



## [12] 发明专利申请公布说明书

[21] 申请号 200810009921.9

[43] 公开日 2008 年 8 月 20 日

[11] 公开号 CN 101247393A

[22] 申请日 2008.2.13

[21] 申请号 200810009921.9

[30] 优先权

[32] 2007.2.13 [33] US [31] 11/674,246

[71] 申请人 国际商业机器公司

地址 美国纽约

[72] 发明人 M·A·科 R·J·雷西奥  
J·A·瓦尔加斯

[74] 专利代理机构 北京市中咨律师事务所

代理人 于 静 李 峰

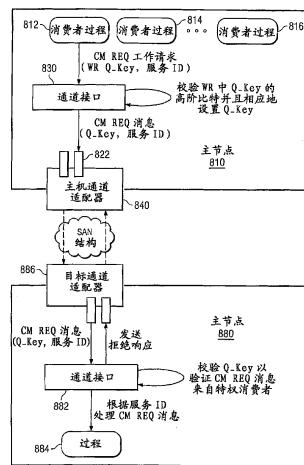
权利要求书 3 页 说明书 21 页 附图 8 页

## [54] 发明名称

用于防止 IP 欺骗和促进专用数据区的解析的  
系统和方法

## [57] 摘要

本发明提供了一种用于在系统和网络连接请求中防止 IP 欺骗和促进对专用数据区的解析的系统和方法。利用所述系统和方法，诸如 Q\_Key 的高阶比特这样的标识符被用于确定通信连接请求是否始于特权过程。第二标识符用于指定通信连接请求的专用数据区是否利用了预定结构或格式的预定字段。只有当所述第一标识符指定该请求始于特权过程时，才准许进行对所述请求的处理。基于所述第二标识符的设置，从所述专用数据区的预定字段中检索特定信息，以便在建立所请求的通信连接时使用。



1. 一种在数据处理系统中用于处理通信连接请求的方法，所述方法包括：

确定通信连接请求是否具有指示所述通信连接请求始于特权过程的第一标识符；

如果所述通信连接请求始于特权过程，则确定所述通信连接请求是否具有指定所述通信连接请求的专用数据区利用预定格式下的预定字段的第二标识符；

依照所述预定格式的预定字段，处理所述通信连接请求的专用数据区中的信息；以及

使用所述通信连接请求的专用数据区中所处理的信息来建立通信连接。

2. 根据权利要求 1 的方法，其中所述通信连接请求是通信管理请求消息，其具有专用数据区以及在其中提供了所述第一和第二标识符的报头。

3. 根据权利要求 2 的方法，其中所述第一标识符是 Q\_Key。

4. 根据权利要求 2 的方法，其中所述第二标识符是通信管理请求消息中的服务标识符。

5. 根据权利要求 1 的方法，其中所述数据处理系统是系统区域网中的主节点，并且其中所述方法是在所述主节点的通信适配器中实现的。

6. 根据权利要求 1 的方法，其中所述专用数据区中的信息包括：所述专用数据区的预定字段中的源 IP 地址或目的 IP 地址中的至少一个。

7. 根据权利要求 1 的方法，其中所述通信连接请求包括：使用远程直接存储器访问操作从另一数据处理系统传递的一个或多个不可靠数据报分组。

8. 根据权利要求 1 的方法，其中如果所述通信连接请求始于特权过程，则将所述第一标识符设置成第一值，以及如果所述通信连接请求始于非特权过程，则将所述第一标识符设置成第二值，并且其中只有操作系统或特

---

权过程可以将所述第一标识符设置成所述第一值。

9. 根据权利要求 1 的方法，其中在所述数据处理系统的第一主节点中实现所述确定步骤、处理步骤和建立步骤，并且其中所述方法进一步包括：

在所述数据处理系统的第二主节点中，接收来自运行在远程主节点中的过程的通信连接请求；

在所述第二主节点中，确定所述过程是否是特权过程；

如果所述过程不是特权过程，则在所述第二主节点中修改所述通信连接请求，以便将所述第一标识符设置成与关联于所述过程的队列对的第一标识符的值相对应的值；以及

将所述通信连接请求从所述第二主节点发送至所述第一主节点。

10. 根据权利要求 9 的方法，其中如果所述第一标识符被设置成指示所述通信连接请求始于特权过程，那么将所述第一标识符设置成与由所述过程发送并在所述第二主节点中接收的通信连接请求中所提供的第一标识符的值相对应的值。

11. 根据权利要求 1 的方法，其中确定通信连接请求是否具有指示所述通信连接请求始于特权过程的第一标识符包括：

确定所述通信连接请求是否以与所述数据处理系统的操作系统相关联的队列对为目标；

确定所述第一标识符是否对应于与所述操作系统相关联的队列对；以及

如果所述第一标识符并不匹配于与所述操作系统相关联的队列对的标识符，则拒绝所述通信连接请求。

12. 根据权利要求 11 的方法，其中确定通信连接请求是否具有指示所述通信连接请求始于特权过程的第一标识符进一步包括：

如果所述通信连接请求并不以与所述操作系统相关联的队列对为目标，则确定所述第一标识符是否设置了高阶比特；

如果所述第一标识符设置了高阶比特，则确定所述第一标识符是否与所述通信连接请求的目标队列对的标识符相匹配；以及

如果所述第一标识符并未设置高阶比特，或者所述第一标识符并不与所述通信连接请求的目标队列对的标识符相匹配，则拒绝所述通信连接请求。

13. 一种计算机系统，其包括用于实现权利要求 1 至 12 中任何一项的方法的装置。

## 用于防止 IP 欺骗和促进专用数据区的解析的系统和方法

### 技术领域

本申请一般涉及一种改进的数据处理系统和方法。更具体而言，本申请针对的是一种用于在系统区域网连接请求中防止网际协议（IP）欺骗和促进专用数据区的解析的系统和方法。

### 背景技术

在组网协议中，提供用于确保只有特权或受信应用才能够访问特定资源的保护是重要的。换句话说，组网协议能够依赖于这样的事实是重要的，即可以相信特定资源是不可被可能有意或无意破坏网络操作的应用或连接到网络的数据处理系统访问的。

举例来说，在诸如因特网这样的传输控制协议（TCP）网络中，非特权应用一般不能够伪造源网际协议（IP）地址，并且通常被防止使用特定的源端口。此外，不允许非特权客户机绑定（即建立软件链路）到任何地址，并且不允许其发送原始以太网分组（即未通过 TCP/IP 编程接口处理的数据分组）来绕过主机栈（host stack）。这是因为原始套接字不可通过非特权应用访问。因此，TCP daemon（后台程序）过程能够作为 IP 连接建立的一部分假设所提供的远程 IP 地址是有效的，除非伪造者（即给出未认证 IP 地址的非特权应用）在远程客户机具有根访问（root access）。

在系统区域网环境中，例如 InfiniBand™ 网络体系结构环境，通过 InfiniBand™ 体系结构规范卷 1 和 2，版本 1.2 中所描述的多个机制促进了 TCP/IP 通信，该规范可从 [www.infinibandta.org/specs/](http://www.infinibandta.org/specs/) 处的 InfiniBand™ 贸易协会获得。出于讨论本发明的目的，假设人们熟悉 InfiniBand™ 规范，其可以很容易从 InfiniBand™ 贸易协会获得，并且因而，在此不再提供对 InfiniBand™ 网络上的 TCP/IP 通信中所涉及的所有机制的详细解释。

当应用需要将要在连接建立期间使用的 IP 地址时，例如对于在 InfiniBand™ 网络环境上的 TCP/IP 连接，通常在通信管理请求 (CM REQ) 消息的专用数据区中传达 IP 地址。然而，由于提供给接收 CM REQ 消息的监听者的远程 IP 地址不可以被假定为是认证了的 (authentic)，因此这并没有实现 TCP/IP 连接建立的语义。换句话说，随着允许 TCP daemon 信任由远程客户机提供的远程 IP 地址的 TCP/IP 网络一起使用的相同保护机制并不是关于 InfiniBand™ 网络中 CM REQ 消息的专用数据区而呈现的。因此，伪造者可以在 CM REQ 消息的专用数据区中插入未认证的 IP 地址，并且由此获取对远程系统资源的未授权访问。因而，无法知道用户模式应用只不过没有虚构 IP 地址并且将其提交作为正常的 CM REQ 专用数据的一部分。这是因为，在 InfiniBand™ 网络中，用户空间消费者可以提供想要作为 CM 专用数据来使用的任何数据。因而，有可能用户空间消费者可以潜在地采用此来进行 IP 欺骗。也就是说，用户空间消费者可以将未授权的 IP 地址放入专用数据区，并且由此能够建立通信连接并且访问用户空间消费者不应该访问的远程资源。

## 发明内容

说明性实施例提供了一种机制来消除用户空间消费者通过在 InfiniBand™ 网络中的通信管理请求 (CM REQ) 消息的专用数据区内提供未授权的 IP 地址来进行网际协议 (IP) 欺骗的能力。此外，说明性实施例提供了一种协议，由此可以读取 CM REQ 消息的专用数据区中所提供的数据，并且将其解释用于 TCP/IP 连接建立和通信。

利用说明性实施例的机制，CM REQ 仅被限制于特权消费者。这可以通过检查所述 CM REQ 中所提供的 Q\_Key，由连接建立事务的被动方 (passive side) 来检验。利用说明性实施例的机制，仅准许指示所述 CM REQ 来自特权消费者的受控 Q\_Keys 建立对于 TCP/IP 通信的通信连接。在一个说明性实施例中，通过使用具有指定 CM REQ 消息的始发方是否是特权应用的这些 Q\_Keys 的高阶比特的 Q\_Keys，有可能确保只有特权消费

者才可以通过所述 CM REQ 消息来处理连接建立。如果设置了 Q\_Key 的高阶比特，那么所述 Q\_Key 与关联于特权应用的特权队列对 (privileged queue pair) 相关联。通过检查所述 Q\_Key 的该高阶比特，所述连接建立的被动方有可能能够确定在 CM REQ 消息中提供的 IP 地址是否是认证了的并且可以被信任。这消除了用户空间消费者可以提供要用作 CM REQ 中的专用数据的任何东西的可能性。

