要，但将仅成为整个通信体系结构内的一部分。

在互联网工程任务组（IETF）技术规范 RFC 3261 内定义的会话启动协议（SIP）提供了一种用于在互联网上建立多媒体会话的新兴标准。其基本性能在于任何通信会话的建立修改和结束，因此其是一种信令协议。SIP 还提供了个人移动性，这意味着消费者可经由单个地址通达，而不论其当前对于所述网络的附加点在哪里。

为了支持多媒体业务、无缝移动和有效多方会议，必须增强 IP 层。移动 IP 允许终端在不同移动网络之间自由移动。

SIP 用于建立、修改和终止会话。其通过允许用户动态向所述网络注册他的通信地址，即 SIP URI（统一资源指示符），提供个人移动。会话通常是若干将被交换的实时传送协议（RTP）流。通常，会话是语音、音频和视频流的组合，但其还可能包括共享应用。基本 SIP 网络包括四种类型的单元，即用户代理（UA）、代理服务器、转发服务器和注册服务器。用户代理通常位于诸如 IP 电话、个人计算机或移动设备的端点内。所述用户代理启动请求并提供响应。通常，UA 还具有对于媒体处理的接口，以及对于提供所述用户接口的实际应用程序的接口。代理服务器是中介物，其接收并转发例如向其提供路由或

其它业务的请求。改发服务器仅通过请求其始发者将请求改发至新地址来响应请求。注册服务器通过从所述 UA 接受注册，跟踪所述消费者的实际接触点。注册服务器和 SIP 注册进程通常提供所述消费者能够经由单个地址从任何位置通达时的用户移动。在这种情况下，它们集合全球移动通信系统 (GSM) 网络内的原籍位置注册 (HLR) 功能。每个消费者都是域的一部分，每个域都运行至少一个注册服务器，所述注册服务器在其消费者得以注册的情况下了解所述用户的位置。

SIP 使用互联网邮件通用的地址格式，即“用户@域”。域部分用于找到消费者的正确域，而用户部分用于分辨域内的各个用户。SIP 包括请求和响应消息，所述请求和响应消息包括标题字段、接收者地址、发送者地址等，所述标题字段例如用于定义随后将请求发送到何处。此外，SIP 消息可能包括用于传送特定于用户或业务的信息的有效负荷部分。

应当注意的是，RTP 与 SIP 消息流并不流经相同路径，而是在相关设备之间直接流动。因此，可在所述 UA 之间直接发送后续 SIP 请求。在 IMS 内，后续 SIP 消息流经记录在初始请求的记录-路由标题内的路径，而轮询网络节点可能退出，而其它网络单元停留在路径上。另一方面，中间的代理服务器可能要求在呼叫期间内保持在信令路径上。如果代理为所述呼叫提供一些业务，则这可能是有用的。

如今，第三代合作项目 (3GPP) 例如将其技术规范 TS 23.228 内的 IMS 规定为可结合不同网络使用的接入独立子系统。IMS 将 SIP 用于会话启动。基本上，IMS 仅是 SIP 网络的一个实例。其具有若干代理和注册员。所述 UA 位于终端设备或用户设备 (UE) 内。当两个设备建立会话时，它们经由呼叫状态控制功能或呼叫会话控制功能 (CSCF) 单元彼此通话，而 RTP 媒体流并不经过 CSCF，而是直接在所述设备之间流动。应用服务器 (AS) 是处理业务的 SIP 单元，所述业务例如是增强型呼叫控制、出现或瞬间消息发送。所述 AS 可能终止会话/事务。所述 AS 可能还代表用户或业务启动会话/事务。

然而，可能存在这样的情况，AS 在接收诸如 SIP INVITE 消息

时，并不了解其应当开始始发还是终止业务，或服务 CSCF (S-CSCF) 并不了解呼入请求消息开始始发还是终止会话/事务。此外，可能需要其它信息来平衡所述网络内的负载。此外，出于连接处理服务器(CPS)内，尤其是 S-CSCF 和轮询 CSCF (I-CSCF) 内负载分担的目的，提供快速简单的算法以发现所接收 SIP 请求是否为 SIP 会话内的第一个，或所接收请求或响应属于哪个 SIP 会话很重要。如今，SIP 并不提供这种有效方法。为了识别呼叫识别、源和目的地所识别的 SIP 对话，即呼叫路线，网络单元或 UE 必须比较每个 SIP 消息的对应标题字段，并确定所述呼叫路线是否已经存在。这意指复杂的字符串比较和数据库询问。保持大量的并行呼叫路线的网络单元需要许多资源。此外，万一网络单元内出现故障，需要关于现有会话的信息。

#### 发明内容

因此，本发明的目的是为分组数据网络提供一种有效的负载控制方案。

此目的是通过一种控制分组数据网络中的处理负载的方法实现的，所述方法包括步骤：

在消息的预定字段中设定负载控制信息；

通过所述分组数据网络路由所述消息；

在所述消息的路由路径上，和/或在对于所述消息的响应消息的路由路径上，检查所述控制信息；以及

响应于所述检查步骤的结果，选择所述数据分组网络的处理资源。

此外，此目的是通过一种用于控制分组数据网络内的处理负载的网络设备实现的，所述设备包括：

检查装置，用于检查设置在消息或响应消息的预定字段内的负载控制信息；以及

选择装置，用于响应于所述检查装置，选择用于所述消息或响应消息的处理资源。

此外，以上目的是通过一种用于将消息传送到分组数据网络的设备实现的，所述设备被设置为将负载控制信息设定在所述消息的预定字段内，所述负载控制信息用于选择数据所述分组数据网络的处理资源。

此外，以上目的是通过一种分配分组交换网络内的负载控制信息的方法实现的，所述方法包括步骤：

- 在第一网络单元内生成第一负载控制信息；
- 将所述第一负载控制信息设定到消息的预定字段内；
- 将所述消息路由到第二网络单元；
- 将所述第一负载控制信息存储在所述第二网络单元内；
- 在所述第二网络单元内生成第二负载控制信息；
- 将所述第二负载控制信息设定到第二消息的预定字段内；
- 将所述第二消息路由到所述第一网络单元；以及
- 将所述第二负载控制信息存储在所述第一网络单元内。

此外，以上目的是通过一种控制分组数据网络内的处理负载的系统实现的，所述系统包括：

第一网络单元，用于在将被在所述分组数据网络内路由的消息的预定字段内设定负载控制信息；以及

第二网络单元，用于在所述消息的路由路径上检查所述负载信息；以及响应于所述检查步骤的结果，选择所述分组数据网络的处理资源。

最后，以上目的是通过一种用于分配分组交换网络内的负载控制信息的系统实现的，所述系统包括：

第一网络单元，用于生成第一负载控制信息，并将所述第一负载控制信息设定到消息的预定字段内；以及

第二网络单元，用于接收所述消息，存储所述第一负载控制信息，生成第二负载控制信息，将所述第二负载控制信息设定到第二消息的预定字段内，以及将所述第二负载控制信息路由到所述第一网络单元；其中所述第一网络单元用于存储所述第二负载控制信息。

