

文章编号: 0529-6579 (2000) 03-0039-05

# LMDS 系统实现可靠的 IP 组播传送设计<sup>\*</sup>

周杰英, 张光昭

(中山大学无线电电子学系, 广东 广州 510275)

**摘 要:** 设计了 LMDS 系统实现可靠的 IP 组播传送方案, 采用带有反馈的 M/M/1 排队模型对该方案进行了性能分析, 计算机仿真结果表明了性能分析的正确性.

**关键词:** LMDS; RMTP; NACK; M/M/1 排队模型

**中图分类号:** TN 915.02 **文献标识码:** A

LMDS (local multipoint distribution system) 系统是一个利用 28 GHz 频段无线电波进行视线通信的点对多点双向广播式通信系统<sup>[1]</sup>. IP 组播 (multicast) 技术在点对多点通信时能节省主干共享链路上的带宽资源. IP 组播的 RMTP (reliable multicast transport protocol) 协议用于文档资料、白板传送等实时性要求不高但要求正确、有序传送的可靠 IP 组播业务. 本文参考 RMTP 协议, 设计 LMDS 系统实现可靠的 IP 组播传送方案, 并对这一方案的性能进行分析和计算机仿真.

## 1 LMDS 系统实现可靠的 IP 组播传送方案设计

LMDS 双向数字通信系统中各用户端间通信需通过头端中转, 若某用户端要可靠有序地组播 IP 分组给其它用户端接收, 则首先必须可靠地将 IP 分组上传给头端, 若传送不正确需重传. 其次是头端将收到的 IP 分组封装在 MAC 广播分组内下行广播, 各用户端主机接收后在 IP 层进行 IP 组播地址过滤. 若用户端成员主机收到不正确的或序号有误的 IP 组播分组, 需向头端发 NACK 消息, 通知头端重新组播.

假设 LMDS 系统中各用户端主机本身工作正常, 在 LMDS 系统通信范围内, 若各成员主机中有一个收到错误的或顺序不对的 IP 分组, 则其它成员也必定收到有误的分组. 因此, 这些成员中只要有一个成员向头端发 NACK 消息即可, 而不需要所有成员向头端发 NACK 消息, 以免系统陷入 NACK 消息的内爆.

每个 IP 组播的用户端主机为其参加的 IP 组播组设置一个变量 Flag. LMDS 系统运行 IP 组播管理协议 (IGMP) 时, 头端的组播路由器每隔一定时间向用户端组播 query 消息, 各成员收到后向头端发 report 消息, 可能只有一个或少数几个成员发回 report 消息, 头端从中选出一个下行组播. 各成员每收到一个由头端中转来的自己为该 IP 组播组发的 report 消

<sup>\*</sup> 收稿日期: 1999-12-20; 作者简介: 周杰英 (1966~), 女, 在职博士生, 讲师.

息，就将 Flag 变量置 1；每收到同组其它成员发出后由头端中转的 report 消息，就让其 Flag 变量置 0。各成员收到有错或序号不正确的 IP 组播分组时，先检查 Flag 变量，若 Flag 为 1 则向头端发 NACK 消息，Flag 为 0 不需发 NACK 消息。

2 头端可靠有序的 IP 分组缓存机制

2.1 IP 分组缓存结构

头端在内存开辟一个 IP 分组缓冲区，用于存放某用户端上传后等待可靠组播给其它用户端的 IP 组播分组，并相应建立一个如表 1 所示的 IP 分组映射表。映射表中的每一条记录分别映射缓冲区中缓存的一个 IP 分组，其中 IP 分组检测位（D2D1D0）的意义如表 2 所示。

表 1 头端可靠且有序的 IP 分组缓冲区映射表  
Tab 1 The mapping table of headend RMTP IP packet buffering

| IP 分组检测位<br>(D2 D1 D0) |   |   | IP 分组序号 | IP 分组存储地址 | IP 分组字节数 |
|------------------------|---|---|---------|-----------|----------|
| 0                      | 0 | 0 | 0       | *****     | XXXXXXXX |
| 0                      | 0 | 0 | 1       | *****     | XXXXXXXX |
| ...                    |   |   | ...     | ....      | ....     |