此外，作为说明性实施例的进一步的特征，为了限制可以在系统区域网 (SAN) 环境 (例如 InfiniBand<sup>TM</sup> 网络) 中的 CM REQ 专用数据字段中传递的信息的类型，使用了新的服务标识符 (ID)。使用该新的服务 ID 通知 InfiniBand<sup>TM</sup> 结构 (即交换机、通道适配器等) 将要以定义的方式解释所述 CM REQ 专用数据字段。因而，举例来说，用于连接建立的 IP 地址和其它关键信息在 CM REQ 专用数据区中具有其自己的定义字段。因此，连接建立事务的被动方知道在 CM REQ 的专用数据区中的什么地方获得在主动方与被动方之间建立 TCP/IP 连接所必需的信息。

在一个说明性实施例中，提供了一种用于处理通信连接请求的方法。所述方法可以包括：确定通信连接请求是否具有指示所述通信连接请求始于特权过程的第一标识符，并且如果所述通信连接请求始于特权过程，则确定所述通信连接请求是否具有指定所述通信连接请求的专用数据区利用预定格式下的预定字段的第二标识符。所述方法可以进一步包括：依照所述预定格式的预定字段，处理所述通信连接请求的专用数据区中的信息。此外，所述方法可以包括：使用所述通信连接请求的专用数据区中所处理的信息来建立通信连接。

所述通信连接请求可以是通信管理请求消息，其具有专用数据区以及在其中提供了所述第一和第二标识符的报头 (header)。所述第一标识符可以是 Q\_Key。所述第二标识符可以是通信管理请求消息中的服务标识符。如果所述通信连接请求始于特权过程，则可以将所述第一标识符设置成第一值，并且如果所述通信连接请求始于非特权过程，则可以将所述第一标识符设置成第二值。只有操作系统或特权过程可以将所述第一标识符

设置成所述第一值。

数据处理系统可以是系统区域网中的主节点。可以在所述主节点的通道适配器中实现所述方法。所述专用数据区中的信息可以包括所述专用数据区的预定字段中的源网际协议（IP）地址或目的IP地址中的至少一个。所述通信连接请求可以包括使用远程直接存储器访问（RDMA）操作从另一数据处理系统传递的一个或多个不可靠数据报分组。

可以在所述数据处理系统的第一主节点中实现所述确定步骤、处理步骤和建立步骤。所述方法可以进一步包括：在所述数据处理系统中的第二主节点中，接收来自运行在远程主节点中的过程的通信连接请求，以及在所述第二主节点中，确定所述过程是否是特权过程。所述方法还可以包括：如果所述过程不是特权过程，则在所述第二主节点中修改所述通信连接请求，以便将所述第一标识符设置成与关联于所述过程的队列对的第一标识符的值相对应的值。此外，所述方法可以包括：将所述通信连接请求从所述第二主节点发送至所述第一主节点。如果所述第一标识符被设置成指示所述通信连接请求始于特权过程，那么可以将所述第一标识符设置成与由所述过程发送并在所述第二主节点中接收的通信连接请求中所提供的第一标识符的值相对应的值。

确定通信连接请求是否具有指示所述通信连接请求始于特权过程的第一标识符可以包括：确定所述通信连接请求是否以与所述数据处理系统的操作系统相关联的队列对为目标；确定所述第一标识符是否对应于与所述操作系统相关联的队列对；以及如果所述第一标识符并不匹配于与所述操作系统相关联的队列对的标识符，则拒绝所述通信连接请求。确定通信连接请求是否具有指示所述通信连接请求始于特权过程的第一标识符可以进一步包括：如果所述通信连接请求并不以与所述操作系统相关联的队列对为目标，则确定所述第一标识符是否设置了高阶比特；如果所述第一标识符设置了高阶比特，则确定所述第一标识符是否与所述通信连接请求的目标队列对的标识符相匹配；以及如果所述第一标识符并未设置高阶比特，或者所述第一标识符并不与所述通信连接请求的目标队列对的标识符相匹

配，则拒绝所述通信连接请求。

在其它的说明性实施例中，提供了一种计算机程序产品，其包括具有计算机可读程序的计算机可用介质。当在计算设备上执行时，所述计算机可读程序使得所述计算设备实现以上关于方法说明性实施例所概括的操作中的各种操作及其组合。

在另一说明性实施例中，提供了一种装置。所述装置可以包括处理器以及耦合于所述处理器的存储器。所述存储器可以包括这样的指令，即当由所述处理器执行时，该指令使得所述处理器实现以上关于方法说明性实施例所概括的操作中的各种操作及其组合。

鉴于以下对本发明的示例性实施例的详细描述，本发明的这些以及其它的特征和优点将得以描述，或者对本领域的普通技术人员来说将变得显而易见。

#### 附图说明

当结合附图阅读时，通过参照以下对说明性实施例的详细描述，将最好地理解本发明以及优选使用模式，及其进一步的目的和优点，在附图中：

图 1 是可以在其中实现说明性实施例的示例性方面的分布式计算机系统的示例图；

图 2 是依照一个说明性实施例说明了主处理器节点的示例性软件和硬件方面的示例图；

图 3 是依照一个说明性实施例的主机通道适配器的软件模型的示例图；

图 4 是依照一个说明性实施例说明了用于 SAN 上的节点的软件管理模型的示例图；

图 5 是依照一个说明性实施例说明了工作和完成队列处理的示例框图；

图 6 是根据一个说明性实施例说明了三种处理器数据报通信服务的示例图；

图 7 是说明了具有用于指示 Q\_Key 是否与特权应用相关联的控制比特的 Q\_Key 的示例图；

图 8 是依照一个说明性实施例用于处理 CM REQ 消息的示例框图；

图 9 是依照一个说明性实施例说明了 CM REQ 消息的结构化专用数据区的示例框图；

图 10 是依照一个说明性实施例概括了连接建立请求的主动方的示例操作的流程图；以及

图 11 是依照一个说明性实施例概括了连接建立请求的被动方的示例操作的流程图。

## 具体实施方式

文中的说明性实施例提供了这样的机制，其用于确保系统区域网中连接建立请求的源是特权源，以及用于限制可以在该系统区域网中使用的连接建立请求中所提供的信息的类型。确保这样的请求来自特权源涉及在请求的报头中提供标识该请求是否始于特权源的标识符。该标识符仅可由特权应用访问，从而使得非特权应用不可以修改该标识符的设置。以这样的方式，仅特权应用可以将自己标识为具有特权，并且能够建立通信连接。因此，连接建立请求的接受方可以信任该请求中的地址信息和其它信息。

在一个说明性实施例中，该系统区域网是提供基于 InfiniBand™ 的网际协议 (IP) (IPoIB) 功能性的 InfiniBand™ 网络。在这样的系统中，CM REQ 消息用于通过队列对在消费者之间建立通信连接。在这样的实施例中，可以将标识符提供为 Q\_Key 中的比特（例如高阶比特），其中，提供该 Q\_Key 作为该 CM REQ 消息的部分。当设置该 Q\_Key 时，CM REQ 消息的接收方可以信任 CM REQ 消息始于特权消费者，并且因而可以信任该 CM REQ 消息的专用数据区中提供的信息。

虽然将具体参照其中系统区域网 (SAN) 是 InfiniBand™ 网络的示例性实施例来描述关于 SAN 的说明性实施例，但是这些实施例仅仅是说明性的，并且并不限于可以在其中实现说明性实施例的机制的网络的类型。为

了描述说明性实施例，假设本领域的普通技术人员熟悉通常可从 InfiniBand™ 贸易协会 (IBTA) 获得的 InfiniBand™ 体系结构规范。因而，在此并不提供对 InfiniBand™ 体系结构规范的详细讨论。

现参照附图并且特别参照图 1，其说明了可以在其中实现说明性实施例的示例性方面的分布式计算机系统的示例性实施例。提供图 1 中所表示的分布式计算机系统 100 仅出于说明的目的，并且下面描述的说明性实施例可以在众多其它类型和配置的计算机系统上实现。举例来说，实现说明性实施例的计算机系统可以从具有一个处理器和几个输入/输出 (I/O) 适配器的小型服务器到具有成百上千个处理器和数千个 I/O 适配器的特大型并行超级计算机系统。此外，可以在由因特网或内联网连接的远程计算机系统的基础设施中实现说明性实施例。

如图 1 中所示，分布式计算机系统 100 包括系统区域网 (SAN) 113，其是分布式计算机系统内的高带宽、低时延网络互连节点。分布式计算机系统 100 中可以包括超过一个的 SAN 113，并且每个 SAN 113 可以包括多个子网络 (子网)。

文中将节点定义成依附于网络的一个或多个链路的任何组件。在所说明的分布式计算机系统中，节点包括主处理器 101、独立磁盘冗余阵列 (RAID) 子系统 103、I/O 适配器 105、交换机 109A-109C、路由器 111 等。图 1 中所说明的节点仅出于说明的目的，因为 SAN 113 可以连接任何数目和任何类型的独立节点。节点中的任何一个均可以充当端节点，文中将其定义为分布式计算机系统 100 中发起或最终消耗消息或帧的设备。

SAN 113 是在分布式计算机系统 100 内支持 I/O 和处理器间通信 (IPC) 这二者的通信和管理基础设施。如图 1 中所说明的，分布式计算机系统 100 包括允许很多设备在安全、远程受管环境中利用高带宽和低时延并行传送数据的交换式通信结构 (即链路、交换机和路由器)。端节点可以通过多个端口进行通信，并且通过 SAN 113 利用多条路径。对于容错和增加带宽的数据传送，可以采用通过 SAN 113 的路径和多个端口的可用性。

SAN 113 包括交换机 109A-109C 和路由器 111。交换机 109A-109C 将

多个链路连接在一起，并且允许使用小报头目的地局部标识符（DLID，Destination Local Identifier）字段在 SAN 113 内将分组从一个链路路由至另一链路。路由器 111 能够使用大报头目的地全局唯一标识符（DGUID，Destination Globally Unique Identifier）将帧从第一子网中的一个链路路由至第二子网中的另一链路。路由器 111 可以通过广域网（WAN）、局域网（LAN）等连接耦合于其它主机和/或其它路由器。

