



(12)发明专利

(10)授权公告号 CN 105791145 B

(45)授权公告日 2019.10.25

(21)申请号 201410802647.6

H04L 12/46(2006.01)

(22)申请日 2014.12.18

(56)对比文件

(65)同一申请的已公布的文献号
申请公布号 CN 105791145 A

CN 104144120 A, 2014.11.12,
CN 102263700 A, 2011.11.30,
US 7760719 B2, 2010.07.20,
US 2004090966 A1, 2004.05.13,

(43)申请公布日 2016.07.20

(73)专利权人 南京中兴新软件有限责任公司
地址 210012 江苏省南京市雨花台区宁南
街道紫荆花路68号

审查员 王鑫

(72)发明人 谢密 张启申 孟庆明

(74)专利代理机构 北京康信知识产权代理有限
责任公司 11240
代理人 余刚 梁丽超

(51)Int.Cl.

H04L 12/803(2013.01)

H04L 12/743(2013.01)

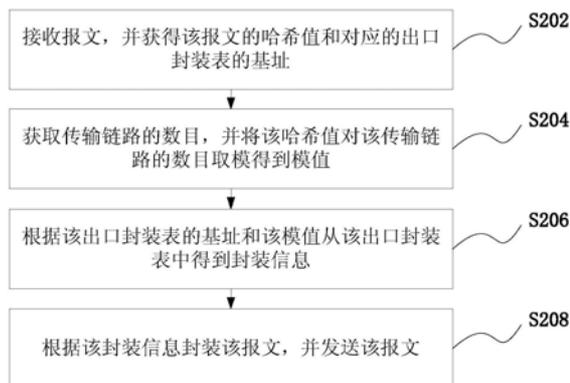
权利要求书2页 说明书6页 附图5页

(54)发明名称

基于等价多路径ECMP的报文传输方法及装置

(57)摘要

本发明公开了一种基于等价多路径ECMP的报文传输方法及装置,其中,该方法包括接收报文,并获得该报文的哈希值和对应的出口封装表的基址;获取传输链路的数目,并将该哈希值对该传输链路的数目取模得到模值;根据该出口封装表的基址和该模值从该出口封装表中得到封装信息;根据该封装信息封装该报文,并发送该报文,以至少解决现有技术中传输链路无法均匀分担报文传输的问题。



1. 一种基于等价多路径ECMP的报文传输方法,其特征在于,包括:
接收报文,并获得所述报文的哈希值和对应的出口封装表的基址;
获取传输链路的数目,并将所述哈希值对所述传输链路的数目取模得到模值;
根据所述出口封装表的基址和所述模值从所述出口封装表中得到封装信息;
根据所述封装信息封装所述报文,并发送所述报文。

2. 根据权利要求1所述的方法,其特征在于,根据所述出口封装表的基址和所述模值从所述出口封装表中得到封装信息包括:

获取分担策略表的基址,其中,所述分担策略表包括所述传输链路在所述出口封装表的偏移地址;

根据所述分担策略表的基址和所述模值从所述分担策略表中得到所述出口封装表的偏移地址;

根据所述出口封装表的基址和所述偏移地址从所述出口封装表中得到封装信息。

3. 根据权利要求2所述的方法,其特征在于,所述获取传输链路的数目包括:

从存储的传输链路数目表中获取所述传输链路的数目。

4. 根据权利要求3所述的方法,其特征在于,在所述传输链路故障时,更新所述传输链路数目表,并将所述分担策略表中所述传输链路对应的所述偏移地址删除。

5. 根据权利要求2至4任一项所述的方法,其特征在于,根据所述分担策略表的基址和所述模值从所述分担策略表中得到所述出口封装表的偏移地址包括:

根据所述分担策略表的基址和所述模值得到所述分担策略表索引;

根据所述分担策略表索引查询所述分担策略表索引对应的分担策略表,并得到所述出口封装表的偏移地址,其中,所述分担策略表存储有所述分担策略表索引和所述出口封装表的偏移地址的对应关系;

根据所述出口封装表的基址和所述偏移地址从所述出口封装表中得到封装信息包括:

根据所述出口封装表的基址和所述偏移地址得到所述出口封装表索引;

根据所述出口封装表索引查询所述出口封装表索引对应的出口封装表,并得到所述封装信息,其中,所述出口封装表存储有所述出口封装表索引和所述封装信息的对应关系。

6. 一种基于等价多路径ECMP的报文传输装置,其特征在于,包括:

获取模块,用于接收报文,并获得所述报文的哈希值和对应的出口封装表的基址,以及获取传输链路的数目;

处理模块,用于将所述获取模块获取的哈希值对所述获取模块获取的传输链路的数目取模得到模值;

所述获取模块还用于,根据所述获取模块获取的出口封装表的基址和所述处理模块得到的模值从所述出口封装表中得到封装信息;

封装发送模块,用于根据所述获取模块获取的封装信息封装所述报文,并发送所述报文。

7. 根据权利要求6所述的装置,其特征在于,所述获取模块还用于,获取分担策略表的基址,其中,所述分担策略表包括所述传输链路在所述出口封装表的偏移地址,并根据所述分担策略表的基址和所述模值从所述分担策略表中得到所述出口封装表的偏移地址,并根据所述出口封装表的基址和所述偏移地址从所述出口封装表中得到封装信息。

8. 根据权利要求7所述的装置,其特征在于,所述获取模块还用于,从存储的传输链路数目表中获取所述传输链路的数目。

9. 根据权利要求8所述的装置,其特征在于,所述处理模块还用于,在所述传输链路故障时,更新所述传输链路数目表,并将所述分担策略表中所述传输链路对应的所述偏移地址删除。

10. 根据权利要求7至9任一项所述的装置,其特征在于,

所述获取模块还用于,根据所述分担策略表的基址和所述模值得到所述分担策略表索引,并根据所述分担策略表索引查询所述分担策略表索引对应的分担策略表,并得到所述出口封装表的偏移地址,其中,所述分担策略表存储有所述分担策略表索引和所述出口封装表的偏移地址的对应关系;

所述获取模块还用于,根据所述出口封装表的基址和所述偏移地址得到所述出口封装表索引,并根据所述出口封装表索引查询所述出口封装表索引对应的出口封装表,并得到所述封装信息,其中,所述出口封装表存储有所述出口封装表索引和所述封装信息的对应关系。

基于等价多路径ECMP的报文传输方法及装置

技术领域

[0001] 本发明涉及通信领域,具体而言,涉及一种基于ECMP (Equal Cost Multi-Path,等价多路径)的报文传输方法和装置。

背景技术

[0002] ECMP (Equal Cost Multi-Path,等价多路径)是改进IP网络中基于最短路径的内部网关路由协议不足的一种技术。在传统最短路径算法中,如果两条或多条最短路径同时存在,算法会选择最优的一条。而ECMP作了改进,流量在多条等价路径中分配。在MPLS (Multi-protocol Label Switching,多协议标签交换)网络部署ECMP后,需要在链路故障时达到电信级的保护倒换要求。

[0003] 在相关技术中,图1是根据相关技术中的一种基于ECMP的报文传输的方法的流程图,如图1所示,报文传输装置通过对报文目的MAC (Media Access Control,介质访问控制)、源MAC、目的IP、源IP等字段进行哈希后获得哈希值,出口封装表中存在多个连续条目,对应多条等价转发路径,当报文目的MAC、源MAC、目的IP、源IP等字段变化时,经过哈希模块得到不同的哈希值,通过出口封装表基址加上哈希值,就可以对应到不同的出口封装表条目,以达到将报文分担到多条路径上转发的目的。当MPLS (Multiple Protocol Label Switch,多协议标签交换协议)网络中形成N个MPLS ECMP时,所有MPLS ECMP共用分担链路。对不同的分担链路数,现有技术不能直接取全部哈希值,只能舍弃一部分比特位,按实际链路数来取,比如实际链路数为32时,取的比特位数为5 (2的5次方),可能哈希值变化不在取的5bit以内,即对不同的报文字段做哈希后,哈希值的低5bit不变,这样会使得链路分担不均,造成分担流量精度超标。