表 2 头端 IP 分组检测位的意义  
Tab 2 The meaning of headend IP packet detecting bits

| D2 | D1 | D0 | D2D1D0 为相应值时对应的意义                      |
|----|----|----|----------------------------------------|
| 0  | 0  | 0  | 该记录为空白记录，当存储新的 IP 分组条目后，D2D1D0 值改为 001 |
| 0  | 0  | 1  | 该记录对应的 IP 分组正等待组播                      |
| 0  | 1  | 0  | 该记录对应的 IP 分组当前已被组播，但不知是否成功             |
| 0  | 1  | 1  | 该记录对应的 IP 分组当前已被成功组播，根据需要可以被淘汰         |
| 1  | 0  | 0  | 该记录对应的 IP 分组以前已被成功组播过，目前又有后加入组播的成员请求组播 |
| 1  | 0  | 1  | 该记录对应的已成功组播过的 IP 分组再次播出，但不知是否成功        |

头端 IP 分组缓冲区中未存储 IP 分组时，IP 分组映射表中的每条记录为空白记录，每条记录的 D2D1D0=000。随着接收分组的增加，空白记录越来越少，甚至没有空白记录；之后随着一些旧分组的淘汰，又会有新的空白记录出现。当头端收到新的等待下行组播的 IP 分组时，搜索缓冲区是否有放得下该分组的空白空间，即检查映射表中是否有 D2D1D0=011 的可以被淘汰的纪录，若有则根据新分组待存储的字节数按序号由小到大的次序淘汰一定数量的纪录，并将这些记录对应分组所占空间整理成一块连续存储区域，之后缓存新分组，填写映射表中所淘汰记录中的一个记录，剩余的被淘汰记录变成新的空白记录。若没有可以被淘汰的记录，或者可以被淘汰的 IP 分组空间加上空白空间之和都不够存放新分组则不接收该分组，并向发送端发 NACK 消息。

2.2 IP 分组输出调度策略

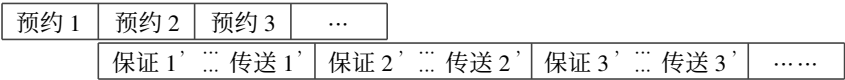
头端下行组播时从映射表中按 IP 分组序号由小到大的顺序挑选出 D2D1D0=001 的 IP 分组传送，并将它们的 D2D1D0 值改为 010。所选分组若传送不成功，则其 D2D1D0 值改

回 001, 否则其 D2D1D0 值改为 011. 若有的成员接收正常组播的 IP 分组的同时, 还想接收以前已成功组播过的一些 IP 分组, 则头端检查 D2D1D0=011 的分组中是否有这些分组, 若有则将这些分组的 D2D1D0 值改为 100. 之后, 头端按 D2D1D0=001 的分组优先, D2D1D0=100 的分组次之; 或者按其它调度策略同时组播这 2 组分组. 下行组播 D2D1D0=100 的旧分组时, 仍按序号由小到大的次序组播, 其分组组播后, D2D1D0 改为 101; 若成功组播, D2D1D0 改为 011; 若不成功, D2D1D0 改为 100.

头端下行组播时, 先选择某一速率组播, 若有某接收成员的缓冲区达到上限, 即可向头端提出降速要求; 若有成员的缓冲区达到下限, 则可向头端提出升速要求, 头端给予响应. 若既有成员要求降速, 又有成员要求升速, 则这两类成员接收能力相差太大, 需分成多个“群”, 采用不同的速率组播同一组信息, 也即进行“分群的可靠组播传送”(group reliable multicast transport). 适当增加映射表中 IP 分组检测位的位数, 定义好检测位的意义, 采用类似于同时可靠组播新、旧 IP 分组的方法, 可以实现“分群的可靠组播传送”.

3 发送端可靠有序的 IP 分组上传策略

LMDS 系统中上行传送一般采用用户端向头端预约上行帧传送时隙、头端分配上行帧时隙、用户端在分配的上行帧时隙中传送信息 3 个步骤串行完成. 这一策略显然不能满足实时传送业务的需求, 对于实时性要求不高的 RMTP 业务也是一种浪费. 本文提出适用于可靠 IP 组播分组上传的并串行策略如下所示.



支持这一上传策略的用户端 IP 分组缓存机制类似于头端, 但其 IP 分组映射表中 IP 分组检测位的意义不同, 如表 3 所示. 用户端收到待上传的 IP 分组时, 先存入 IP 分组缓冲区, 并修改 IP 分组映射表. 用户端每隔一定时间检查映射表, 看是否有未向头端上报的 IP 分组, 如果有就将这些分组的字节数之和  $N_i$  上报头端. 头端为每个 RMTP 业务流设置一个初值为 0 的待上传字节数计数器, 每收到一个上传的字节数  $N_i$ , 对应的计数器加  $N_i$ ; 每成功接收  $M_i$  个字节则让计数器减  $M_i$ . 只要业务流的计数器的值不为 0, 则头端就需根据字节数的多少等因素为该业务流分配上行带宽, 之后, 用户端实现可靠的 IP 分组上传. 并、串行上传策略与串行策略相比, 头端将及时地分配上行带宽, 并因此将加快上传速度.