在 SAN 113 中，主处理器节点 101 和 I/O 节点 106 包括至少一个通道适配器（CA）以便与 SAN 113 连接。主处理器节点 101 包括中央处理器（CPU）119 和存储器 121。在一个实施例中，每个 CA 是向 SAN 113 上所传输的源或接收分组足够详细地实现 CA 接口（例如在以上参照的 InfiniBand™ 体系结构规范中所提供的）的端节点。如所说明的，存在两种 CA 类型，主 CA（HCA）117 和目标 CA（TCA）127。通过通用计算节点来使用 HCA 117 以便访问 SAN 113。在一种实现中，以硬件实现 HCA 117。在 HCA 117 的硬件实现中，HCA 硬件卸载许多 CPU 和 I/O 适配器通信开销。HCA 117 的硬件实现还准许通过交换式网络的多个并行通信，而无需与通信协议关联的常规开销。在 SAN 113 中使用 HCA 117 还向分布式计算机系统 100 的输入/输出（I/O）和处理器间通信（IPC）消费者提供了零处理器复制数据传送，而不涉及操作系统核心过程。HCA 117 和 SAN 113 的其它硬件提供了可靠、容错的通信。

I/O 底盘（chassis）106 包括 I/O 适配器底板和含有适配卡的多个 I/O 适配器节点 105。图 1 中所说明的示例性适配卡包括 SCSI 适配卡 123A、到光纤通道集线器和 FC-AL 设备的适配卡 123B、以太网适配卡 123C、图形适配卡 123D，以及视频适配卡 123E。在不背离本发明的精神和范围的情况下，利用图 1 中所示的机制可以实现任何已知类型的适配卡。I/O 底盘 106 在 I/O 适配器底板中还包括交换机 109B，以便将适配卡 123A-123E 耦合于 SAN 113。

RAID 子系统 103 包括微处理器 125、存储器 126、目标通道适配器（TCA）127，以及多个冗余和/或条式存储盘 129。

在所说明的 SAN 113 中，每个链路 115 均是在任何两个网络元件（例如端节点、交换机 109A-109C 或路由器 111）之间的全双工通道。合适的链路 115 可以包括但不限于：铜缆、光缆，以及底板和印刷电路板上的印刷电路铜迹线（trace）。链路 115 和交换机 109A-109C 等的组合进行操作以便提供 SAN 113 的节点之间的点到点通信。

图 2 中总体说明了示例性主处理器节点 101 的软件和硬件方面。主处理器节点 101 包括执行一组消费者过程 201 的一个或多个处理器。主处理器节点 101 包括具有端口 205 的 HCA 117。每个端口 205 均连接至 SAN 113 的链路 115。端口 205 可以连接至一个 SAN 子网或多个 SAN 子网。利用消息和数据服务 203，消费者过程 201 通过动词接口（verbs interface）207 将消息传送至 SAN 113。动词接口 207 通常是利用操作系统专用程序接口来实现的。利用 SAN 113 的 InfiniBand<sup>TM</sup> 实现，在先前所参照的 InfiniBand<sup>TM</sup> 体系结构规范中指定动词接口 207。

图 3 中说明了 HCA 117 的软件模型。HCA 117 包括一组队列对（QPs）301，其通过端口 205 将消息传送至子网。单个 HCA 117 可以支持数千个 QPs 301。相比较而言，I/O 适配器中的 TCA 127 通常支持少得多数目的 QPs 301。还说明了子网经营管理（SMA，subnet management administration）模型 209、管理分组 211，以及多个虚通道（virtual lane）213，其连接传输层与端口 205。

现转至图 4，其说明了用于 SAN 113 上的节点的软件管理模型。SAN 体系结构管理设施提供了子网管理器（SM，Subnet Manager）303A、子网管理（SA，Subnet Administration）模块 303B，以及支持多个通用管理服务的基础设施。管理基础设施包括在每个节点中操作的子网管理代理（SMA，Subnet Management Agent）307。管理基础设施定义了允许附加通用服务代理的通用服务接口。此外，SAN 体系结构定义了用于在管理器与管理代理之间通信的公共管理数据报（MAD，management datagram）消息结构。

SM 303A 负责初始化、配置和管理交换机、路由器和通道适配器。可

以在诸如通道适配器或交换机的其它设备内实现 SM 303A。将 SAN 113 的一个 SM 303A 专用为主 SM，并且其负责发现子网拓扑，利用各种局部标识（LID）号、全局标识（GID）号、子网前缀和分区键（P\_Keys）来配置每个通道适配器端口；利用 LID、子网前缀及其转发数据库来配置每个交换机，以及为子网维护端节点和服务数据库，以便向 LID/GID 解决服务以及服务目录提供全局唯一标识（GUID）号。因而，利用子网管理（SM）303A 和子网管理（SA）模块 303B 来完成 SAN 113 和 SAN 组件（例如 HCA 117、TCA（或端节点）127、交换机 109A-109C，以及路由器 111）的管理。子网管理分组（SMPs）用于通过端节点 305 的管理代理 307 来发现、初始化、配置和维护 SAN 组件。SAN SA 分组由 SAN 组件使用，以便查询和更新子网管理数据。通过基于主机的端节点 309 中的用户管理控制台 311 提供对子网管理的一些方面的控制。

SAN 113 提供了 I/O 所要求的高带宽和可扩缩性，并且还支持处理器间通信（IPC）所要求的极低的时延和极低的 CPU 开销。用户过程可以绕过操作系统（OS）核心过程，并且直接访问诸如 HCA 117 的网络通信硬件，其启用了高效消息传递协议。SAN 113 适于当前的计算模型，并且是用于新形式的 I/O 和计算机群集通信的构造块。SAN 113 允许 I/O 适配器节点 105 在其自身之间通信或者与分布式计算机系统中的任何或全部的处理器节点 101 进行通信。在 I/O 适配器附于 SAN 113 的情况下，所得到的 I/O 适配器节点 105 具有大体上与分布式计算机系统中的任何处理器节点 101 相同的通信能力。

对于可靠的消息服务类型，诸如主处理器节点 101 和 I/O 适配器节点 105 这样的端节点生成请求分组并且接收确认分组。交换机 109A-109C 和路由器 111 将分组从源传递至目标（或目的地）。除了不同的 CRC 报尾字段之外（在网络中的每个传送阶段均对其进行更新），交换机 109A-109C 向前传递未修改的分组。在路由分组时，路由器 111 更新不同的 CRC 报尾字段并且修改报头中的其它字段。

在 SAN 113 中，硬件提供了一种消息传递机制，其可以用于在通用计

算节点之间的处理器间通信 (IPC) 和输入/输出 (I/O) 设备。消费者通过分别将发送/接收消息置于 SAN 通道适配器 (CA) 上的发送/接收工作队列 (WQ) 来访问 SAN 113 消息传递硬件。

文中将消息定义成数据交换的应用定义单元，其是协作过程之间通信的基本单元。文中将分组 (或帧) 定义成由组网协议报头 (和报尾) 封装的一个数据单元。报头通常提供用于通过 SAN 113 而引导分组 (或帧) 的控制和路由信息。报尾通常含有用于确保在破坏内容的情况下不传递帧的控制和循环冗余校验 (CRC) 数据。

消费者使用 SAN 动词来访问 HCA 功能。解释动词并直接访问 CA 的软件被称为通道接口 (CI)。将发送/接收工作队列 (WQ) 分派给消费者作为队列对 (QP)。可以通过五个不同的传输类型来发送消息：可靠连接 (RC)、可靠数据报 (RD)、不可靠连接 (UC)、不可靠数据报 (UD)，以及原始数据报 (RawD)。消费者通过 SAN 发送和接收工作完成 (WC) 从完成队列 (CQ) 检索这些消息的结果。源 CA 关注分段出站消息并且将其发送至目的地。目的地或目标 CA 关注重新装配入站消息并将其放入由目的地消费者指定的存储空间。下面的附图中说明了这些特征。

现参照图 5，其说明了工作和完成队列处理的框图。每个 QP 301 提供了到发送工作队列 (SWQ) 407 和接收工作队列 (RWQ) 409 的输入。SWQ 407 发送通道和存储语义消息，并且 RWQ 409 接收通道语义消息。消费者调用动词(在动词接口 207 内)，以便将工作请求 (WR) 放到工作队列 (WQ) 中。发送 WR 403 是通道语义操作，以便将一组本地数据段 417 推到 (push) 远程节点的接收 WQE 405 所引用的数据段。发送 WR 的数据段 417 中的每一个均含有虚拟相连存储区域。用于引用本地数据段 417 的虚拟地址处于创建了本地 QP 301 的过程的地址上下文中。

如图 5 中所示，已经由消费者过程 401 放到 WQ 上的 WR 403 被称为工作队列元素 (WQE) 405。WQE 405 由 HCA 117 中的硬件 415 来执行。SWQ 407 含有描述将在 SAN 结构上传输的数据的 WQE 405。RWQ 409 含有描述在哪里放置从 SAN 113 接收到的输入通道语义数据的 WQE 405。

在一个实施例中，RWQ 409 仅支持一种类型的 WQE 405，其被称为接收 WQE。接收 WQE 提供了对于向其中写入输入发送消息（incoming send message）的本地存储空间进行了描述的通道语义操作。接收 WQE 包括描述了若干虚拟相连存储空间的分散列表（scatter list）。将输入发送消息写入这些存储空间。虚拟地址处于创建了本地 QP 301 的过程的地址上下文中。

动词接口 207 还提供了一种用于从完成队列 411 中检索完成的工作的机制。完成队列 411 含有完成队列元素（CQE）413，其含有关于先前完成的 WQE 405 的信息。采用完成队列 411 来为多个 QP 301 创建单点完成通知（a single point of completion notification）。CQE 413 含有足够的信息来确定完成的特定 WQE 405 以及 QP 301。完成队列上下文（未示出）是含有指向长度以及管理各个完成队列 411 所需要的其它信息的指针的信息块。

通过 SAN 结构在使用数据报型消息的过程之间共享的队列需要保护键（protection key）来验证请求方使用在收端接收到的队列的权力。在 SAN 113 内所利用的保护键之一被称为队列键（Q\_Key）。Q\_Key 机制允许应用认证其利用特定的通信资源的权利力，例如发送和接收队列。为了促进应用的认证（即，使用所接收的队列），Q\_Keys 通常能够由应用来设置。因为应用能够设置 Q\_Key，因此需要一种更强的认证，即该认证不能够被未经授权访问通信资源的应用伪造。

在 SAN 113 中，OS 运行为特权类程序，并且应用运行为非特权类。应用要求 OS 进行具有特权的特定操作，例如 QP 上下文建立。提供了 Q\_Key，其是 OS 可控的，并且在没有验证应用具有使用队列资源的权限的情况下，防止从应用过程级访问队列资源。通过利用大得足以使应用过程很难猜测正确的键的键来控制未授权的访问。为队列生成“受控 Q\_Key”，但并不能从应用过程级对其进行操纵，除非操作系统（OS）给予应用这样做的权限。

