

卫星网络路由器高速数据转发设计

周 东, 王翠莲, 李 珂, 张 犛, 冯国平

(北京空间飞行器总体设计部, 北京 100094)

摘要: 卫星网络由多颗卫星和地面站组成, 星间通过高速激光链路通信; 卫星网络的核心之一就是路由器; 高性能路由器的典型特点为数据路径和控制路径的分离; 控制路径处理与高层路由协议相关的数据包, 数据路径处理需要转发的数据包; 数据路径是路由器的关键路径, 直接影响着路由器的整体性能; 在调研路由器技术发展历程之后, 分析了高性能路由器典型结构及相关关键技术, 考虑目前卫星网络系统需求、软硬件环境约束条件, 对现有技术进行了优化和适应性修改, 确定了卫星网络路由器中数据路径的实现方案; 该方案满足当前卫星网络应用需求, 且经简单扩展后, 还可满足后续更高性能卫星网络路由器的设计需求。

关键词: 卫星网络; 路由器; 数据转发

A Design of High-Speed Data Forwarding for Router of Satellite Network

Zhou Dong, Wang Cuilian, Li Ke, Zhang Lu, Feng Guoping

(Beijing Institute of Spacecraft System Engineering, Beijing 100094, China)

Abstract: The satellite network consists of several satellites and ground stations. The satellites communicate with each other through high-speed laser links. One of the core components of the satellite network is the routers. A typical feature of high performance routers is the separation of the data paths and the control paths. The control paths process data packets of the high-level routing protocol, the data paths process the data packets that need to be forwarded, and the data forwarding of different interfaces is performed independently. The critical paths of the routers are the data paths which affect the overall performance of the routers directly. After investigating the development history of the router technology, we discussed the typical structure and the key technologies of high-performance routers. Considering the requirements of the current satellite network system and the constraints of software and hardware environment, we have optimized the existing technology adaptively and confirmed the implementation method of the data paths of the satellite network router. The scheme meet the present requirements of satellite network, it can also meet the design requirements of subsequent high-performance routers after simple expansion.

Keywords: satellite network; network router; data forwarding

0 引言

卫星网络包括卫星系统、地面站等多种网络节点, 构成一个典型的多层、天地一体的空间信息网络, 具备全球数据接入服务能力。星间通过高速激光链路通信。卫星网络的核心之一就是路由器。卫星系统搭载激光终端和高速路由器进行在轨验证, 同时也为我国天地一体化网络的建设提供支撑。

在讨论了路由器发展历程之后, 对高性能路由器典型结构、路由查找关键技术、内部交换关键技术进行了调研, 结合卫星网络的路由需求、软硬件环境的设计约束条件, 开展了卫星网络路由器数据转发设计。所实现数据转发设计具有路由效率高、交换效率高的特点, 并具备一定的可扩展性, 经简单扩展后, 还可满足后续更高性能空间路由器的设计需求。

1 路由技术调研

路由器出现在 20 世纪 80 年代末, 用于互联不同类型的网络的通用组网设备, 工作在 OSI 参考模型的第三层, 适应

于大型组网对性能、容量和安全性的要求。

高性能路由器结构的一个典型特点是数据路径和控制路径的分离。输入接口接收来自外部高速网络接口的数据包, 根据转发表信息, 通过交换网络将数据包传输到目的输出接口, 发送到外部网络。转发表由路由引擎下载到各个接口板上, 一般的数据包转发不再通过路由引擎, 实现了数据路径与控制路径的独立, 不同接口的数据转发也是独立进行的。既提高了系统的转发性能, 又增加了数据路径的可靠性。数据路径处理需要转发的数据包, 是整个路由器的关键路径, 它的实现好坏直接影响着路由器的整体性能^[1-2]。数据路径设计中的 2 个关键问题是路由查找和内部交换。高性能路由器典型结构如图 1 所示。

1.1 路由查找

路由查找的主要功能: 基于 IP 数据包的目的地址, 使用最长前缀匹配规则, 对 IP 路由表进行查询, 获知对应的输出端口。路由查找的主要方法有以下 3 种^[3-4]:

(1) 基于软件的查找方法: 软件实现一般采用基于 Trie 的二叉树结构或基于键树 (Key Tree) 的算法。在键树中, 从搜索键中顺序抽取相应数位或字母字符决定分支去向。该方法实现简单, 适用性好, 可以用于任何长度关键字的查找。但查找效率低, 存储效率较低, Trie 树查找过程实际上就是在长度空间内的顺序查找操作^[5]。