表 3 用户端 IP 分组检测位的意义  
Tab 3 The meaning of subscriber IP packet detecting bits

| D1 | D0 | D1D0 为相应值时对应的意义                       |
|----|----|---------------------------------------|
| 0  | 0  | 该记录为空白记录, 当存储新的 IP 分组条目后, D1D0 值改为 01 |
| 0  | 1  | 该记录对应的 IP 分组未上报未上传                    |
| 1  | 0  | 该记录对应的 IP 分组已上报未上传                    |
| 1  | 1  | 该记录对应的 IP 分组已上报且已传送, 但不知是否传送成功        |

## 4 性能分析

对于本文提出的 LMDS 系统实现可靠 IP 组播传送方案, 可用图 1 所示的数学模型来分析发送用户端、头端以及接收用户端(分别用节点 1、2、3 表示)可靠的 IP 组播分组的队长分布、整个系统的队长分布和一个可靠传送的 IP 组播分组在系统内的平均逗留时间. 发送用户端和头端为带反馈的 M/M/1 排队模型, 接收用户端为不带反馈的 M/M/1 排队模型.

外部 IP 分组按泊松分布规律到达发送用户端, 到达率为  $\lambda$ , 之后在网络内传送. 3 个节点对 IP 分组的服务时间都服从负指数分布, 服务率分别为  $\mu_1$ 、 $\mu_2$ 、 $\mu_3$ .  $p_i$  为分组从节点  $i$  发送后由于接收不正确需重传的概率,  $q_i = 1 - p_i$ , 为 IP 分组从节点  $i$  成功传送的概率.  $p_3 = 0$ ,  $q_3 = 1$ . 设节点  $i$  的总平均到达率为  $\mu_i x_i$  (外部到达率  $\lambda_i$  及从其它节点到达节点  $i$  的全部到达率之和), 由 Jackson 定理<sup>[3]</sup> 可得

$$\begin{cases} \mu_1 x_1 = \lambda + p_1 \mu_1 x_1 \\ \mu_i x_i = q_{i-1} \mu_{i-1} x_{i-1} + p_i \mu_i x_i \end{cases} \quad i = 2, 3 \quad (1)$$

由(1)式可解得

$$x_i = \lambda / (\mu_i q_i) \quad i = 1, 2, 3 \quad (2)$$

$x_i$  实际上是各节点有反馈时的网络忙闲

度, 要求  $x_i$  小于 1 且不能接近与 1, 系统才能稳定. 若  $n_i$  表示第  $i$  个节点处的队长, 由文献<sup>[4]</sup> 可推得, 在统计平衡条件下, 系统处于状态  $(n_1, n_2, n_3)$  的概率  $p(n_1, n_2, n_3)$  为:

$$P(n_1, n_2, n_3) = \frac{1}{G(n)} \prod_{i=1}^3 (x_i)^{n_i} \quad (3)$$

$$\text{式中, } G(n) = \sum_{n_1+n_2+n_3 \leq n} \prod_{i=1}^3 x_i^{n_i} \quad (4)$$

若各节点的缓冲区足够大(一般能存储几十个 IP 分组就可看作无穷大了), 则可看作  $n \rightarrow \infty$ , 而如前所述稳定网络中  $x_i < 1$ . 在  $x_i < 1$ ,  $n \rightarrow \infty$  时, (4) 式可推得:

$$\lim_{n \rightarrow \infty} G(n) = \sum_{n_1, n_2, n_3} \prod_{i=1}^3 x_i^{n_i} = \frac{1}{1-x_1} \cdot \frac{1}{1-x_2} \cdot \frac{1}{1-x_3} \quad (5)$$

将(5)式代入(3)式可得:

$$P(n_1, n_2, n_3) = \prod_{i=1}^3 (1-x_i) x_i^{n_i} \quad (6)$$

第  $i$  个服务台处平均队长为

$$N_i = \sum n_i P(n_1, n_2, n_3) = \sum_{n_1, n_2, n_3} n_i (1-x) x_i^{n_i} \prod_{j \neq i} (1-x_j) x_j^{n_j} = \frac{x_i}{1-x_i} \quad 1 \leq i \leq 3 \quad (7)$$