发明内容

[0004] 本发明提供了一种基于ECMP的报文传输方法及装置,以至少解决相关技术中传输链路无法均匀分担报文传输的问题。

[0005] 根据本发明的一个方面,提供了一种基于ECMP的报文传输方法,包括:接收报文,并获得所述报文的哈希值和对应的出口封装表的基址;获取传输链路的数目,并将所述哈希值对所述传输链路的数目取模得到模值;根据所述出口封装表的基址和所述模值从所述出口封装表中得到封装信息;根据所述封装信息封装所述报文,并发送所述报文。

[0006] 优选地,根据所述出口封装表的基址和所述模值从所述出口封装表中得到封装信息包括:获取分担策略表的基址,其中,所述分担策略表包括所述传输链路在所述出口封装表的偏移地址;根据所述分担策略表的基址和所述模值从所述分担策略表中得到所述出口封装表的偏移地址;根据所述出口封装表的基址和所述偏移地址从所述出口封装表中得到封装信息。

[0007] 优选地,所述获取传输链路的数目包括:从存储的传输链路数目表中获取所述传输链路的数目。

[0008] 优选地,在所述传输链路故障时,更新所述传输链路数目表,并将所述分担策略表中所述传输链路对应的所述偏移地址删除。

[0009] 优选地,根据所述分担策略表的基址和所述模值得到所述分担策略表中得到所述出口封装表的偏移地址包括:根据所述分担策略表的基址和所述模值得到所述分担策略表索引;根据所述分担策略表索引查询所述分担策略表索引对应的分担策略表,并得到所述出口封装表的偏移地址,其中,所述分担策略表存储有所述分担策略表索引和所述出口封装表的偏移地址的对应关系;根据所述出口封装表的基址和所述偏移地址从所述出口封装表中得到封装信息包括:根据所述出口封装表的基址和所述偏移地址得到所述出口封装表索引;根据所述出口封装表索引查询所述出口封装表索引对应的出口封装表,并得到所述封装信息,其中,所述出口封装表存储有所述出口封装表索引和所述封装信息的对应关系。

[0010] 根据本发明的另一方面,提供了一种基于ECMP的报文传输装置,包括:获取模块,用于接收报文,并获得所述报文的哈希值和对应的出口封装表的基址,以及获取传输链路的数目;处理模块,用于将所述获取模块获取的哈希值对所述获取模块获取的传输链路的数目取模得到模值;所述获取模块还用于,根据所述获取模块获取的出口封装表的基址和所述处理模块得到的模值得到所述出口封装表中得到封装信息;封装发送模块,用于根据所述获取模块获取的封装信息封装所述报文,并发送所述报文。

[0011] 优选地,所述获取模块用于,获取分担策略表的基址,其中,所述分担策略表包括所述传输链路在所述出口封装表的偏移地址,并根据所述分担策略表的基址和所述模值得到所述分担策略表中得到所述出口封装表的偏移地址,并根据所述出口封装表的基址和所述偏移地址从所述出口封装表中得到封装信息。

[0012] 优选地,所述获取模块用于,从存储的传输链路数目表中获取所述传输链路的数目。

[0013] 优选地,所述处理模块还用于,在所述传输链路故障时,更新所述传输链路数目表,并将所述分担策略表中所述传输链路对应的所述偏移地址删除。

[0014] 优选地,所述获取模块用于,根据所述分担策略表的基址和所述模值得到所述分担策略表索引,并根据所述分担策略表索引查询所述分担策略表索引对应的分担策略表,并得到所述出口封装表的偏移地址,其中,所述分担策略表存储有所述分担策略表索引和所述出口封装表的偏移地址的对应关系;所述获取模块用于,根据所述出口封装表的基址和所述偏移地址得到所述出口封装表索引,并根据所述出口封装表索引查询所述出口封装表索引对应的出口封装表,并得到所述封装信息,其中,所述出口封装表存储有所述出口封装表索引和所述封装信息的对应关系。