收稿日期: 2018-07-04; 修回日期: 2018-07-31。

作者简介: 周 东(1984-), 男, 四川南充人, 硕士, 工程师, 主要从事空间信息化及信息处理方向的研究。

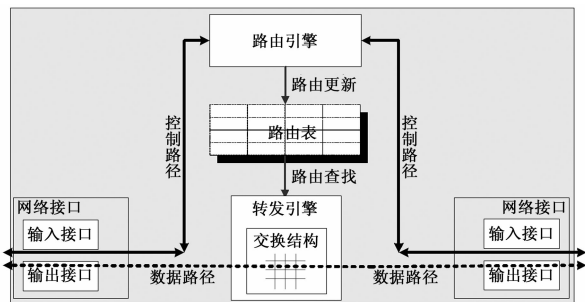


图 1 高性能路由器典型结构

(2) 基于大容量 RAM 的快速路由查找方法：随着网络传输速度大规模提升，基于 CPU 的纯软件处理满足不了要求。引出了基于大容量 RAM 的快速路由查找方法。将路由查找表项存储到大容量 SRAM 中。以“多分支 Trie 树查找算法”为例，在 RAM 中存储两种查找表结构，分别为 Table16 和 TableNext。Table16 表保存路由地址前缀小于等于 16 的表项；TableNext 表保存路由地址前缀大于 16 的表项。对于任何一个 IP 地址，查找过程最多只需要访问两次查表 (Table16 和 TableNext) 即可获得转发信息，大大加快了路由查找的速度^[6]。

(3) 基于 TCAM 路由查找方法：TCAM (Ternary Content Addressable Memory) 三元内容可寻址存储器，是近年来在各种高端设备上逐渐广泛使用的一种技术。TCAM 采用并行查找技术，与查找速度和表项的条目数量无关。采用 TCAM 可以实现每秒上亿次的查找。TCAM 器件在速度和容量上都有很大提高，能够满足大规模高速路由查找的需求。基于 TCAM 的查找速度快，但是实现代价较高，功耗较大，表项的更新复杂^[7]。

1.2 交换结构

交换结构是高性能路由器的核心，完成输入端口和输出端口的连接，是影响路由器速度和容量的关键因素。根据路由器中使用的交换结构不同，路由器可分为共享总线、共享存储器、交叉开关阵列等类型^[8-9]，交换结构分类如图 2 所示。

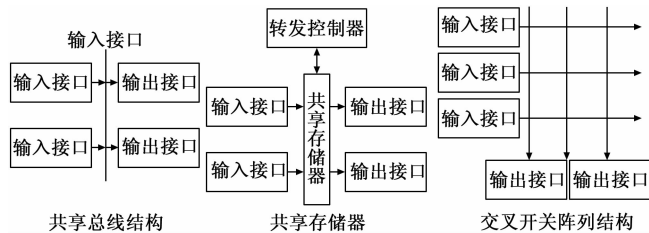


图 2 交换结构分类

(1) 共享总线结构：数据包在路由器中通过共享总线传输。共享总线采用时分复用方式，连接在共享介质上的某一输入、输出端口只在指定的时间片进行数据传输。这种方式交换速率受限于总线的带宽。

(2) 共享存储器结构：使用了大量的高速 RAM 来存储输入数据，由于数据首先由输入端口存入共享存储器，再

从共享存储器传输到输出端口，其交换速率受限于存储器的访问速度。包头信息传输到转发控制器，由转发控制器来决定读取哪个数据包传送到输出端口。如果要想实现输出排队，存储器的操作速度必须多倍于端口速度，这将受限于存储器性能而难以扩展。

(3) 交叉开关阵列结构：交叉开关阵列结构的路由器中，多个数据分组可以同时通过不同的线路进行交换，从而极大地提高系统吞吐量，使系统性能得到显著提高。系统的最终交换带宽取决于交叉开关阵列和调度器的能力，而不是取决于互连介质。可实现吞吐量和交换延迟的最优化，满足高速路由器对容量和性能的要求。

2 总体方案

在调研高性能路由器技术发展历程、典型结构及相关关键技术的基础上，考虑目前卫星网络系统需求、软硬件环境约束条件，对现有技术进行了优化和适应性修改，确定卫星网络路由器数据转发实现方案如下：

(1) 路由器结构：选用第三代高性能路由器作为卫星网络路由器的设计结构，控制路径和数据路径分离。控制路径根据网络拓扑动态生成路由表，将路由表项注入到数据路径的转发表中；数据路径从 IP 数据包中提取包头信息，查询转发表获得下一跳端口，将数据交换到输出线路上。数据路径实现最长前缀匹配、数据缓存、内部交换等功能。路由器采用 CPU+FPGA 的实现方案，其中路由管理 CPU 完成控制路径的功能，路由处理 FPGA 完成数据路径的功能；