因此，负载分担或负载均衡信息可被在消息内路由到对应网络节点或服务器，从而减少负载控制功能所需的资源量。此外，在网络单元出现故障的情况下，仍可提供关于现有会话的信息。

所述预定字段可能经由 SIP 消息的分支。所述负载均衡信息可能是从另一预定字段拷贝到所述预定字段的。

所述预定字段可能是地址标题的用户部分的子字段，例如 SIP 路由标题的 URI。因此，附加信息可被在所述用户部分内传递。尤其是，公司秘密信息可被在例如在信息中加密和/或标记和/或编码的一个或多个子字段内传递。此外，例如可使用所述用户部分的一个或多个子字段来执行网络单元内的路由，例如选择正确的呼叫状态模型（例如用于始发或终止会话/事务情况）。所述负载控制信息和/或其它控制信息可能是特定于用户和/或业务和/或服务器的信息，其例如在一个或多个子字段内从 S-CSCF 传递到 AS。作为选择，IP 地址可能被在子字段内传递，其可能是所述域部分内的主机地址。

因此，多个子字段可被设置在所述用户部分内，用于传递不同类型的所述负载控制信息和/或其它控制信息。尤其是，所述用户部分可能被语法分析并划分为子字段。作为选择，所述子字段的结构、顺序和使用中的至少一个可能是在标准化基础上预定的。所述子字段可能由预定比特串、字符、字符串或其它分隔符分隔。

所述选择步骤可用于负载均衡。虚拟地址用于将所述消息分配给预定处理器节点。所述虚拟地址可由多个处理器节点共享。这些处理器节点可能具有基于 IP 的蜂窝网络的呼叫状态控制功能。因此，提供了一种确保平均负载均衡的机制，而用户可在整个注册期间内成为处理器节点。通过使用虚拟地址，群集自身可为群集节点执行负载均衡。

此外，所述负载控制信息可能包括第一端口号，所述第一端口号指示用于接收请求消息的第一端口。此外，所述负载控制信息可能包括第二端口号，所述第二端口号指示用于接收响应消息的第二端口。由此，可在提供负载均衡信息的情况下，在所指示端口处接收请求和响应。

另一方面，所述负载控制信息可能包括第一信息和任选第二信息。所述第一与任选第二信息可能被设置在从 SIP 消息的路由标题字段、通路标题字段或接触标题字段中选择的标题字段内。所述第一信息可能指示所述消息的会话是否已经存在。所述任选第二信息可能随后指示现有会话的识别。所述第一与第二信息可能是对于其它网络而言无意义的隐藏信息。

因此，提供了两种选择。在第一种选择中，仅检测所述消息是否为呼叫路线内的第一个。因此，可提供对于会话内第一消息的简单快速辨识。无需对其它网络单元或终端做出改变。这种方案甚至可能是在非标准化基础上设置的。在第二种选择中，基于所述第二信息检测附加会话识别。因此，除了第一种选择的以上优点之外，还可提供对于所述会话的简易快速识别。

尤其是，所述第一和/或第二信息可能被设置为标题字段的地址（例如 SIP URI）中的用户部分的一部分，标题字段的主机名称的一部分，标题字段的域名的一部分，所述标题字段的参数，所述标题的端口号，或所述消息的扩展标题字段，其中所述端口号可用于区分第一消息与后续消息。作为选择，所述第一和/或第二信息可能被设置在所述消息的有效负荷部分内。

然后，可能响应于检测所述第一信息，提取所述第二信息，所述第二信息可用于所述选择步骤。

用于控制处理负载的网络设备可能包括基于 IP 的蜂窝网络的呼叫状态或呼叫会话控制功能。所述选择装置可能被设置为选择所述消息被分配给的预定处理器节点。因此，除了虚拟地址之外，所述负载控制信息规定所述处理器节点。

所述选择装置可能被设置为启动新会话的创建。

用于传送所述消息的设备可能还包括基于 IP 的蜂窝网络的呼叫状态控制功能或呼叫会话控制功能。所述会话控制功能可能是服务呼叫会话控制功能、轮询呼叫会话控制功能或代理呼叫会话控制功能。所述设备可能被设置为在所述消息的标题地址的用户部分内，或作为

选择在主机名称、域名、标题参数、端口号或标题部分的扩展标题字段内，或在所述消息的有效负荷部分内，设置所述负载控制信息。

#### 附图说明

以下将基于优选实施例，参照附图描述本发明，在附图中：

图 1 示出了其内可实施本发明的 IMS 网络体系结构；

图 2 示出了根据第一优选实施例的 SIP URI 的结构；

图 3 示出了根据第一优选实施例的第一信令实例的信令和处理图；

图 4 示出了根据第一优选实施例的指示第二信令实例的信令和处理图；

图 5 示出了根据第二优选实施例的信令和处理图；

图 6 示出了根据第二优选实施例的负载共享机制的第一实例的流程图；以及

图 7 示出了根据第二优选实施例的负载共享机制的第二实例的流程图。

#### 具体实施方式

以下将基于图 1 所示的 IMS 网络体系结构描述优选实施例。

应当注意的是，图 1 仅示出了描述本发明所需的 IMS 部件。例如，图 1 并未示出无线电接入网与核心网。根据图 1，终端设备或 UE 10 连接到 P-CSCF 20，所述终端设备可能是移动或蜂窝终端设备，所述 P-CSCF 20 被设置在所述 UE 10 的受访域 70 内，并提供基本 IP 连接及其之下的移动管理。所述 UE 10 使用 SIP 与类似于 SIP 代理服务器的 P-CSCF 20 通信。在这种情况下，所述 UE 10 的消费者或用户在 P-CSCF 20 位于的受访域 70 内漫游。所述 P-CSCF 20 的职责为提供紧急呼叫，以及其它需要特别了解所述受访域 70 的业务。除了无线电接入网络之外，还可使用备选接入网络。除了移动或蜂窝终端设备之外，还可使用其它类型的终端，尤其是在备选接入网络内。

此外，S-CSCF 40 始终位于所述用户或消费者的原籍域 80 内，并承担 SIP 注册和代理服务器的职责，从而可经由 P-CSCF 20 使用 SIP 在 S-CSCF 40 处注册 UE 10。此外，I-CSCF 30 被设置为另一 SIP 代理服务器，主要负责找到特定用户或消费者的正确 S-CSCF。可在每次注册时动态分配所述 S-CSCF，以便实现有效负载均衡和错误滞留。应用服务器 (AS) 60 被设置为处理提供给 UE 10 的业务的 SIP 单元。可出于不同目的提供独立 AS。最后，原籍用户服务器 (HSS) 50 被设置用于简表管理和鉴权。