现参照图 6，其说明了三种处理器数据报通信服务。三个处理器，处

理器 501、处理器 502 和处理器 503，通过各种过程的发送/接收 (QP) 消息相互通信。说明了四个这样的过程，举例来说，包括处理器 502 上的过程 C 和过程 D 以及处理器 503 上的过程 E。在操作期间，远程过程可以同时尝试通过使用数据报型消息与过程 A 进行通信。数据报型消息包括允许在端节点共享 QPs 的特性。

如图 6 中所说明的，过程 C、D 和 E 试图与过程 A 的 QP\_4 进行通信。每个数据报消息均含有 Q\_Key 作为消息的一部分。将 Q\_Key 与关联于 QP\_4 的 Q\_Key 进行比较。如果匹配，则将请求放入 QP\_4 的接收队列 511 中，否则将其默默丢弃（例如，在不可靠数据报型服务的情况下）或将否定确认 (NAK) 型消息发送回发送方（例如，在可靠数据报型服务的情况下）。通常，请求访问 QP 的应用过程可以在 Q\_Key 处猜测多次，直到猜到正确的 Q\_Key。为了防止一个或多个过程 C、D 或 E 获得对 QP\_4 的未授权访问，需要不能被排除的过程 (excluded process) 猜测得到的 Q\_Key。通过提供不能被请求过程生成的 Q\_Key，防止过程 C、D 或 E 中的一个或多个可能成功尝试对 Q\_Key 的正确猜测。

在 Q\_Key 中提供了附加比特，并且将该附加比特指定为特权比特（或控制比特）。附加比特通过允许对 Q\_Key 设置应用级访问限制来增强 Q\_Key 功能性。图 7 说明了具有控制比特 703 的 Q\_Key 701（例如，字符序列）。因而，在 SAN 操作期间，禁止应用级代码在发送工作请求或修改 QP 请求上生成 Q\_Key，除非可信的 HCA 代码首先将应用标识为具有使用受控 Q\_Key 的权限。

在优选实施例中，受控 Q\_Key 是具有为 OS 保留（即，仅可以由 OS 来改变值）的专用最高阶比特或附加比特的 Q\_Key。因此，创建了两类 Q\_Key，受控类和不受控类。对于受控类将控制比特设置为 1，而对于不受控类将控制比特设置为 0。仅准许 OS 以及由 OS 给予特权的消费者过程在工作请求 (WR) 中提交作为受控 Q\_Key 的 Q\_Key。其它用户空间消费者仅可以提交具有未设置的高阶比特的 Q\_Key 的 WR。这防止了用户空间消费者向特权模式 QP 发送消息，因为被动态将校验高阶比特（其不能由

用户空间消费者调节)并且只是成功处理了设置了该比特的消息。

说明性实施例的机制利用受控 Q\_Key 来检验连接请求始于授权或特权消费者。以这样的方式,在维持类似于 TCP 的安全级别的安全级别时,有可能通过 SAN 进行网际协议 (IP) 通信。该安全级别确保连接建立的被动方可以信任连接请求的发送方。

将 InfiniBand<sup>TM</sup>作为示例性 SAN,在其中实现了说明性实施例的机制,通过使用不可靠数据报 (UD) 队列对 (QP),由 InfiniBand<sup>TM</sup> 结构来传递通信管理 (CM) 消息。这样的 CM 消息,即 UD 分组,是使用远程直接存储器访问 (RDMA) 操作来传递的,在 RDMA 操作中,直接将数据从一个存储器传递到另一存储器,而不涉及主节点的处理器。在 UD 分组中使用数据报扩展传输层报头 (DETH)。DETH 尤其含有 UD 分组所导向的队列对的目的地 Q\_Key。该 Q\_Key 具有高阶比特,可以将其设置来指定 UD 分组的始发方是否具有特权。

举例来说,当消费者生成工作请求 (WR) 时,该消费者指定包括在 WR 中的 Q\_Key。如果消费者是特权消费者,那么由该消费者设置 WR 中的 Q\_Key 的高阶比特。不是特权消费者的消费者不能够设置该 Q\_Key 的高阶比特。因而,与非特权消费者关联的 Q\_Key 未被设置并且指示 WR 的源是非特权消费者。

通道接口 (CI) 检查 WR 中的该 Q\_Key,并且基于高阶比特的设置,确定输出分组的 DETH 是否含有来自与消费者关联的 QP 的 Q\_Key,或者来自工作请求 (WR) 的 Q\_Key。再者,具有最高比特设置的 Q\_Key 被认为是受控 Q\_Key,并且通道适配器不允许消费者任意指定受控 Q\_Key。OS 维持对受控 Q\_Key 的控制,因为其可以仅为特权消费者配置受控 Q\_Key 的 QP 上下文。这允许特权模式代码实现这样的策略,即仅向用户空间消费者提供具有未设置的高阶比特的 Q\_Key。

因而,依照说明性实施例的机制,举例来说,对于 CM REQ 消息,CM REQ 中的 DETH 中的 Q\_Key 的高阶比特通知通道接口 (CI)(其工作是解释在通过通道适配器的通信中使用的动词) CM REQ 消息是否始于特

权消费者。利用说明性实施例，只有特权消费者，即 OS 已经给予特权状态的应用或主节点的 OS，可以使用 CM REQ 消息建立通信连接。如果 Q\_Key 的高阶比特指示特权消费者是 CM REQ 消息的源，那么该通信连接的被动方可以信任在该 CM REQ 消息的专用数据区中提供的信息。如果 Q\_Key 的高阶比特指示非特权消费者是 CM REQ 消息的源，那么被动方可以不信任在专用数据区中提供的信息，即该信息可能是用户空间应用生成的，并且因而可能是欺骗信息。因此，可以拒绝该 CM REQ 消息。

图 8 是依照一个说明性实施例用于处理 CM REQ 消息的示例框图。如图 8 中所示，主节点 810 包括在其上运行的多个消费者过程 812-816。消费者过程 812 希望在主节点 880 上与过程 884 建立 TCP/IP 通信连接。因此，消费者过程 812 通过通道接口 830 将 CM REQ 工作请求递送到其在与主机通道适配器 840 相关联的本地 QP 822 中的发送队列。作为该 CM REQ 工作请求的一部分，消费者过程 812 提供 Q\_Key，并且设置该 CM REQ 工作请求的 Q\_Key 的高阶比特。另外，消费者过程 812 利用新的服务标识符来限制 CM REQ 工作请求，其中该新的服务标识符具有格式化专用数据区，如下文将较为详细讨论的。

在通道接口 830 中从消费者过程 812 接收 CM REQ 工作请求。通道接口 830 检查该 CM REQ 工作请求，并且确定在该 CM REQ 工作请求中提供的 Q\_Key 是否指示该 CM REQ 工作请求始于特权消费者，例如 OS 或已经由 OS 授予特权状态的过程。举例来说，通道接口 830 可以检查 Q\_Key 中高阶比特的状态，以便确定是否设置了高阶比特。如果设置了高阶比特，那么通道接口 830 可以确定消费者过程 812 是特权消费者过程。如果没有设置高阶比特，那么通道接口 830 可以确定消费者过程 812 是非特权消费者过程。

由于 OS 控制谁可以使用特权 Q\_Key，因此唯一可以设置 Q\_Key 中高阶比特的时间是当消费者过程 812 是特权消费者过程并且在该 Q\_Key 中具体设置了高阶比特的时候。否则，将不设置高阶比特，其对于非特权消费者过程是缺省的。

如果没有设置工作请求中 Q\_Key 的高阶比特，则通道接口 830 指示主机通道适配器 840 的本地 QP 822 将 CM REQ 工作请求中所提供的 Q\_Key 嵌入到由本地 QP 822 所发出的 CM REQ 消息的 DETH 中。如果设置了 CM REQ 工作请求中 Q\_Key 的高阶比特，则通道接口 830 指示本地 QP 822 改为嵌入其自己的 Q\_Key。

使用这些机制，Q\_Key 本身便不能由于体系结构而被欺骗。也就是说，由于 OS 控制 QP 创建并且用户级应用不能改变 QP 上下文，因此 OS 具有控制使用特权 Q\_Key 的装置。如果非特权用户级应用设置了工作请求中 Q\_Key 的高阶比特，那么本地 QP 将嵌入其自己的 Q\_Key 而不是工作请求中所提供的 Q\_Key。对于非特权用户级应用来说，QP 上下文中的 Q\_Key 是非特权 Q\_Key。

当目标主节点 880 接收到 CM REQ 消息时，其使用该 CM REQ 消息的 DETH 中的 Q\_Key 来验证输入的 CM REQ 消息。目标主节点 880 的通道接口 882 检查该 Q\_Key，并且确定作为该 CM REQ 消息的目标的队列对是否是队列对 1(QP1)。QP1 是 InfiniBand™ 体系结构中的特别队列对，其被分派给 OS，并且因而处理特权或受信通信。如果 CM REQ 消息指向 QP1，则通道接口 830 确定该 Q\_Key 是否是与 QP1 相关联的指定 Q\_Key，例如，0x80010000。如果 CM REQ 消息指向 QP1 并且该 Q\_Key 是指定 Q\_Key，那么准许继续由目标主节点 880 处理该 CM REQ 消息。如果 CM REQ 消息指向 QP1，但 Q\_Key 不是与 QP1 相关联的指定 Q\_Key，那么可以将拒绝响应返回给发起 CM REQ 消息的主节点 810。

如果 CM REQ 消息的目标 QP 不是 QP1，那么目标主节点 880 的通道接口 882 验证 CM REQ 消息的 DETH 具有设置了高阶比特的 Q\_Key，其中该高阶比特在 CM REQ 消息始于特权模式的情况下可以被设置仅是 QP。如果 CM REQ 消息的 Q\_Key 具有设置了的高阶比特，那么在 CM REQ 消息中的 Q\_Key 与目的地 QP 的 Q\_Key 匹配的情况下准许目标主节点 880 继续处理 CM REQ 消息。否则，如果 Q\_Key 并不具有设置了的高阶比特或者 CM REQ 消息中的 Q\_Key 并不与目的地 QP 的 Q\_Key 相匹配，

