



[12] 发明专利申请公布说明书

[21] 申请号 200780022101.7

[43] 公开日 2009年6月24日

[11] 公开号 CN 101467405A

[22] 申请日 2007.3.16

[21] 申请号 200780022101.7

[30] 优先权

[32] 2006.7.12 [33] EP [31] 06014496.1

[86] 国际申请 PCT/EP2007/052527 2007.3.16

[87] 国际公布 WO2008/006629 英 2008.1.17

[85] 进入国家阶段日期 2008.12.12

[71] 申请人 西门子公司

地址 德国慕尼黑

[72] 发明人 L·科内

[74] 专利代理机构 中国专利代理(香港)有限公司  
代理人 胡莉莉 刘春元

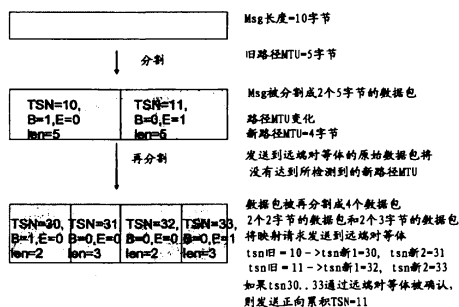
权利要求书1页 说明书3页 附图1页

[54] 发明名称

用于在网络中发送数据包的方法

[57] 摘要

流控制传输协议(SCTP)将数据包发送到远端对等体,其中数据包的大小等于或小于最大传输单元(MTU)。MTU可以通过SCTP利用RFC1191(或者类似技术)的路径MTU发现算法来发现。然而,SCTP关联内的某个路径的MTU可以随着时间的过去而变化。如果路径MTU增长,则不存在问题。如果路径MTU减小,则大小等于旧路径MTU的数据包将不能被递送到远端对等体,因为具有较小MTU的链路或路由器将丢弃该数据包。因此,所有这样的数据包将保持未确认,并且针对将流的数据包按序递送到未被递送的数据包所属的地方,所有这样的数据包将导致线头阻塞。这导致关联的故障。本发明通过由于SCTP关联中的路径MTU减小而再分割SCTP数据包来解决该问题。



1. 一种用于在网络中发送数据包的方法，该方法包括以下步骤：
  - a) 发送连接的数据包，其中数据包的大小等于或小于最大数据包大小 (MTU)，
  - b) 有规律地发现最大数据包大小 (MTU)，
  - c) 由于发现最大数据包大小 (MTU) 减少，再分割该连接的数据包，使得数据包大小与新的最大数据包大小 (MTU) 相匹配。
2. 根据权利要求 1 所述的方法，还包括以下步骤：
  - a) 在某路径上发送连接的数据包，
  - b) 发现每个路径的最大数据包大小 (MTU)。
3. 根据权利要求 1 或 2 所述的方法，还包括以下步骤：
  - a) 给连接的每个数据包分配传输顺序号 (TSN)，
  - b) 将再分割参数朝向连接的远端对等体发送，以指示什么旧传输顺序号被新传输顺序号替换。
4. 根据权利要求 1、2 或 3 所述的方法，还包括以下步骤：  
不发送 MTU 对于目前的路径 MTU 太大的数据包。
5. 根据权利要求 1、2、3 或 4 所述的方法，还包括以下步骤：  
根据 SCTP 发送数据包。

## 用于在网络中发送数据包的方法

### 1. 本发明要解决的问题

1. 类似流控制传输协议 (= SCTP) 的传输协议将数据包发送到远端对等体 (remote peer), 其中数据包的大小等于或小于最大传输单元 (= MTU)。在 SCTP 的情况下, 路径的 MTU 可以利用 RFC1191 (或类似技术) 的路径 MTU 发现算法来发现。然而, SCTP 关联内的某个路径的 MTU 可随着时间的过去而变化。如果路径 MTU (某个路径的 MTU) 在 SCTP 关联期间增长, 则不存在问题。如果路径 MTU 减小, 则大小等于旧路径 MTU 的数据包将不能被递送到远端对等体, 因为具有较小 MTU 的链路或路由器将丢弃该数据包。因此, 所有这样的数据包将保持未确认, 并且针对将流的数据包按序递送到未被递送的数据包所属的地方, 所有这样的数据包将导致线头阻塞 (head of line blocking)。这导致关联的故障。