以下将基于图 2 至 6 描述本发明的第一优选实施例。

在第一优选实施例中，SIP URI 的用户部分的内容用于负载控制，例如会话控制和负载均衡。尤其是，所述 SIP URI (统一资源识别符) 的用户部分可分为例如用于控制和/或路由目的的子字段。在 SIP 中，所述路由标题通常包括仅具有域部分的 SIP URI，从而使得所述用户部分自由用于其它目的。

图 2 示出了根据第一优选实施例的 FQDN 或 SIP URI 100 的结构。所述 SIP URI 100 包括由“@”符号分隔的用户部分 120 和域部分 140，类似于电子邮件地址。SIP 所寻址的对象为所述 SIP URI 100 所识别的主机处的用户。所述用户部分 120 用于传递用户姓名或电话号码，而所述主机或域部分 140 传递域名或数字网络地址。

由于并未在路由标题和通路标题内使用所述用户部分，因此所述用户部分可分为 121, 122, ...12n，其可由适当的分隔符 130，例如比特串、字符或字符串分隔，其中所述字符可能是可打印和/或不可打印的字符或比特串。子字段 121 至 12n 的顺序和使用可以是预定的或标准化的，如果并不作为特定于实施方式的信息的话。

至于所述用户部分 120 内的子字段 121 到 12n 的设置与结构，可使用三种选择。

根据第一种选择，所述用户部分 120 可被设置为包括子字段 121 到 12n 的单个字段。然后在需要时传送所述单个字段，并将其划分为子字段 121 到 12n。这提供了如果在相同网络内生成和使用所述字段

则无须标准的优点。

根据第二种选择，所述用户部分 120 可能结构上包括子字段 121 到 12n，而所述子字段 121 到 12n 的语法和语义被预定义或标准化。在这种情况下，甚至可在不同网络内生成并使用所述子字段 121 到 12n。

根据第三种选择，可使用所述第一与第二种选择的组合。

以下是 SIP URI 100 的实例，其中第二子字段用于信令会话/事务情况，而第一子字段用于信令特定于实施方式的信息。所述分隔符 130 由校正符“\_”形成。

**57BC442C\_terminating@s-cscf2.ims.sonera.fi**

因此，“terminating”被作为会话/事务情况信令，而“57BC442C”被作为特定于实施方式的信息信令。

以下将参照图 3 和 4 描述根据第一优选实施例的负载控制机制的第一与第二实例。

所述负载控制机制被设置为确保平均的负载均衡，如果网络单元或部分由若干处理器节点实施的话。在根据图 1 的 IMS 网络体系结构内，所述 UE 10 使 P-CSCF 20 作为所述网络的接触点。在所述 UE 10 与 P-CSCF 20 之间，IP 安全功能 IPsec 用于完整和秘密保护。此外，压缩功能可用于压缩所述 UE 10 的 FQDN 或 SIP URI 100 的前缀或用户部分 120 内的信令信息。为了实现高容量和快速响应时间，UE 10 的特定用户的 IPsec 数据和压缩数据必须存储在存储器内，例如易失性存储器或随机存取存储器（RAM）内。结果，用户应当连接至其注册到的处理器节点。通过使用所述虚拟地址，仅节点或服务器群集自身即可为群集节点执行负载均衡。

图 3 示出了用于在用户注册时分配负载均衡信息（LBI）的负载控制机制的信令和处理图。此处，负载均衡的原理在于在所分配节点处执行压缩。因此，实际上，在所述 UE 10 与 P-CSCF 20 之间的接口处，无法基于 SIP 级信息执行负载均衡，因为其被压缩。在此接口处，合理的负载均衡键是 UE 10 的 IP 地址。当接收到指向 UE 10 的终止

尝试或请求时，关键是在消息从 UE 10 分配给的不同处理节点处接收终止尝试。因此，可在基于 IP 的网络内避免不必要的跳频。

当所述 UE 10 在步骤 1 中传送 SIP 注册消息时，所述 P-CSCF 20 在步骤 2 中生成第一负载均衡信息 LBI (P-CSCF-1)，并将其插入所述注册消息的标题的路径字段的 SIP URL (P-CSCF)。以后在从 S-CSCF 40 接收初始请求时，接收所述路径字段内的第一负载均衡信息 LBI (P-CSCF-1)。所述 P-CSCF 20 还生成第二负载均衡信息 LBI (P-CSCF-2)，并将其插入所述注册消息的通路分支。在所述通路分支处将接收所述第二负载均衡信息 LBI (P-CSCF-2) 以及对于所述注册消息的响应。所述第一与第二负载均衡信息 LBI (P-CSCF-1) 和 LBI (P-CSCF-2) 可能相同或不同。在步骤 3 和 4 内将所述注册消息经由 I-CSCF 30 路由到所述 S-CSCF 40。当所述 S-CSCF 40 从 P-CSCF 20 接收所述注册消息时，所述 S-CSCF 40 在步骤 5 中例如基于呼叫 ID 执行负载均衡。在步骤 6 中，所述 S-CSCF 40 然后将 SIP-URL (P-CSCF) 从路径字段存储到用户数据库，所述 SIP-URL (P-CSCF) 包括 P-CSCF 20 的第一负载均衡信息 LBI (P-CSCF-1)。在步骤 7 中，所述 S-CSCF 40 生成自己的负载均衡信息 LBI (S-CSCF-1)，并将其插入 SIP 200OK 响应消息的业务路由字段的 SIP-URL (S-CSCF)，所述 SIP 200OK 响应消息被在步骤 8 和 9 中经由 I-CSCF 30 发送到 P-CSCF 20。此负载均衡信息 LBI (S-CSCF-1) 将在从 P-CSCF 20 接收初始请求时被接收。此外，所述通路分支包括已从在步骤 4 之后接收的注册消息拷贝的 P-CSCF 20 的 SIP-URL (P-CSCF)。当所述 P-CSCF 20 接收所述 200OK 响应消息时，可在步骤 10 中基于从所述通路分支得到的第二负载均衡信息 LBI (P-CSCF-2)，执行负载均衡。在步骤 11 中，所述 P-CSCF 20 将 SIP-URL (S-CSCF) 存储在数据库内，所述 SIP-URL (S-CSCF) 包括从所述 200OK 响应消息的业务路由字段得到的 S-CSCF 40 的负载均衡信息 LBI (S-CSCF-1)。最后，在步骤 12 中，指示成功注册的 200OK 响应消息被转发给 UE 10。因此，在以上进程之后，所述 P-CSCF 20 在其用户数据中具有