则可以将拒绝响应消息返回给发起方主节点 810。

因而，上述机制确保 CM REQ 消息是发自于特权消费者过程的，而不是由可能进行欺骗的用户空间过程发送的。除了这些保护机制之外，说明性实施例还提供了一种机制，通过该机制可以处理 CM REQ 消息的专用数据区，以便获得必要的 TCP/IP 连接建立信息，例如，源 IP 地址、目的 IP 地址等。特别地，提供了可以包括在 CM REQ 消息的 DETH 中的服务标识符，以便指示专用数据区是根据特定规范进行格式化的。

通常，并不结构化诸如 CM REQ 消息的不可靠数据报的专用数据区。因此，过程可以将它们认为符合专用数据区的任何信息放入专用数据区的任何字段。因而，如果专用数据区用于将 TCP/IP 信息传送给目标主系统 880，则无法确切知道在专用数据区的哪里放置了 TCP/IP 信息或者在专用数据区的哪些字段中有什么信息。利用说明性实施例的机制，以预先确定的方式结构化该专用数据区，从而使得将专用数据区的特定字段指定用于存储 TCP/IP 信息的特殊部分。当其被利用的时候，在 CM REQ 消息的报头中指定该结构化专用数据区。

基于对所接收的 CM REQ 消息的 DETH 中该服务标识符的检测，目标主节点 880 的通道接口 882 可以从 CM REQ 消息的专用数据区提取必要的信息，以便通过系统区域网建立 TCP/IP 连接。服务标识符可以是任何类型的服务标识符，其可以被包括在 CM REQ 消息的报头或 DETH 中。

对所接收的 CM REQ 消息的 DETH 中预先确定的服务标识符的检测通知通道接口 882：在 CM REQ 消息中利用了专用数据区的预先确定的结构。因此，通道接口 882 知道专用数据区的哪些字段含有在主节点 810 上的过程 812 与目标主节点 880 上的过程之间建立 TCP/IP 通信连接所需要的 TCP/IP 信息的哪些部分。

图 9 是依照一个说明性实施例说明了 CM REQ 消息的结构化专用数据区的示例框图。图 9 仅仅是 CM REQ 消息的专用数据区的一种可能的结构，且并不旨在陈述或暗示关于可以结构化专用数据区的方式的任何限制。在不背离说明性实施例的精神和范围的情况下，可以对图 9 中所描绘的结构

进行很多修改。

如图 9 中所示，专用数据区 900 包括特定字段 910-942，其用于存储对于在诸如 InfiniBand<sup>TM</sup> 网络的系统区域网中的主节点的过程之间建立 TCP/IP 通信连接可能是必要的特定 TCP/IP 信息。在所描绘的例子中，字段 910 存储主要版本，字段 912 存储次要版本，字段 914 存储 IP 版本，字段 916 存储基于零的虚拟地址 (ZB) 异常值，字段 918 存储无效发送 (SI) 异常值，字段 920 存储连接偏好 (CP) 值，保留字段 922，字段 924 存储源端口标识符，并且保留字段 926。字段 928-934 存储源 IP 地址，且每个字段存储源 IP 地址的不同部分，如所描绘的。字段 936-942 存储目的 IP 地址，且每个字段存储目的 IP 地址的不同部分，如所描绘的。

TCP/IP 通信连接建立的主动方上的特权消费者（即 CM REQ 消息的源）负责设置预定专用数据字段 910-942 中的值，例如源 IP 地址、目的 IP 地址等。当被动方（即目标主节点）接收到 CM REQ 消息时，其首先验证该 CM REQ 消息来自使用先前所描述的方法和机制的特权消费者。然后，通过检查服务标识符，被动方知道专用数据区含有诸如图 9 中所示的那些字段的预定字段。被动方然后可以依照定义的结构解释专用数据区。

图 10 和 11 是依照一个说明性实施例概括了通信连接建立的主动方和被动方的示例性操作的流程图。应当理解，可以通过计算机程序指令实现流程图说明的每个块，以及流程图说明中块的组合。可以将这些计算机程序指令提供给处理器或其它可编程数据处理装置来产生机器，从而使得在处理器或其它可编程数据处理装置上执行的指令创建用于实现流程图块中所指定的功能的装置。还可以将这些计算机程序指令存储在可以指导处理器或其它可编程数据处理装置以特定方式运行的计算机可读存储器或存储介质中，从而使得存储在计算机可读存储器或存储介质中的指令产生包括实现流程图块中所指定的功能的指令装置在内的制品。

因此，流程图说明的块支持用于实现指定功能的装置的组合、用于实现指定功能的步骤以及用于实现指定功能的程序指令装置的组合。还应当理解，可以通过实现指定功能或步骤的基于专用硬件的计算机系统，或者

通过专用硬件和计算机指令的组合，来实现流程图说明的每个块以及流程图说明中的块的组合。

图 10 是依照一个说明性实施例概括了连接建立请求的主动方的示例性操作的流程图。如图 10 中所示，操作开始于通道接口接收对建立 TCP/IP 通信连接的工作请求(步骤 1010)。通道接口检查该工作请求的 Q\_Key(步骤 1020)，并且确定是否设置了 Q\_Key 的高阶比特(步骤 1030)。如果没有设置高阶比特，那么通道接口指示本地队列对将工作请求中所提供的 Q\_Key 嵌入其发送出的 CM REQ 消息(步骤 1040)。如果设置了高阶比特，那么通道接口指示本地队列对嵌入与本地队列对相关联的 Q\_Key (步骤 1050)。然后该操作结束。

图 11 是依照一个说明性实施例概括了连接建立请求的被动方的示例性操作的流程图。如图 11 中所示，操作开始于对 CM REQ 消息的接收(步骤 1110)。通道接口确定 CM REQ 消息是否以队列对 1 为目标(步骤 1120)。如果是的话，则通道接口确定 CM REQ 消息的 Q\_Key 是否是与队列对 1 相关联的预定 Q\_Key (步骤 1130)。如果是的话，则继续进行对 CM REQ 消息的处理，以便建立 TCP/IP 连接(步骤 1140)。例如，可以在 CM REQ 消息上继续进行依照 InfiniBand<sup>TM</sup> 规范的处理，从而建立 TCP/IP 连接。举例来说，在从远程通道适配器的连接管理器(CM)接收到请求时，本地通道适配器的 CM 确定所请求的服务是否在本地通道适配器(CA)上可用。如果不是，则 CM 将拒绝消息发送回远程 CA 的 CM，其陈述拒绝的原因。如果本地 CA 支持所请求的服务，则本地 CM 创建 QP 来处理其通信通道的末端，对具有请求中所提供的信息的 QP 的上下文进行编程，将新创建的 QP 转变成准备接收状态，并且然后将响应消息与关于新创建的 QP 的信息发送回请求方。

返回步骤 1140，如果 CM REQ 消息的 Q\_Key 不是与队列对 1 相关联的预定 Q\_Key，那么可以将拒绝响应消息返回给 CM REQ 消息的发起方或主动方(步骤 1150)。如果 CM REQ 消息并不指向队列对 1，则通道接口确定是否设置了 CM REQ 消息中 Q\_Key 的高阶比特(步骤 1160)。如

果设置了 Q\_Key 的高阶比特，那么通道接口将 CM REQ 消息中的 Q\_Key 与目的地 QP 的 Q\_Key 进行比较，以便确定是否存在匹配（步骤 1165）。如果存在匹配，则继续进行对 CM REQ 消息的处理（步骤 1140）。如果没有设置 Q\_Key 的高阶比特，那么可以将拒绝响应消息发送回给 CM REQ 消息的发起方或主动方（步骤 1150）。

此后，通道接口检查 CM REQ 消息中的服务标识符（步骤 1160），并且确定该服务标识符是否指定在 CM REQ 消息的专用数据区中使用了预定字段（步骤 1170）。如果 CM REQ 消息中的服务标识符指示利用了预定字段，则通道接口依照预定字段来处理专用数据区中的信息（步骤 1180）。否则，如果服务标识符没有指定利用了预定字段，那么 CM 确定服务 ID 所指定的期望服务是否存在于其关联子系统内（步骤 1185）。如果服务存在，那么按照本领域中公知的正常方式继续处理（步骤 1190）。否则，如果服务并不存在于 CM 的关联子系统内，则可以将拒绝响应消息返回给 CM REQ 消息的发起方或主动方（步骤 1195）。然后该操作结束。

因而，利用说明性实施例的机制，将新的服务标识符用于指示 CM REQ 消息专用数据区含有根据预定结构的预定字段。此外，通过限制 CM REQ 消息仅由特权消费者发送，如受控 Q\_Key 所检验的，被动方可以确定含于 CM REQ 消息的专用数据区中的信息并不是由非特权用户空间消费者设置的。这保证了对 CM REQ 消息的预定专用数据区字段的处理是由特权消费者完成的，并且诸如 IP 地址的信息（其由主动方传递给 CM REQ 消息专用数据区中的被动方）是可以信任的。

应当理解，说明性实施例可以采取全硬件实施例、全软件实施例或者既含有硬件元素又含有软件元素的实施例的形式。在一个示例性实施例中，以软件实现说明性实施例的机制，其包括但不限于固件、常驻软件、微码等。

此外，说明性实施例可以采取可访问于计算机可用或计算机可读介质的计算机程序产品的形式，该计算机可用或计算机可读介质提供由计算机或任何指令执行系统使用的或者与计算机或任何指令执行系统结合使用的