(2) 路由查找：基于软件的查找方法效率较低，而基于 TCAM 的路由查找方法需要专用器件，目前尚无宇航产品。因此使用 RAM 存储转发表项，并实现最长前缀匹配的功能。并借鉴 IP 网三层交换中“流式 IP 交换”策略，经适应性改进后，采取了“基于标记的快速转发策略”，提高数据包路由查询效率；

(3) 交换结构：选用交叉开关阵列的交换结构。输入输出接口之间设置多对多交换网络进行数据包传输。并实现数据缓存、内部传输等功能。并采取“动态优先首 1 查询策略”，提高数据包交换效率。

3 路由查找

路由策略由路由管理 CPU 和路由处理 FPGA 的配合完成，其中路由管理 CPU 负责根据链路状态以及地面注入的路由信息维护接入节点表和骨干路由表。并生成硬件转发表，注入到路由处理 FPGA 中。路由处理 FPGA 依据转发表完成数据的转发^[8]。

3.1 最长前缀匹配

路由处理 FPGA 上设置转发表缓冲区，用于存储转发表项。转发表项包括目的 IP 前缀、前缀长度、端口号等信息，仅需匹配前缀长度字段规定的的数据位。前缀长度较长的转发表项需放置到转发表中靠前位置，路由查表时从转发表第一项开始查找，如此前缀较长的表项首先被查询到。

默认路由表项：路由管理 CPU 填充完有效 IP 转发表项后，在转发表末尾再填入一项默认转发表项，前缀长度为

0。所有与表中前部表项不匹配的 IP 数据包, 均能与之匹配, 并从所设置的默认通道输出。转发表格式如表 1 所示。

表 2 实验设备主要参数

序号	目的 IP 地址	前缀长度	端口
1	FC00:0:0:D300:0001:0:0:0	64	0
2	FC00:0:0:D300:0002:0:0:0	64	0
3	FC00:0:0:D301:0:0:0:0	56	1
4	FC00:0:0:D302:0:0:0:0	56	2
...
N	0:0:0:0:0:0:0:0	0	5

3.2 转发表项管理

路由处理 FPGA 中采用“乒乓”缓存方式解决路由更新和路由查询的冲突。双口 RAM 一端接收路由管理 CPU 写入的转发表项, 另一端供路由查找模块使用。如图 3 所示, 路由管理 CPU 向 DRAM1 注入转发表项时, 路由查找模块使用 DRAM2。表项注入完成后, 路由管理 CPU 发送“注入完成”标识, 切换 DRAM 对应关系。DRAM2 供路由管理 CPU 注入, DRAM1 供路由由查询模块使用。

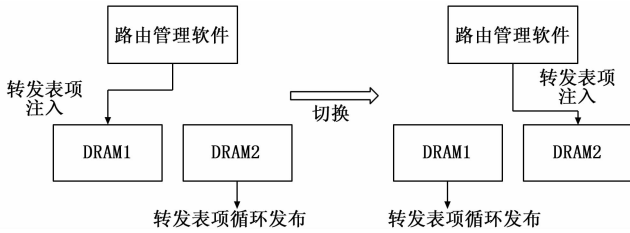


图 3 转发表注入和输出示意图

转发表中的表项通过“循环发布”的方式送给各通道。每个时钟周期从转发表中提取一个表项, 发布到各接收通道。每个周期开始阶段需发出“循环开始标记”, 各通道接收到数据包后, 等待“循环开始标记”后, 依次匹配转发表发布的各个表项。如有匹配则退出路由查询, 并进行数据包传输; 如匹配超时, 则丢弃当前数据包。各接收通道独立进行路由匹配, 某一接收通道故障不影响其他接收通道。转发表循环发布方式如图 4 所示。

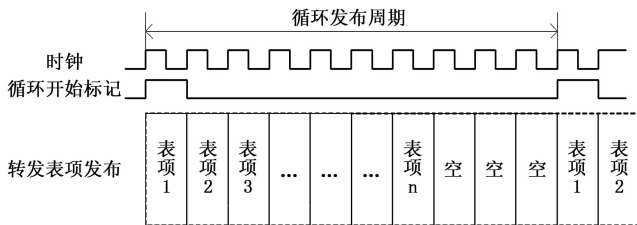


图 4 转发表循环发布

3.3 路由查表的优化

为提高路由查表效率, 路由处理 FPGA 借鉴了 IP 网三层交换中“流式 IP 交换”策略(利用 IP 包头信息来对后续数据业务流进行标记, 记录 IP 地址、端口号、应用层协议类型、服务质量要求等信息, 建立从发端到收端的转发路径。具有同一特征标记的业务流的后续分组直接传送到已