SIP-URL (S-CSCF)，所述用户数据指向 S-CSCF 40，并包括对应负载均衡信息 LBI (S-CSCF-1)。类似地，所述 S-CSCF 40 在其用户数据中具有 SIP-URL (P-CSCF)，所述用户数据指向 P-CSCF 20，并包括对应负载均衡信息 LBI (P-CSCF-1)。所述负载均衡信息因此可由对应的 P-CSCF 20 和 S-CSCF 40 随后使用。例如，当终止尝试发生时，所述 S-CSCF 40 然后从用户数据库中提取所述负载均衡信息，并将其插入对应请求。

图 4 示出了负载控制机制的处理和信令图，所述负载控制机制用于在将启动请求发送到网络时使用负载均衡信息 (LBI)。当所述 UE 10 在步骤 1 内将 SIP 邀请消息发送到 P-CSCF 20 时，基于所述 UE 10 的 IP 地址执行负载均衡。在步骤 2 中，所述 P-CSCF 20 从所述用户数据库中读取先前所存储的 SIP-URL (S-CSCF)，所述 SIP-URL (S-CSCF) 用于在步骤 3 中将所述邀请消息路由到 S-CSCF 40。此外，所述 SIP-URL (S-CSCF) 被插入最上的路由字段，而 SIP-URL (P-CSCF) 被插入记录-路由字段。此外，第一负载均衡信息 LBI (P-CSCF-1) 被插入通路分支。当在 S-CSCF 40 处接收所述邀请消息时，其从最上路由字段中得到 SIP-URL (S-CSCF)，所述 SIP-URL (S-CSCF) 包括其先前所设置的负载均衡信息 LBI (S-CSCF-1)。基于所述负载均衡信息 LBI (S-CSCF-1)，在步骤 4 中执行负载均衡，以发现正确的计算机。当所述 S-CSCF 40 在步骤 5 中发送邀请消息时，其将 SIP-URL (S-CSCF) 插入记录-路由字段，并将其负载均衡信息 LBI (S-CSCF-1) 插入所述通路分支。在步骤 6 中，应用服务器 60 以 200OK 响应消息做出响应，所述 200OK 响应消息包括通路分支内的 P-CSCF 20 和 S-CSCF 40 的负载均衡信息 LBI (P-CSCF-1) 和 LBI (S-CSCF-1)。当所述 S-CSCF 40 接收 200OK 响应消息时，其从所述通路分支得到其负载均衡信息 LBI (S-CSCF-1)，并在步骤 7 中将其用于负载均衡。类似的，当所述 P-CSCF 20 随后在步骤 8 中接收所述 S-CSCF 40 所转发的 200OK 响应消息时，其从所述通路分支得到第一负载均衡信息 LBI (P-CSCF-1)，并在步骤 9 中将其用于负载均

衡。在步骤 10 中，所述 200OK 消息被转发到 UE 10，以确认先前的邀请消息。

当所述应用服务器 60 在步骤 11 中在对话内发送诸如邀请消息的后续 SIP 请求时，其使用路由列表，所述路由列表是所述应用服务器 60 先前基于在步骤 5 之后接收的初始请求的记录路由条目生成的，所述初始请求即邀请消息。所述最上路由条目是包括其内的对应负载均衡信息 LBI (S-CSCF-1) 的 SIP-URL (S-CSCF)。所述第二路由条目是包括其内对应负载均衡信息 LBI (P-CSCF-1) 的 SIP-URL (P-CSCF)。当所述 S-CSCF 40 接收后续邀请消息时，其从所述最上路由条目得到 SIP-URL (S-CSCF) 内的负载均衡信息 LBI (S-CSCF-1)。在步骤 12 中，所述 S-CSCF 40 基于此负载均衡信息 LBI (S-CSCF-1) 执行负载均衡，并删除指向其自身的路由条目。如今，所述最上路由条目指向 P-CSCF 20。在步骤 13 中，后续邀请消息被转发给 P-CSCF 20。当所述 P-CSCF 20 接收后续邀请消息时，其从最上路由条目得到在 SIP-URL (S-CSCF) 内提供的负载均衡信息 LBI (P-CSCF-1)。然后，在步骤 14 中，所述 P-CSCF 20 基于此负载均衡信息 LBI (P-CSCF-1) 执行负载均衡，并删除指向其自身的路由条目。最后，在步骤 15 中，所述邀请消息被转发给 UE 10。

一般而言，在注册阶段内，记录并存储所述路径和负载均衡信息，以随后用于请求。随后的请求应当包括此负载均衡信息，然后基于此负载均衡信息执行负载均衡。因此，在注册阶段期间内插入的任何负载均衡信息随后用于请求。

在所有情况下，所述通路标题字段的通路分支参数可用于传送所述负载均衡功能用来将响应分配给正确处理器节点的类似信息。

此外，不同端口号可用于识别可发现负载均衡信息之处。尤其是，在“路径记录”期间内，所述请求端口被设置为 SIP URI 120 的域部分 140。因此，请求然后被在请求端口处接收，并了解何处读取所述负载均衡信息。类似地，“端口”可被设置用于呼出请求，从而可在此端口处接收所述响应。

在执行对于指向相关网络单元的输入或呼入 SIP 业务的负载均衡功能的情况下，检查目的地 IP 地址是否为虚拟 IP 地址。如果否，则无需负载均衡，例如正常的 L3 可应用于分组。如果目的地 IP 地址是虚拟 IP 地址，则需要负载均衡。存在两种选择，即所述虚拟 IP 地址可能是 P-CSCF 地址，或 S-CSCF 或 I-CSCF 地址。对应的目的地端口信息用于检测所述负载均衡信息的种类和位置。然后，基于所述负载均衡信息执行负载均衡，而结果输出对应于分组应当被转发到的正确处理节点。所述负载均衡信息可能是呼叫 id、UE-IP 地址或先前所生成的负载均衡信息。

在从相关网络单元始发的输出或呼出业务的情况下，检查源 IP 地址是否为 CPS 的虚拟 IP 地址（P-CSCF、S-CSCF 或 I-CSCF）中的一个。如果否，则正常路由发生。如果是，检查其是否为环路地址。如果如此，则确定传送协议，并随后检查目的地端口，以确定请求端口或专用端口。基于所述检查结果，选择用于得到所述负载均衡信息并转发所述 IP 分组的进程。在非环形地址的情况下，再次检查源 IP 地址，以确定其是 S-CSCF/I-CSCF 地址还是 P-CSCF 地址。如果其是 S-CSCF/I-CSCF 地址，则确定传送协议，即传输控制协议（TCP）或用户数据报协议（UDP）。如果指示 TCP，则在缓存之后重新集合 SIP 消息，然后将其转发。如果指示 UDP，则可直接转发 IP 分组。如果其是 P-CSCF 地址，则确定源端口。如果指示客户机端口或请求端口，则确定传送协议，并再次启动以上处理。在指示 UE 不安全/安全客户机/服务器端口的情况下，所述 IP 分组直接由 L3（协议层 3）进程直接转发。