---

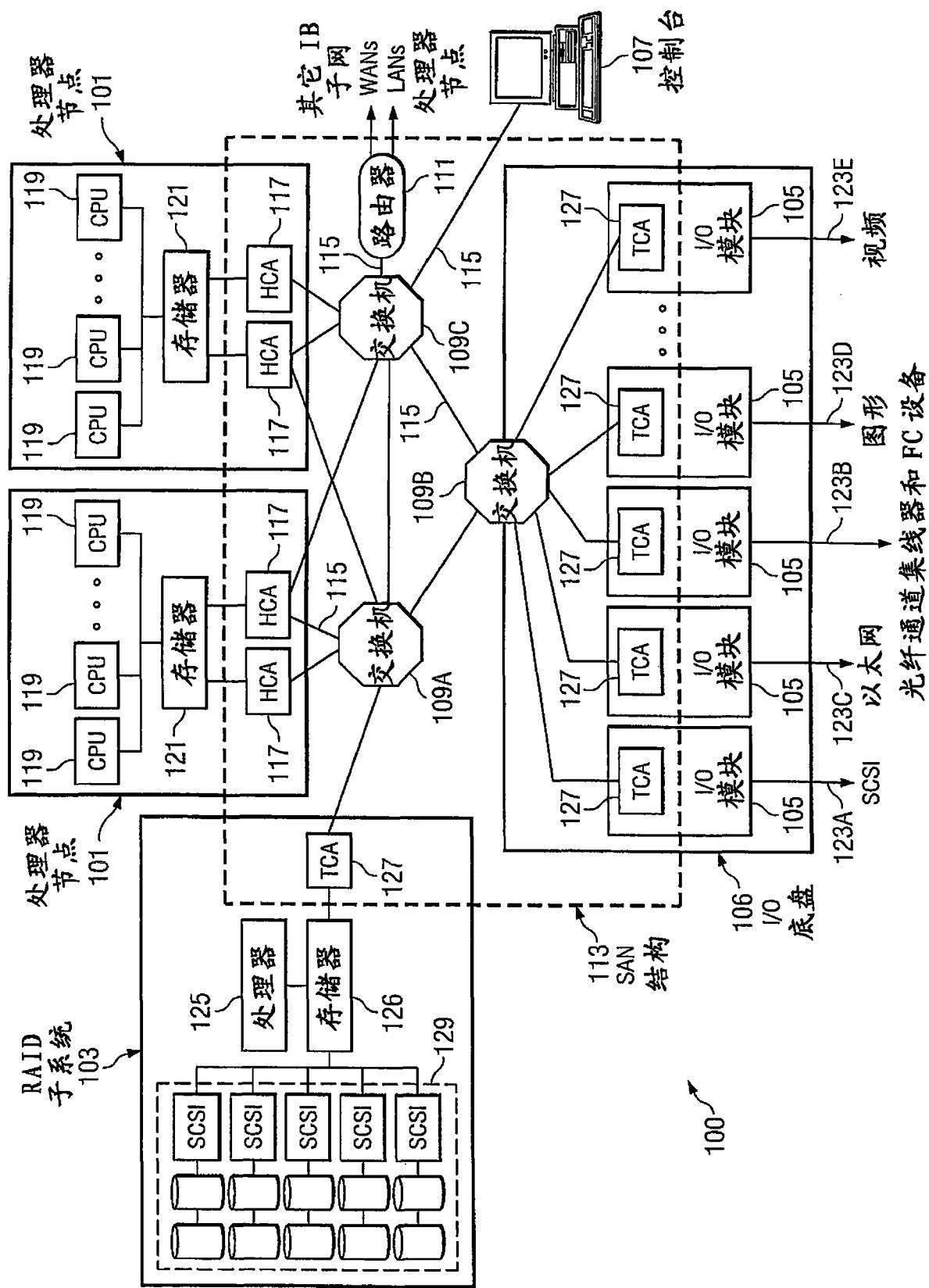
程序代码。对于该描述来说，计算机可用或计算机可读介质可以是能够容纳、存储、通信、传播或传送由指令执行系统、装置或设备使用的或者与指令执行系统、装置或设备结合使用的程序的任何装置。

介质可以是电子、磁性、光学、电磁、红外或半导体系统（或装置或设备）或者传播介质。计算机可读介质的例子包括半导体或固态存储器、磁带、可装卸计算机磁盘、随机访问存储器（RAM）、只读存储器（ROM）、硬磁盘和光盘。光盘的当前的例子包括只读光盘存储器（CD-ROM）、读/写光盘（CD-R/W）和DVD。

适于存储和/或执行程序代码的数据处理系统可以包括通过系统总线直接地或间接地耦合于存储元件的至少一个处理器。存储元件可以包括在程序代码的实际执行期间所采用的局部存储器、大容量存储器，以及为了减少在执行期间必须从大容量存储器检索代码的次数而提供对至少一些程序代码的临时存储的高速缓冲存储器。

输入/输出或 I/O 设备（包括但不限于键盘、显示器、指点设备等）可以直接地或者通过插入 I/O 控制器耦合于系统。网络适配器也可以耦合于系统，从而使得数据处理系统能够适于通过介入专用或公用网络耦合于其它的数据处理系统或远程打印机或存储设备。调制解调器、电缆调制解调器和以太网卡正是几种当前可用类型的网络适配器。

已经出于说明和描述的目的给出了对本发明的描述，且并不旨在以所公开的形式穷举或限制本发明。对本领域的普通技术人员来说，很多修改和变形将是显而易见的。选择和描述实施例是为了最好地解释本发明的原理、实际应用，以及使本领域的普通技术人员能够针对适于预期的特定用途的各种实施例以及各种修改来理解本发明。



1

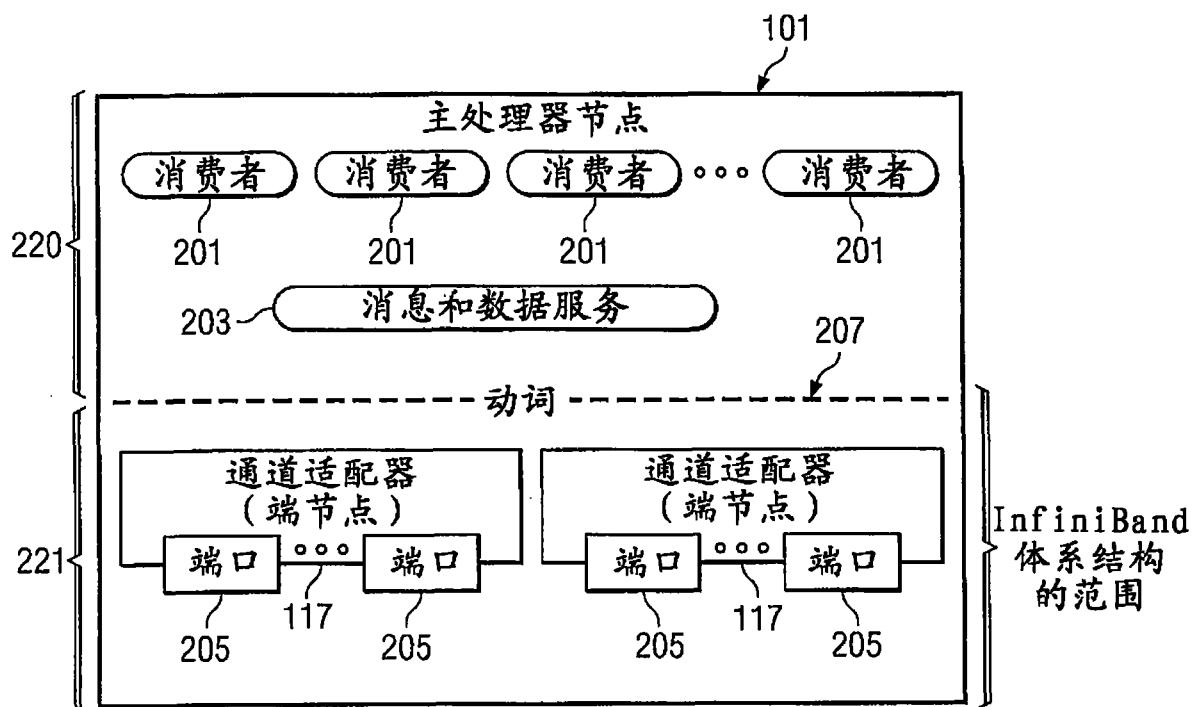


图 2

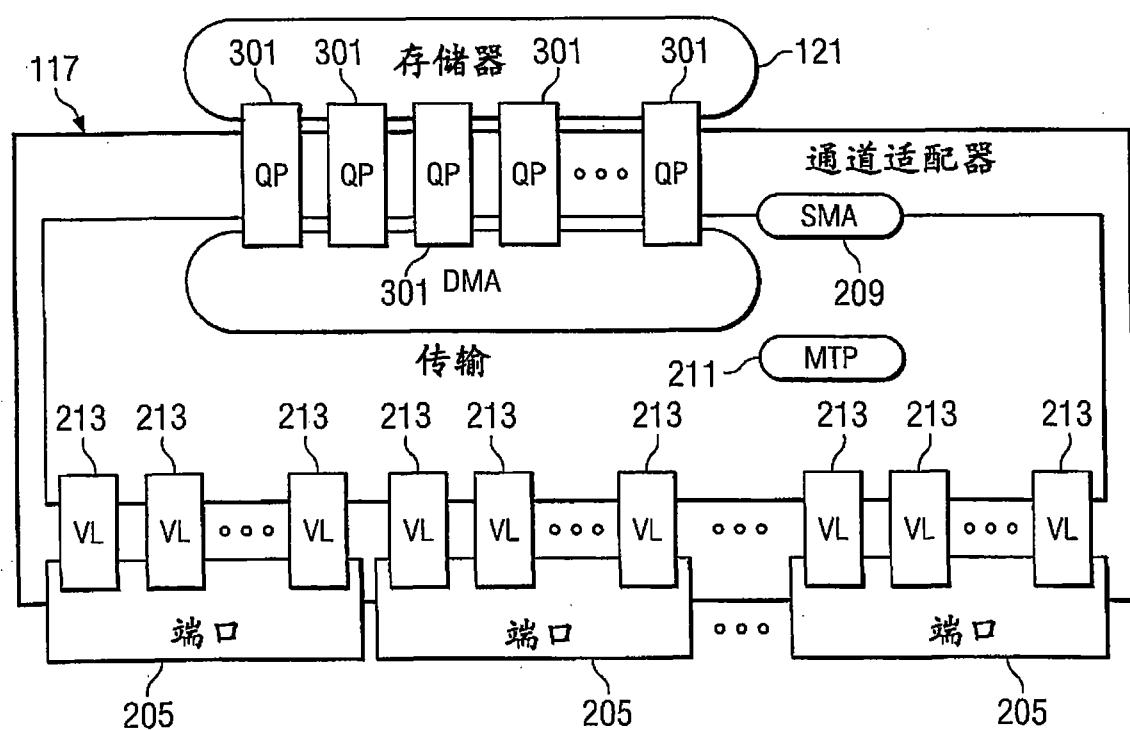
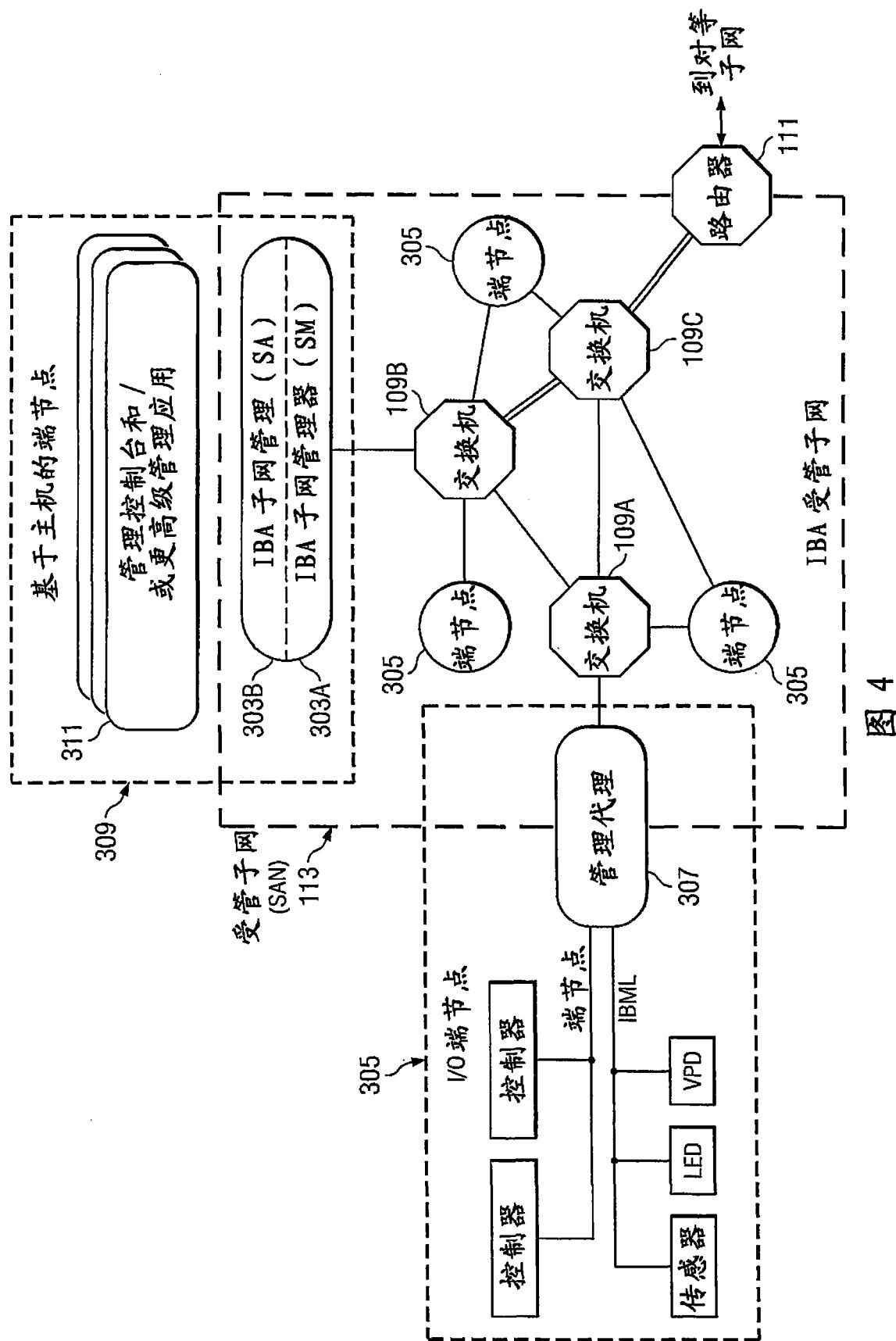


