

文章编号: 1005-8451 (2018) 02-0047-05

旅客列车无线局域网带宽分配系统的 研究与实现

宫兴琦, 姜会增, 周用贵

(中国铁道科学研究院 标准计量研究所, 北京 100081)

摘要: 旅客列车无线局域网带宽分配系统采用链路发现模块均衡9路链路负载, 采用jhash_3words算法、RoundRobins算法构建用户发现模块, 实时更新系统内真实用户数量。该系统优化了旅客列车无线局域网互联网出口的带宽分配方式, 有效提升了互联网使用效率, 通过了国家铁路产品质量监督检验中心的性能测试与功能测试, 为在高速列车的实际运用奠定了基础。

关键词: 旅客列车无线局域网; 高速铁路; 互联网接入; 带宽分配系统

中图分类号: U285.211 : TP393 **文献标识码:** A

Wireless local area network bandwidth allocation system for passenger train

GONG Xingqi, JIANG Huizeng, ZHOU Yonggui

(Standards & Metrology Research Institute, China Academy of Railway Sciences, Beijing 100081, China)

Abstract: In the wireless local area network(LAN) bandwidth allocation system for the passenger train, the link discovery module was used to balance the load of 9 road links, jhash_3words algorithm and Round Robins algorithm were used to construct the discovery module, update the number of real users in the system in real time. The system optimized the mode of bandwidth allocation, could effectively enhance the efficiency in the use of the Internet. The system passed the performance test and the function test of the national railway product quality supervision and inspection center. It laid a solid foundation for the practical application in the high-speed train.

Keywords: wireless local area network(LAN) for passenger trains; high-speed railway; Internet access; bandwidth allocation system

为旅客提供互联网接入服务是旅客列车无线局域网系统的主要功能之一, 该系统采用移动、联通和电信3家运营商的4G网络, 构建9路车地通信链路, 提升互联网出口带宽。高铁及动车车厢内用户密度高, 车地通信出口带宽较小, 旅客列车无线局域网带宽分配系统作为旅客列车无线局域网接入系统的关键部分, 切实影响旅客接入互联网的实际体验。利用有限的带宽为旅客提供更好的互联网接入服务是带宽分配系统的一个重要功能。

1 带宽分配系统设计

1.1 系统组成

收稿日期: 2017-08-07

基金项目: 中国铁道科学研究院标准计量研究所研发课题 (HH-YFB-2016002)。

作者简介: 宫兴琦, 工程师; 姜会增, 副研究员。

旅客列车无线局域网由中心服务器、单车服务器、无线访问接入点(AP)和无线网桥组成, 中心服务器主要提供旅客互联网接入服务, 单车服务器提供车厢内局域网服务, AP提供无线终端接入服务, 无线网桥连接车厢间的单车服务器, 局域网组成如图1所示^[1]。

旅客列车无线局域网带宽分配系统部署于中心服务器, 由用户发现模块和链路发现模块组成, 主要用于提升互联网接入利用率。

1.2 系统功能

带宽分配系统的用户发现模块主要用于实时更新车上用户的状态, 适应高铁停靠站旅客上下线较频繁的情况, 为链路发现模块提供信息支持; 链路发现模块主要根据在线用户状态和9路车地链路状态为用户选择合适的通信信道。