以下将参照图 5 和 7 描述本发明第二优选实施例。所述第二优选实施例指向负载共享机制，用于以有效方式发现哪个 SIP 业务属于哪个会话，以及诸如 SIP INVITE 请求的请求是初始请求还是再次请求。

图 5 示出了指示根据第二优选实施例的负载均衡机制的第一实例的处理和信令图。第一实例的机制适合于检测诸如 SIP INVITE 的 SIP 请求是否为呼叫路线中的第一个。这是通过在 SIP 请求的记录-路由标

题字段内提供隐藏信息得以实现的。

只要诸如图 1 内 S-CSCF 40 的会话状态 CSCF 接收 SIP 请求并插入记录-路由标题字段（步骤 1），会话状态 CSCF 即将隐藏指示插入记录-路由标题字段（步骤 2），并将带有所述隐藏指示的 SIP 请求转发到目的地地址。在当前情况下，“隐藏”意味着所述指示对于其它网络无意义，例如其内设置指示的网络之外的网络。然而，如果需要，则所增加的指示可能也是其它网络可读的标准化信息。

然后，只要 INVITE 到达，会话状态 CSCF 即检查最上路由标题字段或请求 URI（如果无路由标题）是否包括所述隐藏信息。所述路由标题字段是根据记录-路由标题字段构建的。如果隐藏信息或指示存在，则所述会话业已存在。如果否，则必须在相关会话状态 CSCF 内部创建新会话。

由于 SIP 响应（例如 200OK）在正常情况下属于现有会话，因此无需对其辨别。

所述指示可是主机名的一部分。例如，假定指向所述单元的缺省路由地址为 <scscf17.operator.net>，例如 <B@provider.net; maddr=scscf17.operator.net>，则所述记录路由可能类似于：

**RecordRoute: <B@provider.net;  
maddr=exist.scscf17.operator.net>**

或

**Record-Route: <B@exist.provider.net>**

或

**RecordRoute: <B@exist.scscf17.operator.net>**

此处，隐藏指示可能是作为主机名的一部分的“exist.”。所述用户部分 120 在这些实例中可能是空的。例如，替代 <B@exist.scscf17.operator.net>，以下可能使用：<exist.scscf17.operator.net>。所述路由标题字段是根据记录-路由标题字段构建的。例如：

**Record-Route: <B@exist.scscf17.operator.net>**

提供

**Route: <B(@exist.scscf17.operator.net)>**

作为选择，如 RFC3261 内所定义，所述 S-CSCF 40 可能将“rr-param”加入记录-路由标题字段。

**Record-Route="Record-Route" HCOLON rec-route\*(COMMA rec-route)**

**rec-route = name-addr\*( SEMI rr-param )**

**rr-param = generic-param**

**generic-param = token [ EQUAL gen-value ]**

**gen-value = token / host / quoted-string**

**token = 1\*(alphanum /“-”/“、”/“!” /“%”/“\*” /“\_”/“+”/“`” /“'”/“\_”**

**Route = "Route" HCOLON route-param \*(COMMA route-param) route-param = name-addr\*( SEMI rr-param)**

对应实例可能类似于：

**RecordRoute:<B@provider.net;maddr=scscf17.operator.net;existing>**

或

**Record-Route: <B@provider.net; existing>**

或

**Record-Route: <B@scscf17.operator.net: existing>**

此外，所述用户部分 120 可能是空的。所述路由标题字段被根据记录路由标题字段构建。

如果接收到“存在”，则可轻易检测到所述请求属于现有会话。

根据另一选择，不同端口号可能用于第一和后续请求。例如，假定指向相关单元或节点的缺省路由可能是 <B@provider.net;maddr=scscf17.operator.net>。然后，所述记录路由条目可能类似于：

**RecordRoute:<B@provider.net;maddr=scscf17.operator.net:19373>**

或

**Record-Route: <B@>provider.net:19373>**

或

**Record-Route: <B@scscf17.operator.net:19373>**

此处，所述用户部分 120 可能依然为空。根据所述记录-路由标题字段构建所述路由标题字段。例如

**Record-Route:<B@provider.net;maddr=scscf17.operator.net:19373>**

给出

**Route:<B@provider.net;maddr=scscf17.operator.net:19373>**

所有到达端口 19373 的请求可能被识别为属于现有会话。

根据另一选择，可能在 SIP 内使用新的专有或非标准化扩展标题字段来传递信息。新标题条目的实例可类似于：

**CSCF-session: existing in scscf17.operator.net**

然而，在这种情况下，UA 可能必须支持此特征，即从所述 SIP 请求将所述新标题字段拷贝到 SIP 响应和后续 SIP 请求。

根据其它选择，所述 SIP 请求的有效负荷部分可用于传递隐藏指示。

图 6 示出了指示根据第一实例的负载共享机制的示意流程图。当初始请求由单元处理时，记录-路由标题可被在发送之前插入所述请求。可将隐藏信息加入此记录-路由标题。然后，这些记录-路由标题被在响应内传送回所述请求的始发者。当所述始发者取得此响应时，其取得所述记录-路由标题，并将它们拷贝到路由列表。当所述始发者发送后续请求时，其取得所述路由列表，并将所有条目作为路由标题插入后续请求。然后，可将后续请求发送出去。当下一个服务器接收所述请求时，其可从所述路由标题发现其先前插入初始请求的相同 SIP URI。通路标题同样被插入所述请求。响应于所述请求接收相同通路标题，所述通路标题内的隐藏信息可用于发现所述响应必须被传递到的实例或单元。

因此，根据图 6，如果接收到 SIP 请求，则为隐藏指示检查路由

标题字段，或新标题字段或有效负荷部分（步骤 S201）。如果在步骤 S202 中检测到隐藏指示，则会话业已存在，无需创建，且可将请求分配给现有会话的呼叫线路（步骤 S204）。另一方面，如果未检测到隐藏指示，则在步骤 S203 中创建新会话。

根据负载共享机制的第二实例，基于所述隐藏指示检测内部会话识别符（ISId）。此备选包括第一实例的以上机制，即如果无法发现 ISId，则可假定所述请求属于新呼叫路线。

在第二实例中，在记录-路由标题字段和 SIP 请求的通路标题字段内提供了隐藏指示。

因此，参照图 5，只要诸如 S-CSCF 40 的会话状态 CSCF 插入记录-路由或通路标题字段，所述会话状态 CSCF 即会加入包括关于内部会话识别符（ISId）的信息的隐藏指示。

图 7 示出了指示根据第二实例的增强型负载共享机制的示意性流程图。只要消息到达，即检查所述消息是否对应于 SIP 请求（步骤 S301）。如果是，则所述会话状态 CSCF 在步骤 S303 中检查最上路由标题字段是否包括隐藏指示。如果在步骤 S305 中确定所述隐藏指示存在，则所述会话存在，而可提取 ISId，以提供快速会话辨识和分配功能（步骤 S307）。如果否，则必须在步骤 S306 中创建新会话。