图 3



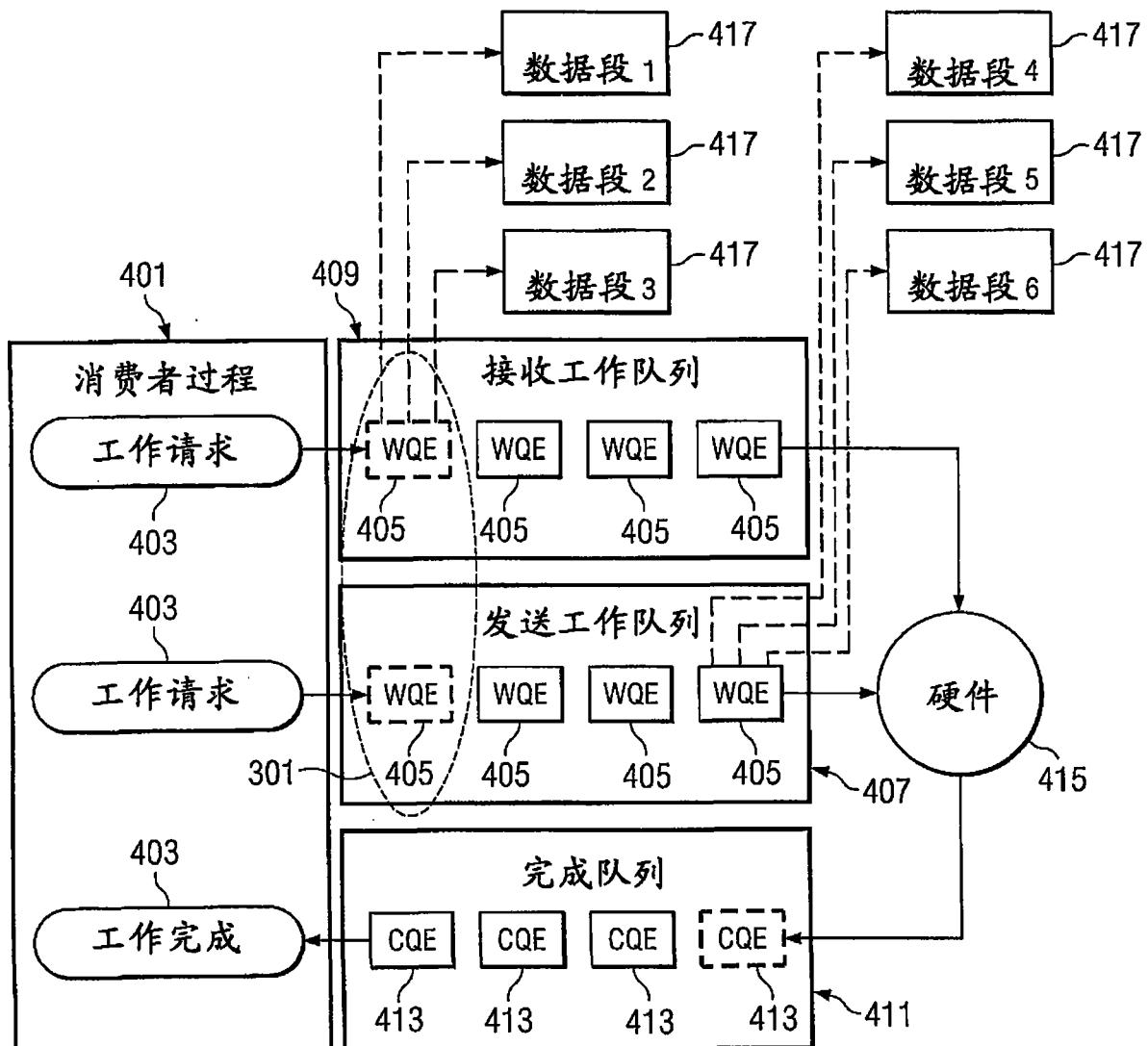


图 5

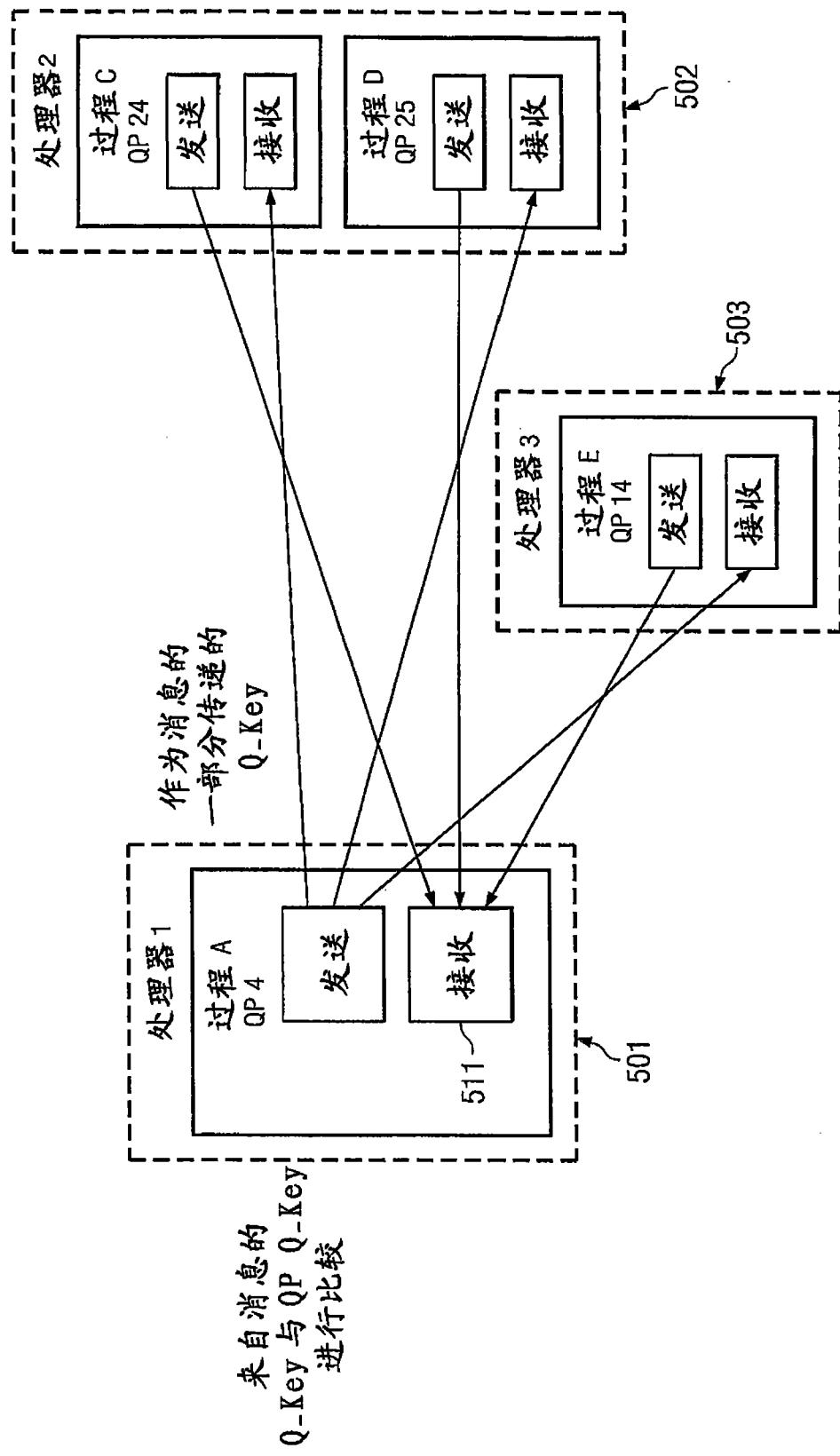
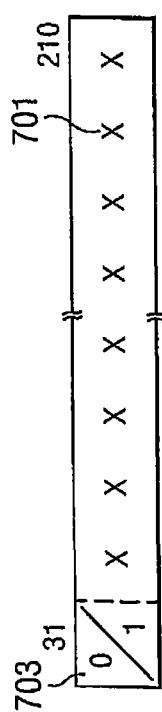