记录的端口, 不必对所有数据包进行路由查找^[8])。经适应性改进后, 采取了“基于标记的快速转发策略”。接收到数据包后, 记录数据包的“特征标记”。当有新的数据包送入时, 将目的 IP 与所记录的“标记”进行匹配, 同时与转发总表发布的表项进行匹配。如果与“标记”匹配, 则直接进入传输环节, 否则需要对转发总表进行查询, 匹配后进入传输环节。当连续接收的数据包为同一目的 IP 时, 路由查询周期仅为 1 个时钟周期。还可根据需要设置多个“特征标记”, 数据包与最近的多个“标记”之一相匹配时, 查询周期也仅为 1 个时钟周期。路由查表流程图如图 5 所示。

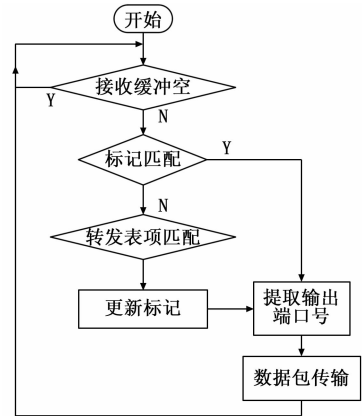


图 5 路由查表流程图

为保证“最长前缀匹配”的策略的实现, “标记”的格式为目的 IP 地址全部字段和输出通道号, 需匹配目的 IP 的全部字段, 且当转发总表发生更新时, 需将所记录的“标记”清除。

4 内部交换

4.1 交换单元

路由处理 FPGA 内部采用多对多交换网络进行数据包交互。图 6 为 MxN 交换单元示意图, 传输控制模块数量为 M, 接收控制模块为 N。每个传输控制模块均能对任一接收控制模块提出申请。每个接收控制模块均能响应任一传输控制模块发出的数据传输请求。传输控制模块、接收控制模块均独立运行。

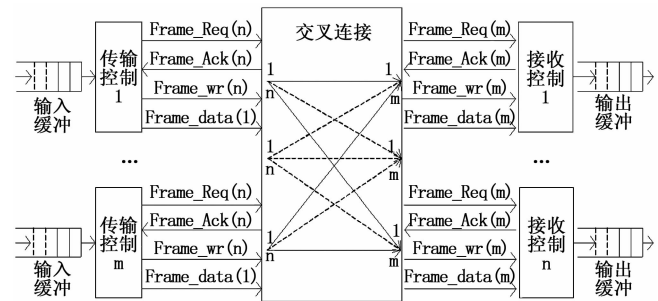


图 6 M x N 交换单元

传输控制模块实现的功能: 当输入缓冲非空时, 向对应接收控制模块发出传输申请 Frame_Req; 等待应答 Frame_Ack; 通过写信号 Frame_Wr、数据信号 Frame_

data 将数据包传输到接收控制模块。

接收控制模块实现的功能：当输出缓冲非满时，检查 M 个传输控制模块送出的申请信号；根据有效传输申请 Frame_Req 的序号，向对应传输控制模块的发出应答信号 Frame_Ack；等待传输控制模块发出数据包，在 Frame_wr 信号有效时，将 Frame_Ack 信号撤销，并将 Frame_data 存入输出缓冲中。

传输控制模块和接收控制模块均设置“超时退出”机制，避免突发异常影响后续数据传输。

4.2 动态优先首 1 查询策略

接收控制模块检查 M 个传输控制模块的申请信号，每个时钟周期可以处理其中一条申请。按照传统轮询方法，轮询一遍需要 M 个时钟周期。即使申请信号中仅有一个有效信号，也需要等待轮询时间到，才能被处理，平均耗时 $\frac{M}{2}$ 个时钟周期。采取“动态优先首 1 查询策略”，在一个时钟周期内计算出所有申请信号中第一个有效信号，并可动态改变各路通道申请信号的排列顺序，以均衡各申请信号的被受理的概率。动态优先排序方法如表 2 所示。

表 2 动态优先排序

查询周期序号 $cknum$	申请信号动态优先排序 $reqtemp$						
0	0	1	2	3	4	...	$m-1$
1	1	2	3	4	...	$m-1$	0
...
$m-1$	$m-1$	0	1	2	3	...	$m-2$