如果在步骤 S301 中未检测到 SIP 请求，则在步骤 S302 中检查所述消息是否对应于 SIP 响应。如果是，则所述会话状态 CSCF 在步骤 S304 中检查最上通路标题字段是否包括隐藏指示。如果在步骤 S305 中确定所述隐藏指示存在，则会话状态 CSCF 从 SIP 响应的最上通路标题字段提取 ISId（步骤 S307）。

因此，甚至可在 CSCF 处提供现有会话的快速识别。

正如第一实例内一样，所述隐藏指示可能是主机名的一部分。例如，假定指向此单元的缺省路由可能是 <scscf17.operator.net>，例如 <B@provider.net;maddr=scscf17.operator.net>。然后，所述路由标题字段如下所示。根据记录-路由标题字段构建所述路由标题字段。

**Route:**

**<B@provider.net;maddr=ISId224497.scscf17.operator.net>**

或

**Route:<B@ISId224497.provider.net>**

或

**Route:<B@ISId224497.scscf17.operator.net>**

此处，所述 ISId 是作为主机名一部分的“224497”。所述用户部分 120 可能为空。

类似地，可能会使用所述通路标题字段，其类似于：

**Via: SIP/2.0/UDP ISId224497.scscf17.operator.net:5060**

所有包括作为主机名一部分的“224497”的 SIP 响应都属于相同的现有会话。

作为选择，出于隐藏或冗余目的，作为主机名一部分的 ISId 可能同样被加密/标记。

作为选择，如 RFC3261 内所定义，所述 S-CSCF 40 可能将“rr-param”加入记录-路由标题字段。

类似地，这用于 SIP 响应，其中如 RFC3261 内所定义，所述通路标题字段可能由“via-extension”扩展：

**Via = ( "Via"/ "v") HCOLON via-param\*(COMMA via-param)**

**via-param =sent-protocol LWS sent-by \*(SEMI via-params)**

**via-params = via-ttl / via-maddr**

**/ via-received / via-branch**

**/via-extension**

**via-ttl = "ttl" EQUAL ttl**

**via-maddr = "maddr" EQUAL host**

**via-received = "received" EQUAL (IPv4address / IPv6address)**

**via-branch = "branch" EQUAL token**

**via-extension = generic-param**

**sent-protocol = protocol-name SLASH protocol-version**

**SLASH transport****protocol-name = "SIP" /token****protocol-version = token****transport = "UDP" /"TCP" / "TLS" / "SCTP"****/other-transport****sent-by = host [ COLON port ]****til = 1\*3DIGIT; 0 to 255****generic-param = token [ EQUAL gen-value ]****gen-value = token / host / quoted-string****token = 1\*(alphanum /"-"/"."/"!"/"%"/\*"****/"\_"/"+"/"/"/"/"/"~"**

例如，所述路由标题字段可能如下所示。根据记录-路由标题字段构建所述路由标题字段。

**Route:<B@provider.net;maddr=  
scscf17.operator.net;ISId=224497>**

或

**Route:<B@provider.net;ISId=224497>**

或

**Route:<B@scscf17.operator.net;ISId=224497>**

此处，ISId 是“224497”，作为路由标题字段的参数。

对应的通路标题字段可能类似于：

**Via: SIP/2.0/UDP scscf17.operator.net:5060: ISId-224497**

在这种情况下，所有包括参数“ISId=224497”的 SIP 响应都属于相同的现有会话。

作为选择，出于隐藏或冗余目的，还会加密/标记作为主机名一部分的 ISId。

根据另一备选，不同端口号可用于所有现有会话：

例如，假定指向相关单元的缺省路由是 <B@provider.net;maddr=scscf17.operator.net>。然后，所述路由标题字段如下所示。

根据记录-路由标题字段构建所述路由标题字段。

**Route:<B@provider.net;maddr=scscf17.operator.net:224497>**

或

**Route:<B@provider.net:224497>**

或

**Route:<B@scscf17.operator.net:224497>**

所述用户部分可能为空。所有到达端口 224497 的 SIP 请求都属于相同现有会话。

类似地，这用于所述通路标题字段，其类似于：

**Via: SIP/2.0/UDP scscf17.operator.net:224497**

此处，所有到达端口 224497 的 SIP 响应都属于相同的现有会话。

然而，同时收听许多端口引起性能或可量测问题。

作为另一选择，可能会在 SIP 内使用新的专有扩展标题字段来传递所述隐藏信息。例如，新的扩展标题条目的实例可类似于：

**CSCF-session: "ISId=224497 in scscf17.operator.net"**

然后，如以上结合第一实例所述，所述 UA 可能必须支持此特征，即从 SIP 请求将所述新标题字段拷贝到 SIP 响应和后续 SIP 请求。

作为另一选择，所述 SIP 请求或响应的有效负荷部分同样可用于此。

以上负载共享机制主要设置在带有对应功能的 CSCF 或其它网络节点内。然而，一般而言，它们可能还实施在诸如 UE 10 的终端设备内。

应当注意的是，本发明并不仅限于上述优选实施例。本发明可实施在任何网络中的具有负载控制功能的任何网络节点内。尤其是，可使用任何分组数据消息或数据报的任何标题或有效负荷字段。此外，可传递任何可用于负载控制的信息。因此，实施例可在所附权利要求书的范围内改变。