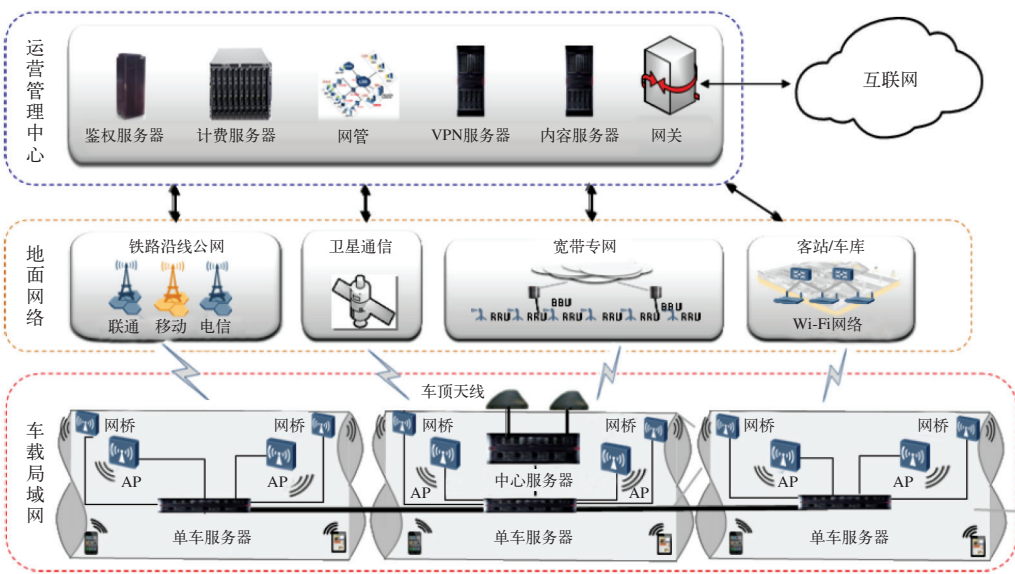


图1 旅客列车无线局域网组成图

2 关键技术

2.1 jhash_3words哈希算法

jhash_3words 算法为 Linux 网络协议中常用的一个哈希算法，适用于网络数据哈希运算。通过用户数据包解析出用户的 mac 地址，并通过 jhash_3words 算法将解析出的 mac 地址进行 hash 散列运算，用户发现模块通过对 jhash_3words 算法进行二次封装，得到 mac 地址散列运算函数 get_tuple_hash。

运算结果更新，存入用户信息链表，信息链表的单个节点记录了用户的状态、最新数据更新状态和权限级别。当用户接入中心服务器后，get_tuple_hash 函数通过用户源 IP 和 mac 地址进行 hash 运算，根据运算结果将用户登入时间和权限等信息更新到指定的链表结点上，如果指定结点不存在则表示为新用户，为其创建新的结点^[2]。如果结点已经存在则表示该用户为老用户，更新其最后一次访问时间。

当内核态通过 hash mac 散列算法将用户列表更新后，内核通过 netlink socket 动态消息通道向用户态应用发送新的用户列表。

2.2 RoundRobin算法

链路发现模块负责数据包在各条可用链路上的并发传输，带宽分配方式的选择原则是并行带宽增大和数据传输可靠。

采用 RoundRobinScheme 算法^[4]，也称为轮询调度算法，该算法把来自用户的请求轮流分配给内

部链路，从 1 开始，直到 N（网卡个数），然后重新开始循环。

RoundRobin 算法的优点是简洁，无需记录当前所有连接的状态，是一种无状态分配。系统使用了 RoundRobin 的权重分配算法，即为每条链路分配不同的比例系数，例如，3 块网卡设置为 3 : 1 : 1 时，将控制数据包在各条链路上的分发比例为 3 : 1 : 1。

3 系统实现

带宽分配系统采用 C 语言开发，开发平台为 Centos6.6，内核版本为 2.6.38，采用的编译器为 gcc。系统的用户发现模块采用内核编程方式，链路发现模块采用应用层编程方式，二者通过 netlink 套接字方式通信。

3.1 用户发现模块实现

当用户通过列车 AP 接入中心服务器后，用户发现模块通过内核 HOOK 获取用户数据包^[3]，根据数据包信息通过 jhash_3words 算法更新用户列表，之后通过 netlink 套接字将消息发送至应用层接收程序，应用层接收程序实时更新用户列表。部分实现代码如下：

```
u32 kmod_hook_init(void)
{
    /* 该钩子对应的处理函数 */
    g_cwlan_in_hook_prer.hook = (nf_hookfn *)
cwlan_in_hook_prer;
    /* 使用 IPv4 的第一个 hook */
    g_cwlan_in_hook_prer.hooknum = NF_IP_
PRE_ROUTING;
    g_cwlan_in_hook_prer.pf = PF_INET;
    g_cwlan_in_hook_prer.priority = NF_IP_PRI_
FIRST; /* 让我们的函数首先执行 */
}
```

```

/* 将用户自己定义的钩子注册到内核中 */
nf_register_hook(&g_cwlan_in_hook_prer);

....

return CWLAN_OK;
}