### 2. 目前为止对该问题的解决方案

到目前为止, 在 SCTP 还未解决该问题。

### 3. 根据本发明的对该问题的解决方案

根据本发明, 由于发现最大数据包大小 (MTU) 减小, 所以传输协议对连接的数据包进行再分割 (在 SCTP 的情况下, 连接被称为关联), 使其大小匹配新的最大数据包大小 (MTU)。

以下, 本发明的 SCTP 实施方案针对用于将消息/数据包递送到应用的所有操作模式来描述 (也参见附图的图)。

根据 SCTP, 每个消息都可以被分割成一个或多个数据包, 并且每个数据包都具有分配给其的传输顺序号 (= TSN)。每个数据包也经由流标识符 (流 id) 被分配给某个流。属于相同流 id 的所有顺序操作的消息必须按次序被递送到对等体应用。SCTP 实施方按应该针对其路径中的每个路径有规律地进行路径 MTU 发现 (这是由 RFC2932 所要求的)。当路径 MTU 的减小已被检测到时, 那么应观察无序的和有序的递送的

后面的动作。检测可以通过 RFC1191 或者类似技术 (RFC1981、draft - ietf - pmtud - method - 06.txt...) 来完成。

- 首先, 新的路径 MTU 必须通过使用路径 MTU 发现来被检测。如果路径 MTU 被减小, 则旧数据包必须被再分割, 使其大小匹配新的路径 MTU (因此, 至少两个新的 TSN 必须被依次分配)。

- 如果针对该消息仅生成单个数据包, 则所要求的再分割可以产生两个或者更多数据包。根据在 RFC2932 中得到的规则, 需要对数据包进行分割。

- 如果原始数据包已经是分割消息的结果, 则根据在 RFC2932 中得到的规则依次针对每个旧 TSN 数据包, 旧数据包必须利用新 TSN 再次被分割。

- 另外, 包含在控制数据包 (SCTP 控制组块 (chunk)) 中的再分割参数被朝向远端对等体发送, 以指示什么旧 TSN 被新 TSN 替换。旧 TSN 将被映射到 2 个或者更多新 TSN, 并且如果要求重编整个消息, 则必须使用在旧 TSN 之前和之后已经接收到的数据包。再分割参数可以被包含在任何已经存在的 SCTP 控制组块或者新的 SCTP 重新映射控制组块中。

- 在新的再分割的数据包被发送到远端对等体之前, 再分割的参数必须被发送到远端对等体。

- 再分割的参数包含旧 TSN 和一个或者多个新 TSN 的列表。这将旧 TSN 映射到新 TSN。再分割的参数可以包含映射列表或者可以在控制组块内或者跨越这些控制组块被重复。

- 通过将具有设置到旧数据包 TSN 的 TSN 的正向累积 (forward cumulative) TSN 组块发送给远端对等体, 使旧数据包 TSN 无效。参见 RFC3758 中的正向累积 TSN 的具体使用。这仅仅可以在使用未完成的 (outstanding) 旧 TSN 和所映射的旧到新 TSN 而已成功地由数据包重编消息之后才完成。