图1

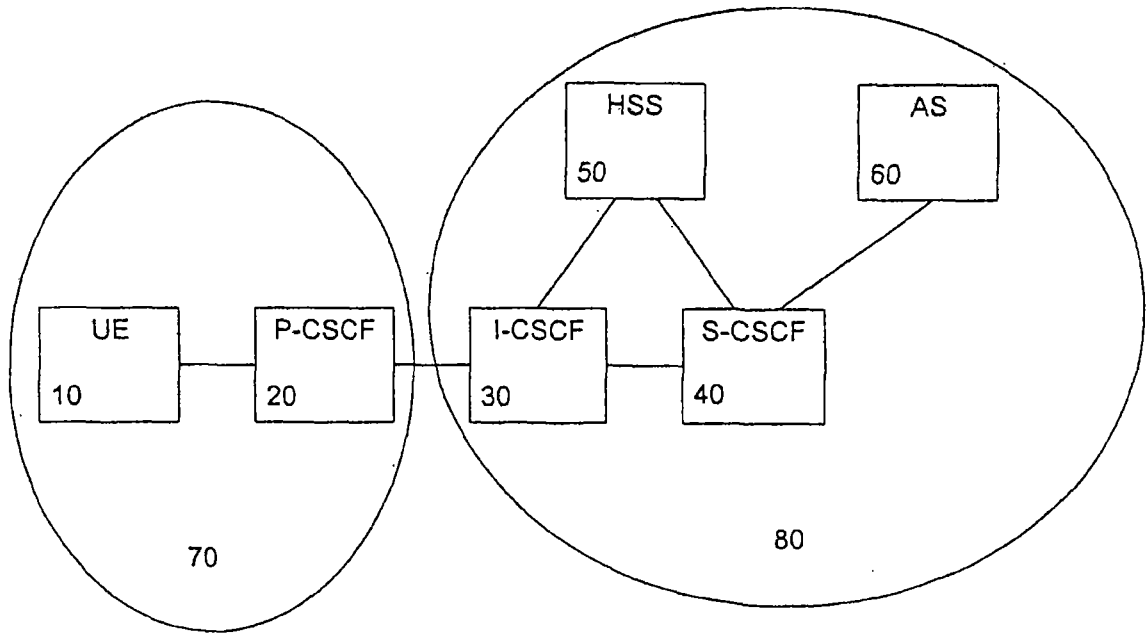


图2

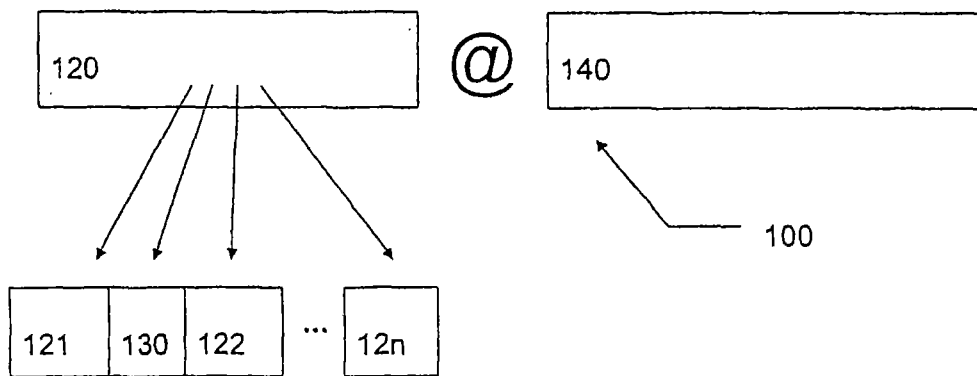


图3

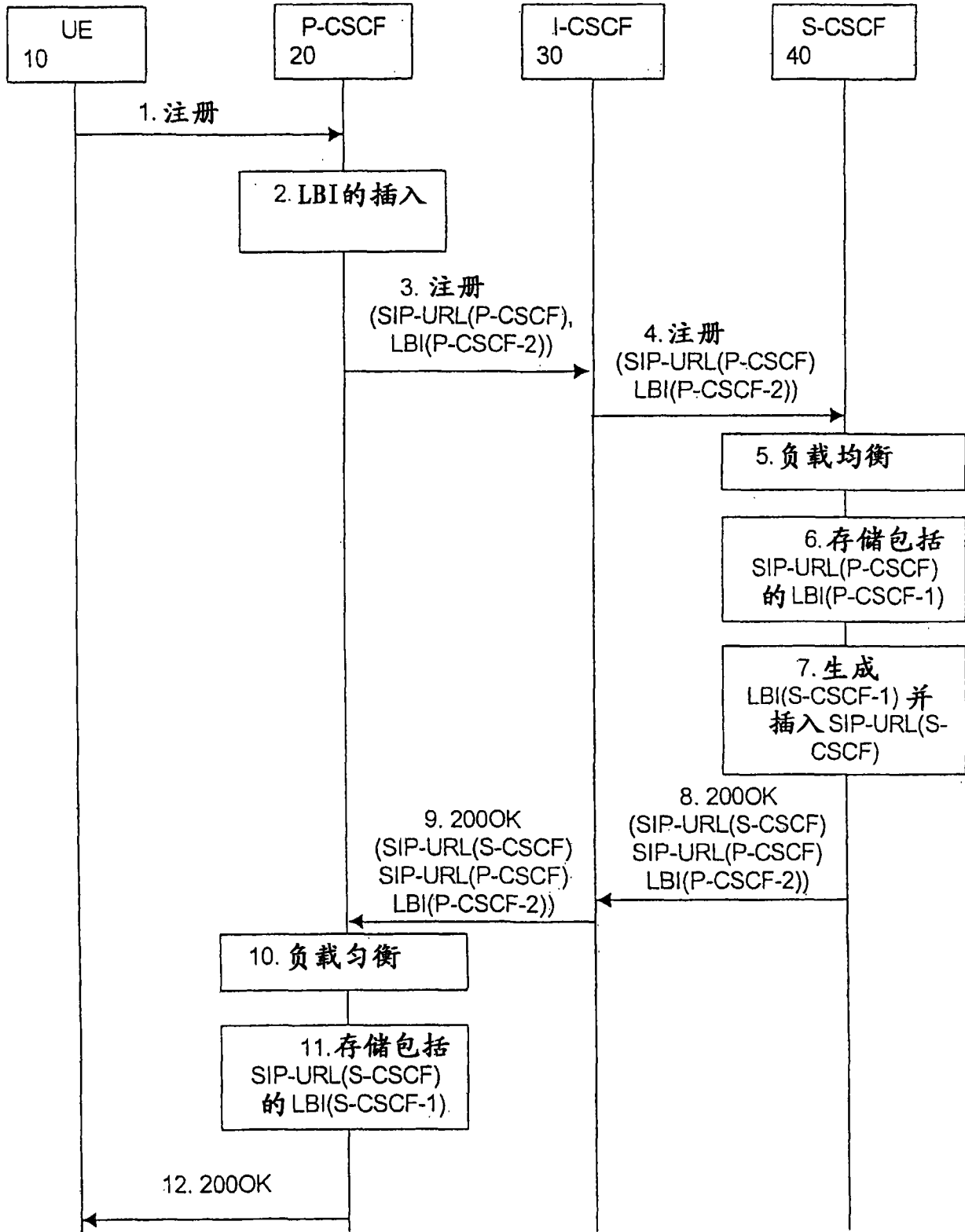