```

jhash_3words 算法进行二次封装,得到高性能 mac 地址散列运算函数 get_tuple_hash,实现代码如下:

```

u32 get_tuple_hash(u32 a, u32 b, u16 c, u16 d,
u32 mask)
{
    a += c;
    b += d;
    c = 0;
    a = a^b;
    b = a;
    return (jhash_3words(a,(a^c),(a|(b<<16)),
0x5123598) & mask);
}

```

此时, get_tuple_hash 函数通过用户源 IP 和 mac 地址进行 hash 运算,根据运算结果将用户登入时间和权限等信息更新到指定的链表结点上,如果指定结点不存在,则表示为新用户,为其创建新的结点。如果结点已经存在,则表示该用户为老用户,更新其最后一次访问时间,其部分实现代码如下:

```

u32 flow_session_match_online_list(wlan_
quintuple_s *quintuple_info)
{
    ....

    // 通过 hash 算法计算散列值
    hash_key = get_tuple_hash(quintuple_
info->ip_hd.saddr,(u32)quintuple_info->eth_hd.h_
source[0],
(u16)quintuple_info->eth_hd.h_source[4],(u16)
quintuple_info->eth_hd.h_source[5], FLOW_
SESSION_HASH_LEN-1);

    hash_head = g_cw_fs_hash + hash_key;

```

```

write_lock_bh(&hash_head->rwlock);
// 遍历链表中所有 node, 查找传入 mac 是否
存在
list_for_each_entry_safe(node, temp_
node,&hash_head->h_head,list)
.....
return user_status;
}

```

当内核态通过 hash mac 散列算法将用户列表更新后,内核通过 netlink socket 动态消息通道向用户态应用发送新的用户列表,netlink 需要定义新的协议号用来标示消息,如下所示:

```

#define NETLINK_WiFi 17
u32 wlan_sendto_umod(u32 type, u8 *buff, u32
datalen,struct sk_buff *skb)
{
    ....

    out_skb = nlmsg_new(NLMSG_DEFAULT_
SIZE, GFP_KERNEL); // 分配足以存放默认大小的
sk_buff

    ....

}

```

3.2 链路发现模块实现

中心服务器使用移动、联通、电信的 3G/4G 网络作为车地通信的网络出口,支持最多 9 条无线链路并发传输用户数据,但是相对于乘客的需求而言,此时带宽仍属于窄带通信,因而如何将 9 路链路充分利用,提高带宽分配效率是提升乘客使用体验的一个重要环节。

链路发现模块的作用在于即时维护各条链路的真实状态,包括链路的可用性、链路 IP 变换,并且在链路状态发生变化时能够将用户调度到可用链路上,对上层应用是透明的,能够保证用户通信的连续性。

链路发现模块需要维护可用链路链表,在 IP 发生变化时通过可用链表来选择数据转移的去处,同时更新可用链表中的内容,IP 切换的过程可以分解为两个部分:

(1) 现有的 IP 地址消失,即旧的链路失效,此时,应该将数据转移到其他链路上去。

(2) 新的 IP 地址出现,即新的可用链路出现,此时,应该将一部分数据分配到新的可用链路上。

在新的 IP 出现时需要新建套接字描述符,并将其代表的链路名称等信息生成一条新的可用链路结构加入到链表中,同时,将代表该链路上的旧 IP 记录删除,注销套接字描述符。

创建套接字的过程并不复杂,关键在于要保证在 IP 切换的过程中数据能够及时、连续地进行交互^[5]。当可用链路的数量发生变化时,自动地将数据进行转移分发,既不能向失效链路分配数据,也不能使新的可用链路闲置。