- 在发送再分割参数之后, 具有旧 TSN 的数据包不必被重新发送给远端对等体。

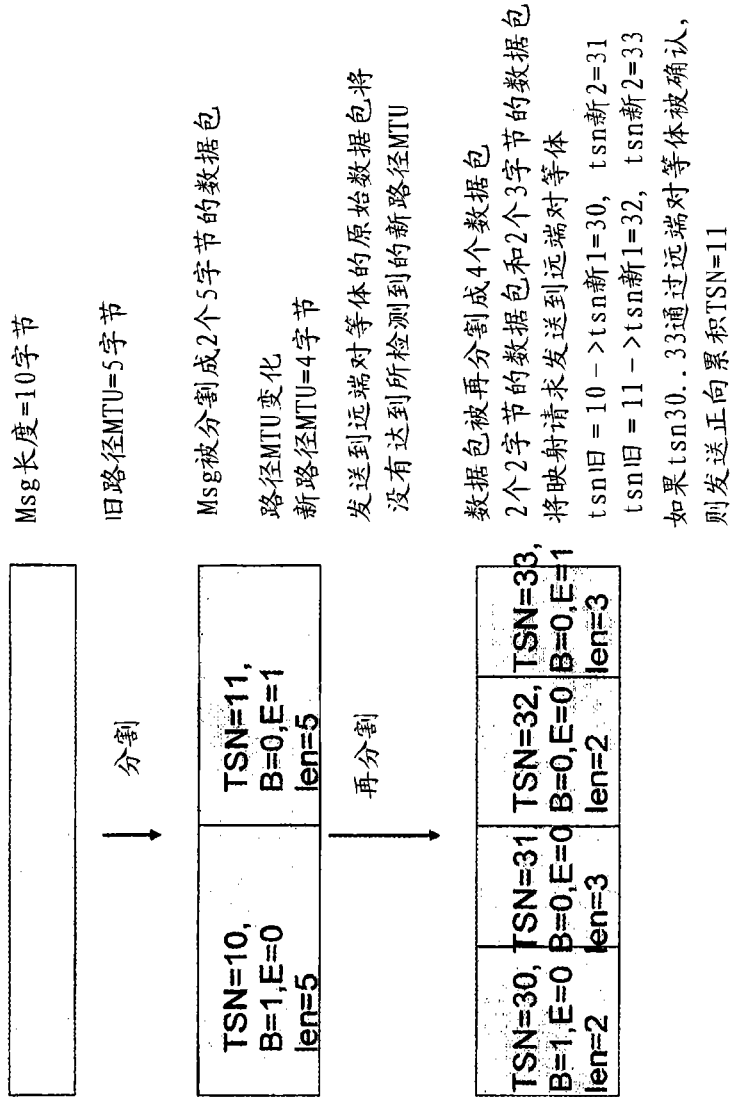
- 在接收到再分割参数之后, 不要求远端对等体发布旧 TSN 的间隙报告 (间隙报告 = 请求发送方对等体再次发送这些数据包, 在这种情况下为旧 TSN)

- 当远端对等体已接收到所有新的 TSN 段并且已重编该消息时，那么该消息必须被递送到该应用。在有序递送的情况下，所递送的消息被继之以包含原始旧 TSN 的消息。在无顺序递送的情况下，具有原始旧 TSN 的消息曾已经被递送到该应用。

- 由于在 SCTP RFC2932 和其他技术中所利用的拥塞算法，再分割之后发送的数据包的数目不准违背关于未完成的数据包的规则。

#### 4. 本发明的效果

如果路径 MTU 在传输期间被减小，则允许 SCTP 不阻塞数据到远端对等体的传输。由于使用视为 TSN 的字节（TCP 简单地分割消息并且使用继之以原始字节计数 TSN 的在中间的字节计数 TSN），所以 TCP 并没有这个问题。（尤其是在非常长的使用期限的连接/关联的情况下），尤其是在单个宿主关联（homed association）中，本发明解决了 SCTP 与 TCP 相比的弱点并且允许 SCTP 对于 TCP 应用更易接受。如果使用多宿主关联，则过大的消息有经过替换路径被发送的可能性（假设路径 MTU 沿着那些路径没有变化）。



图