图4

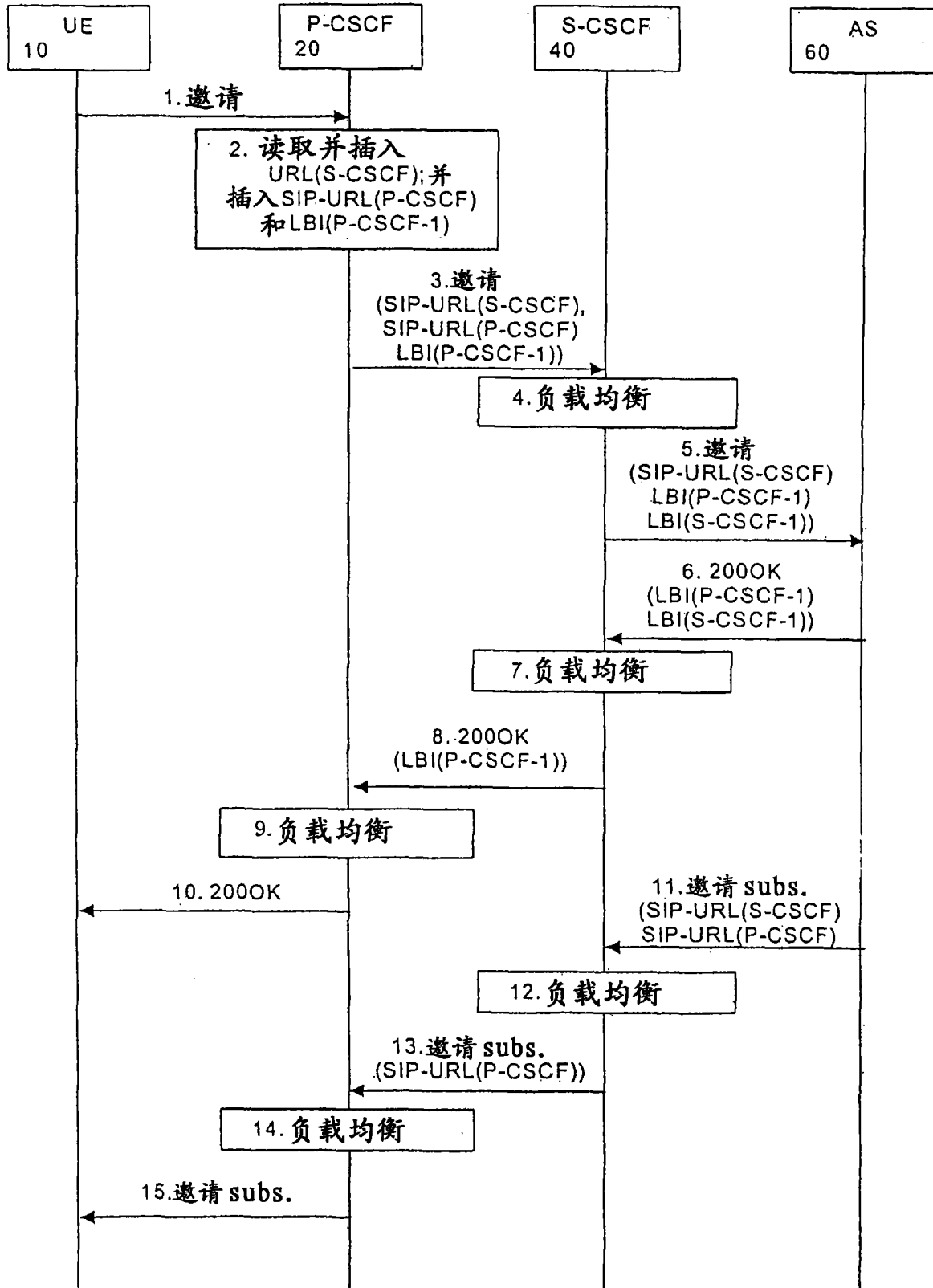


图5

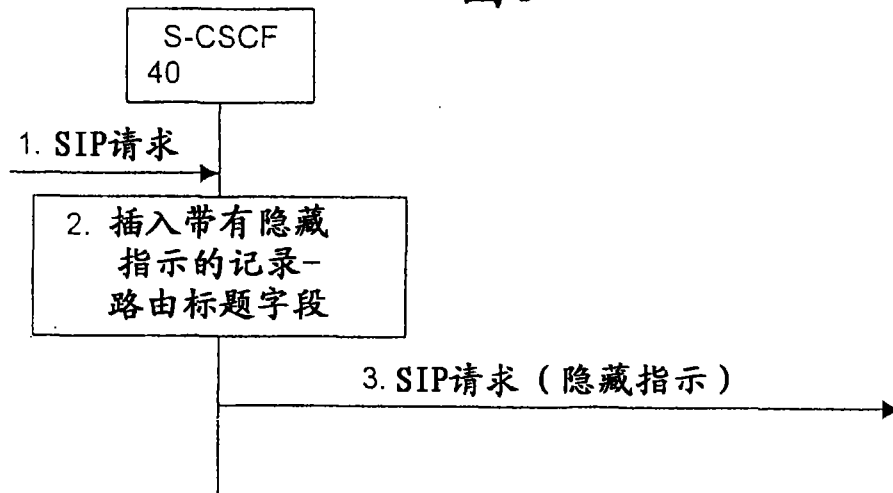


图6

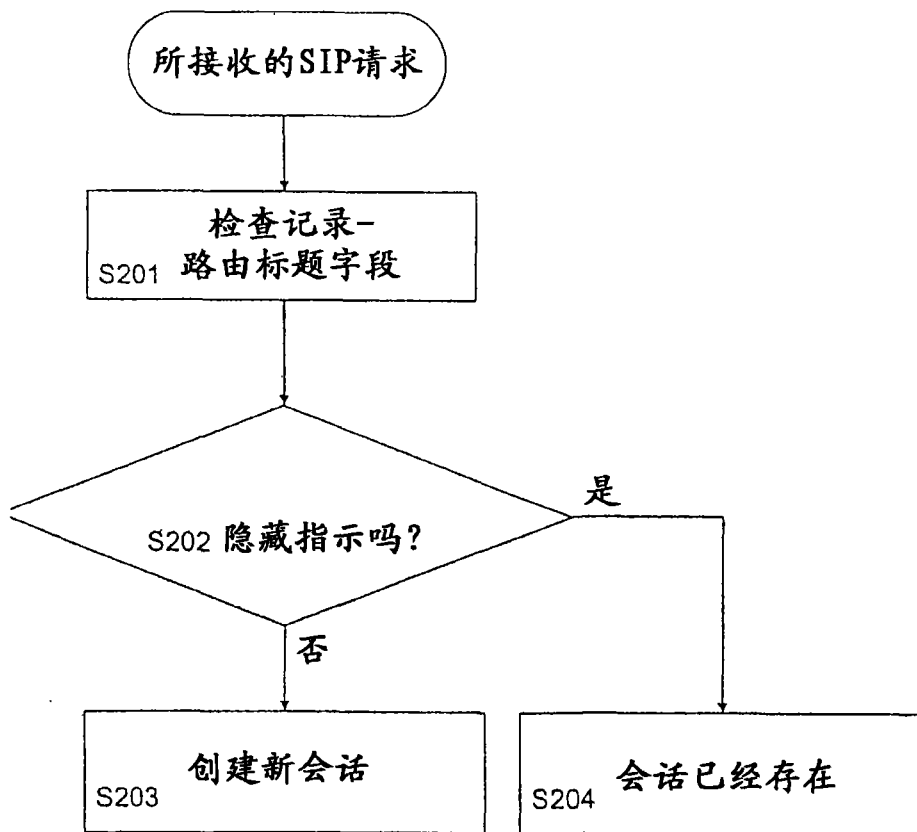


图7