在查询周期 0，申请信号按照 $0 \sim m-1$ 顺序排列；在查询周期 1，申请信号 1 排在第一个位置，申请信号 0 排在末尾；以此类推，在查询周期 $m-1$ ，申请信号 $m-1$ 排在第一个位置，申请信号 $0 \sim (m-2)$ 按序排列。

首 1 查询策略的实现方法如表 3 所示（以八位申请信号为例），计算出申请信号 $reqtemp$ 中第一个为 1 的排列序号 $index$ 。

表 3 首 1 查询策略示例

序号	申请信号排列 $reqtemp$	首 1 序号 $index$	有效信号 $flag$
1	10001001	0	1
2	01001001	1	1
3	00001100	4	1
4	00000001	7	1
5	00000000	0	0

$flag$ 为 1 时，表示申请信号组合中存在有效申请信号。将首 1 查询所得“首 1 序号” $index$ 与动态优先策略中查询周期序号 $cknum$ 相加，并对申请通道总数 M 取模，可得动态优先排序后的第一个有效申请通道序号 $reqnum$ ，计算公式如下所示。

$$reqnum = (cknum + index) \% M \quad (1)$$

5 实验结果与分析

某空间骨干网络采用了该方案的卫星网络路由器，已

成功地实现了星间网络数据帧的高速转发功能，同时支持星内遥控数据上行及遥测数据下行等传输业务。根据测试结果，每路接口工作速率最高可达 1.6 Gb/s，能够满足大部分卫星网络的需求。测试结果如表 4 所示，经扩展后还可以满足未来卫星网络高速数据转发需求。

表 4 高速数据转发性能表

序号	测试项目	设计指标	实测结果
1	端口速率	1Gb/s	1.6Gb/s
2	端口数量	6	6
3	网络协议	TCP/IP	TCP/IP
4	时延	不大于 1ms	不大于 25 μ s

6 结束语

在讨论了路由器发展历程之后，对高性能路由器典型结构、路由查找、内部交换关键技术进行了充分调研，结合卫星网络的路由需求、硬件电路设计约束条件，开展了卫星网络路由器数据路径的设计。采取了基于 RAM 的路由查找策略、“循环发布”的方式，各接收模块独立进行路由匹配；采取了“基于标记的快速转发策略”，实现“一次路由，多次交换”，提高数据包转发效率；数据传输路径采用多对多交换网络进行数据包交互，各通道独立运行；采取“动态优先首 1 查询策略”，在一个时钟周期内受理所有申请信号中第一个有效信号，并可动态改变各路通道申请信号的排列顺序，以均衡各申请信号的被受理的概率。所实现的路由查找功能、内部交换结构，经扩展后还可满足后续更高性能卫星路由器的设计需求。

参考文献：

[1] 朱培栋, 孙志刚, 张晓哲, 等. 高性能路由器设计与实现 [M]. 北京: 国防工业出版社, 2013.

[2] 梁志勇, 徐 恪, 分布式路由器中的路由管理模型 [J]. 清华大学学报 (自然科学版), 2003, 43 (4): 503-506.

[3] 卢锡城, 管剑波. 高速路由与交换技术现状及发展趋势 [J]. 国防科技大学学报, 2005, 27 (5): 1-5.

[4] 熊忠阳, 阳信宏, 张玉芳. 分组 IP 路由由最长前缀匹配查找算法研究 [J]. 世界科技研究与发展, 2011, 33 (6): 1014-1018.

[5] 王振兴, 王智强, 孙亚民, 等. 基于二分搜索 Trie 的 IPv4/IPv6 路由快速查找算法 [J]. 计算机工程, 2005, 31 (2): 108-109.

[6] 殷 科, 邓亚平. 基于 RAM 和 TCAM 存储结构的高速路由查找算法 [J]. 计算机工程与应用, 2005, 20: 159-161.

[7] 蔡立军, 李 杜, 池 鹏, 等. 一种支持 TCAM 规则更新与压缩方法 [J]. 湖南大学学报 (自然科学版), 2014, 41 (8): 94-100.

[8] 罗国明, 沈庆国, 张曙光, 等. 现代交换原理与技术 (第 3 版) [M]. 北京: 电子工业出版社, 2014.

[9] 管剑波, 孙志刚, 卢锡城. 使用多级交换网络进行高性能路由器设计 [J]. 计算机研究与发展, 2005, 42 (6): 965-970.