整个系统平均队长为

$$N = N_1 + N_2 + N_3 = \sum_{i=1}^3 \frac{x_i}{1-x_i} = \sum_{i=1}^3 \frac{\lambda}{\mu_i q_i - \lambda} \quad (8)$$

IP 分组在系统内平均逗留时间为

$$T = \frac{N}{\lambda} = \frac{1}{\lambda} \sum_{i=1}^3 \frac{x_i}{1-x_i} = \sum_{i=1}^3 \frac{\lambda}{\mu_i q_i - \lambda} \quad 1 \leq i \leq 3 \quad (9)$$

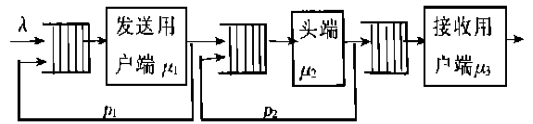


图 1 LMDS 系统可靠 IP 组播传送数学模型

Fig. 1 The mathematical model of LMDS RMTP

5 计算机仿真

对本文提出的 LMDS 系统的 RMTP 方案进行了计算机仿真, 外部 IP 分组按泊松规律到达, 到达间隔服从负指数分布, IP 分组在各节点的服务时间也服从负指数分布. 由仿真结果, 绘出各节点平均队长、IP 分组在系统的平均逗留时间与各结点的网络忙闲度  $x_i$  及 IP 分组外部到达率  $\lambda$  的关系曲线. 将其中一个节点的关系曲线与理论分析的关系曲线对比如图 2、3 所示. 从图中可以看出, 仿真结果与理论分析结果非常接近, 表明理论分析结果准确.

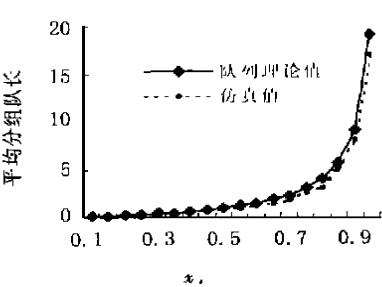


图 2 平均队长与负载( $x_i$ )关系

Fig 2 The Relation between IP packet length and load

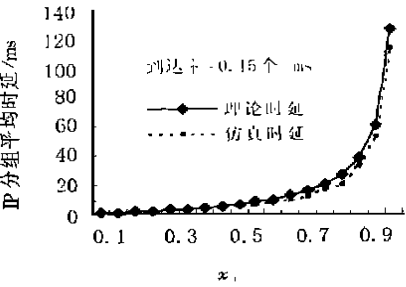


图 3 IP 分组平均时延-负载( $x_i$ )关系

Fig 2 The relation between IP packet delay and load

本文提出了 LMDS 系统实现可靠 IP 组播传送的方案, 分析了该方案的平均缓存队列长度、IP 分组平均时延, 计算机仿真结果表明了理论分析的正确性. 该方案及理论分析结果可用于指导实际 LMDS 上 RMTP 系统的建立.

参考文献:

[ 1 ] IMDS from NORTEL, 1998 <http://www.iec.org>  
[ 2 ] PAUL S, SABNANI K K, LIN J C H. Reliable multicast transport protocol (RMTP)[ J ] . IEEE J on Selected Areas in Communications 1997, 15( 3): 407~421.  
[ 3 ] JACKSON J R. Networks of waiting lines[ J ] . Operations Res, 1957, 5(4): 46~49.  
[ 4 ] GORDON W J, NEWELL G F. Closed queuing systems with exponential servers[ J ] . Operations Research, 15, 1967: 256~265.

Reliable IP Multicast Transport Based on LMDS

ZHOU Jie-ying<sup>\*</sup>, ZHANG Guang-zhao

**Abstract:** A reliable IP multicast transport scenario is proposed which include the scenario of NACK transporting, IP packet buffering and scheduling, upward and downward speed controlling, etc. After analyzing the performance of this scenario, the expression of average buffering length and average IP packet delay is given. The result of computer simulation shows that these expressions are correct.

**Keywords:** LMDS; RMTP; NACK; M/M/1 queuing model

<sup>\*</sup> Department of Radio and Electronics, Zhongshan University, Guangzhou 510275, China.  
(C)1994-2019 China Academic Journal Electronic Publishing House. All rights reserved. <http://www.cnki.net>