以 4 块网卡为例,当 4 条链路变为 3 条,或者 3 条链路变为 4 条之后,需要对网卡进行重新排序,调整各条可用链路的状态,处理流程如图 2 所示。

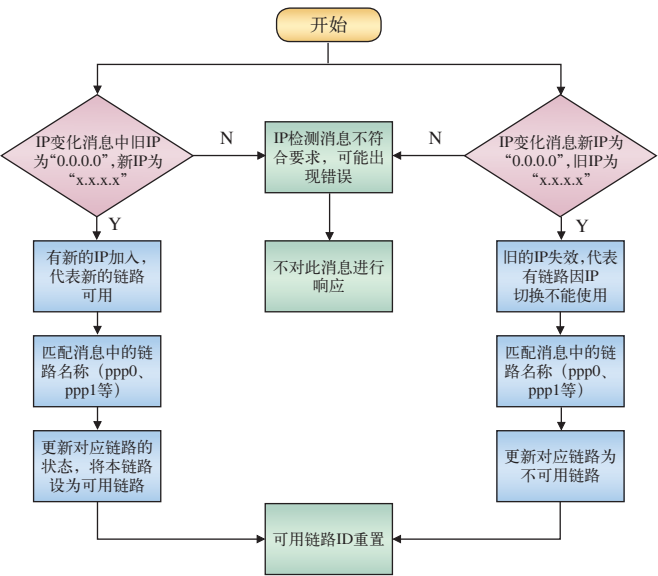


图2 可用链路更新流程图

可用链路模块在数据转移模块中的作用比较重要,可用链路模块维护了现在可用链路的全部信息,为数据包的链路分发提供了重要参考。

链路发现模块负责数据包在各条可用链路上的并发传输,实现并行带宽增加和数据传输可靠。在调度方式的选择上,采用 RoundRobin 算法^[6]。

本文允许其他模块通过调用 RoundRobin 模块对各条链路的权重系数进行调节,其中,sendpercent^[7]为权重系数,链路发现模块将根据 sendprecent 设置

的系数,并进一步通过匹配 ifname 和 device_id 来确定对应的链路,从而正确设置系数。

RoundRobin 算法在高速移动环境下简单易实现,是从链路负载的角度进行考虑。然而从用户使用的角度上讲有着很大的局限性,通过简单的轮询,同一个应用的数据流被分到不同的链路上很可能非但没有增加并行传输的带宽,反而造成了链路带宽的下降。假设有两条链路 A 和 B, A 链路的往返时间(RTT)为 40 ms, B 链路的 RTT 为 80 ms,如图 3 所示。

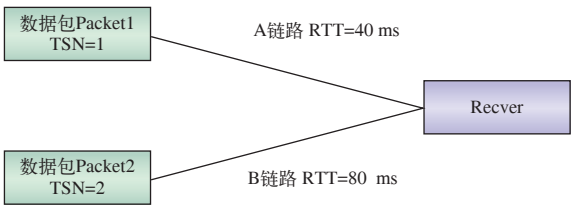


图 3 RoundRobin轮询算法示例

由图 3 可知, B 链路的时延较大,在 Packet1 到达 Recver 后,需要等到 Packet2 也到达 Recver 才能向上层应用进行递交,也就是说,即便 A 链路此时的 RTT 较小,也需要等待 B 链路的数据包,此时对于用户来讲, B 链路的情况较差。在 A 链路不发生堵塞的情况下, B 链路的加入反而拉低了 A 链路的表现。

当某条链路的 IP 失效时,链路发现模块会将发生 IP 变化的链路消息通知数据转移模块,此时需要将此链路设为不可用链路,停止继续分配数据包至此链路,代码如下:

```
// 如果是 EVDO1 的地址发生变化
if(strcmp(infotem.device_id,idname[j])==0)
{
    // 如果消息中新 IP 为 0 说明网卡
    DWON 掉
    if(strcmp("0.0.0.0",inet_ntoa(infotem.
newip))==0)
    {
        // 原先的套接字将失效,轮询比例
        设为 0
        mylist[j].dis=0;
        mylist[j].distem=0;
        printf("%s has been shut down\
```



```
n",mylist[j].device_id);
    }
}
```