图 6

图 7



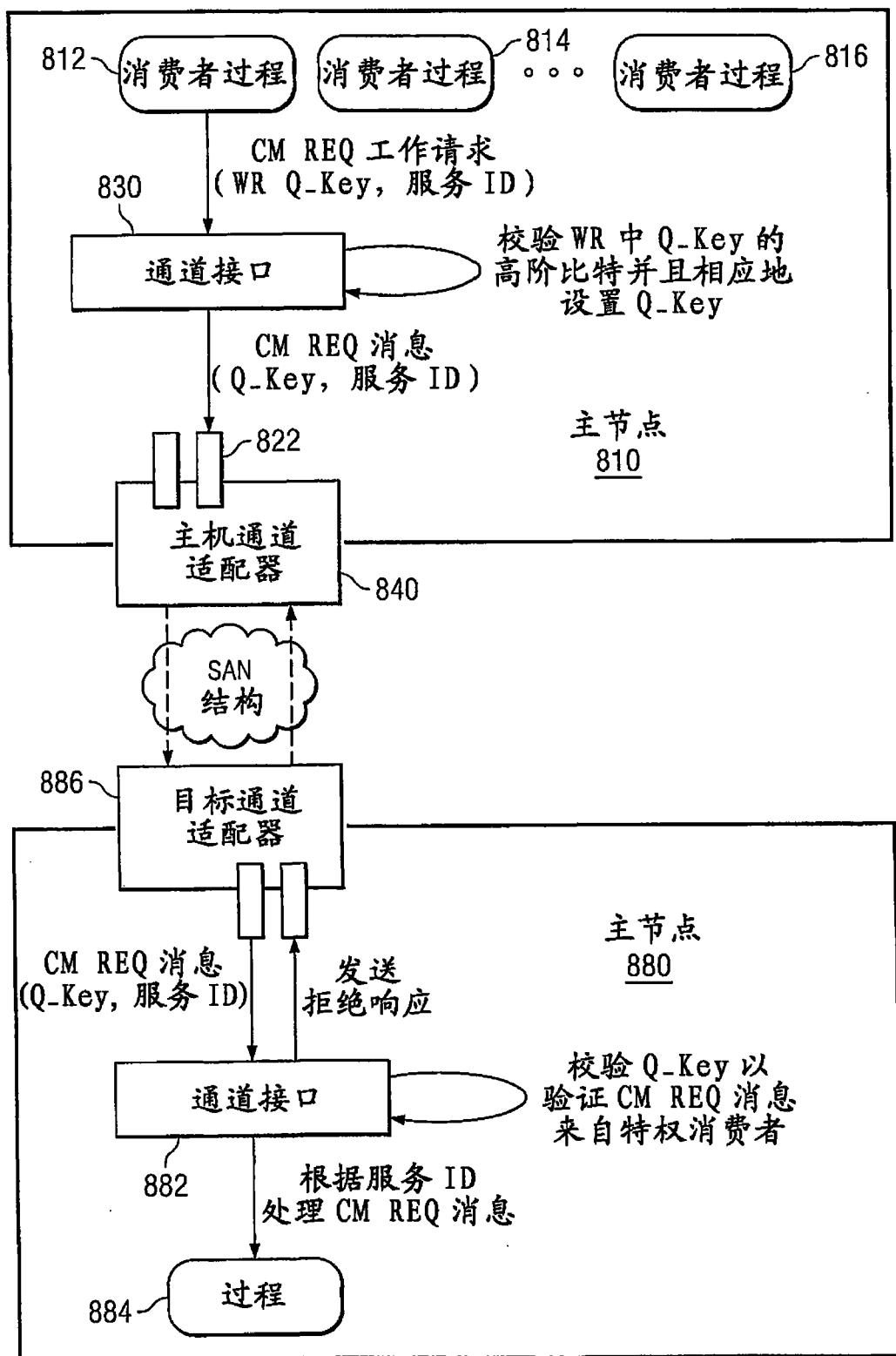


图 8



图 9

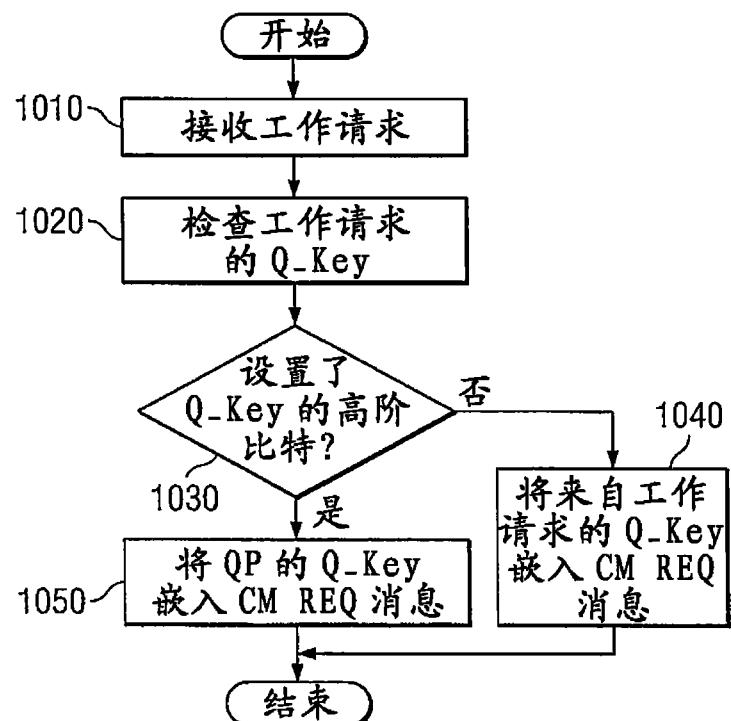


图 10

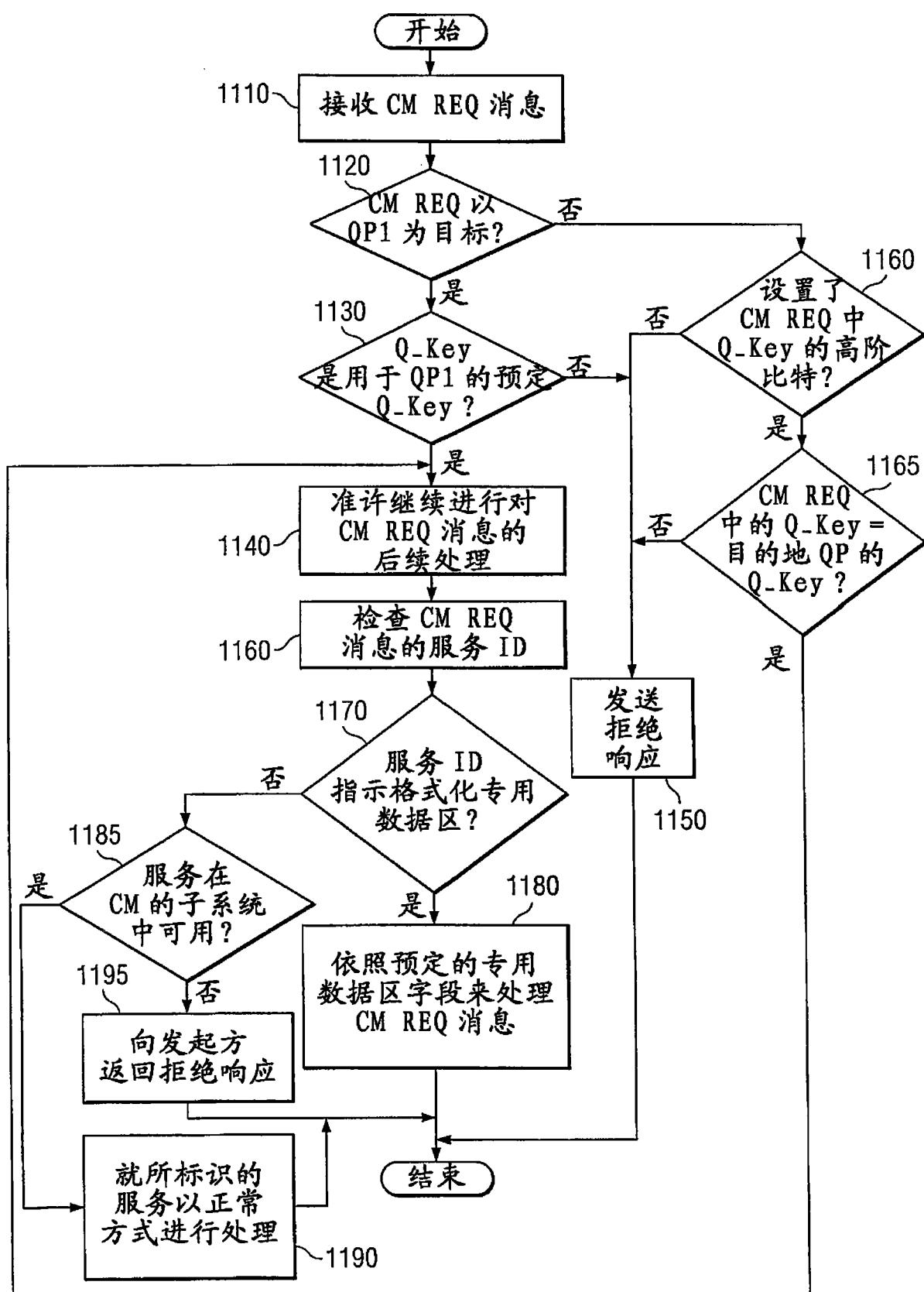


图 11