IP 变化消息中的新 IP 为 0.0.0.0, 说明该网卡的 IP 已经失效, 需要将此链路中的轮询比例 (dis) 属性设为 0, 数据调度模块会检测 dis 的数值, 当为 0 时会停止向此链路分发数据, 此时可用链路的数目发生了变化, 需要更新链路 ID, 确保可用链路链表的状态是最新的, 通过检查 dis 属性设置可用链路的 ID 和可用链路的数量^[5]。

当接收到的 IP 变化的消息中新 IP 不为零时说明有新的可用链路加入, 此时应该将其加入到可用链路链表, 为其分配数据^[8], 代码如下:

```
if(strcmp("0.0.0.0",inet_ntoa(infotem.
newip))!=0&&strcmp(mylist[j].ip,inet_ntoa(infotem.
newip))!=0)
{
    ...
    // 设置新的轮询比例

    mylist[j].dis=1;
    mylist[j].distem=mylist[j].dis;
    // 更新 IP
    strepy(mylist[j].ip,inet_ntoa(infotem.newip));
    printf("%s has been set up\n",mylist[j].device_
id);
    //ip 修改
    netidinit();
    ...
}
```

4 系统测试

带宽分配系统在试验室环境下已经通过功能验证, 经测试, 该系统能够在常数时间内判别大量用户实时状态, 提升 9 路链路的应用效率, 采用运营商的网络进行互联网接入, 具备同时接入中国移动、中国联通和中国电信 3G/4G 网络的能力。每列车中

心服务器上统一配置一套互联网接入设备, 具备多信道带宽汇聚并在车载局域网内进行自适应分配的功能。

5 结束语

旅客列车无线局域网带宽分配系统在试验室环境下已经通过功能验证, 但是在实际情况下, 每条线路的运营商 3G/4G 信号覆盖强度不同, 导致不同的线路出口带宽差别较大。下一步需做的工作是在实际运行环境下进行测试, 通过积累大量用户的行为特征与数据特性, 更好地优化互联网接入功能, 更进一步提升互联网带宽分配效率, 提升用户使用体验。

参考文献:

- [1] 中国铁路总公司. 旅客列车无线局域网系统和安装布线总体要求 (暂行) [Z]. 北京: 中国铁路总公司, 2015.
- [2] 李晨光. Linux 系统网络管理模块的实现 [J]. 铁路计算机应用, 2008, 17 (6): 35-38.
- [3] 张 昊. 高铁车地通信系统级仿真平台设计与多基站协作技术的研究 [D]. 成都: 西南交通大学, 2013.
- [4] 丁良辉. 无线自组织网络中的 TCP 协议研究 [D]. 上海: 上海交通大学, 2008.
- [5] 周开波. 利用 mSCTP 实现 3G 和 WLAN 无缝切换及其性能分析 [D]. 北京: 北京邮电大学, 2012.
- [6] 何 川. 移动自组织网络中的可靠传送协议 [D]. 成都: 电子科技大学, 2013.
- [7] 宫兴琦. 高速移动环境下连接及拥塞控制的研究与实现 [D]. 北京: 北京交通大学, 2014.
- [8] 方 棋, 吉 强, 韩 臻. Linux 环境下基于策略路由实现多线路负载均衡 [J]. 铁路计算机应用, 2006, 15 (9): 35-38.
- [9] Andrea De Vendictis, Andrea Baiocchi, Michela Bonacci. Analysis and enhancement of TCP Vegas congestion control in a mixed TCP Vegas and TCP Reno network scenario[J]. Performance Evaluation, 2003(533).
- [10] Tine Stegel, Janez Sterle, Urban Sedlar, et al. SCTP multi-homing provisioning in converged IP-based multimedia environment[J]. Computer Communications, 2010(3314).

责任编辑 王 浩