[0015] 通过本发明,采用报文传输装置接收报文,并获得该报文的哈希值和对应的出口封装表的基址,获取传输链路的数目,并将该哈希值对该传输链路的数目取模得到模值,并根据该出口封装表的基址和该模值得到所述出口封装表中得到封装信息,并根据该封装信息封装该报文,并发送该报文,解决了现有技术中传输链路无法均匀分担报文传输的问题,从而提高负荷分担精度。

附图说明

[0016] 此处所说明的附图用来提供对本发明的进一步理解,构成本申请的一部分,本发

明的示意性实施例及其说明用于解释本发明,并不构成对本发明的不当限定。在附图中:

- [0017] 图1是根据相关技术中的一种基于ECMP的报文传输的方法的流程图;
- [0018] 图2是根据本发明实施例的一种基于ECMP的报文传输的方法的流程图;
- [0019] 图3a是根据本发明实施例的一种分担策略表的示意图;
- [0020] 图3b是根据本发明实施例的一种传输链路数目表的示意图;
- [0021] 图4是根据本发明实施例的另一种基于ECMP的报文传输的方法的流程图;
- [0022] 图5是根据本发明实施例的一种报文传输装置进行报文传输的示意图;
- [0023] 图6是根据本发明实施例另一种报文传输装置进行报文传输的示意图;
- [0024] 图7是根据本发明实施例的一种报文传输装置的结构示意图。

具体实施方式

[0025] 下文中将参考附图并结合实施例来详细说明本发明。需要说明的是,在不冲突的情况下,本申请中的实施例及实施例中的特征可以相互组合。

[0026] 在本实施例中提供了一种基于ECMP的报文传输的方法,图2是根据本发明实施例的一种基于ECMP的报文传输的流程图,如图2所示,该流程包括如下步骤:

[0027] 步骤S202,接收报文,并获得该报文的哈希值和对应的出口封装表的基址;

[0028] 步骤S204,获取传输链路的数目,并将该哈希值对该传输链路的数目取模得到模值;

[0029] 步骤S206,根据该出口封装表的基址和该模值从该出口封装表中得到封装信息;

[0030] 步骤S208,根据该封装信息封装该报文,并发送该报文。

[0031] 通过上述步骤,解决了现有技术中传输链路无法均匀分担报文传输的问题,从而提高符合分担精度,提高了负荷分担精度。

[0032] 优选地,获取分担策略表的基址,根据该分担策略表的基址和该模值从该分担策略表中得到该出口封装表的偏移地址,根据该出口封装表的基址和该偏移地址从该出口封装表中得到封装信息,这样,可以实现多个ECMP组对应一个分担策略表。

[0033] 其中,图3a是根据本发明实施例的一种分担策略表的示意图,如图3a所示,该分担策略表包括该传输链路在该出口封装表的偏移地址。

[0034] 需要说明的是,该分担策略表包含连续M条表项(M为传输链路数目),该表基于链路接口组合与链路分担权重聚合,即链路接口相同且分担权重相同的ECMP业务共用一个分担策略表。

[0035] 优选地,从存储的传输链路数目表中获取该传输链路的数目。

[0036] 其中,图3b是根据本发明实施例的一种传输链路数目表的示意图,如图3b所示,该传输链路数目表保存了传输链路数目,该表基于链路接口组合与链路分担权重聚合,即链路接口相同且分担权重相同的ECMP业务共用一个传输链路数目表。

[0037] 优选地,在该传输链路故障时,更新该传输链路数目表,并将该分担策略表中该传输链路对应的该偏移地址删除,这样,虽然有多个ECMP组,但切换一个分担策略表即可,实现多到一的切换,避免了对每个出口封装表的切换,显著提高了切换性能。

[0038] 优选地,根据该分担策略表的基址和该模值得到该分担策略表索引,根据该分担策略表索引查询该分担策略表索引对应的分担策略表,并得到该出口封装表的偏移地址,

根据该出口封装表的基址和该偏移地址从该出口封装表中得到封装信息包括:其中,该分担策略表存储有该分担策略表索引和该出口封装表的偏移地址的对应关系。

[0039] 优选地,根据该出口封装表索引查询该出口封装表索引对应的出口封装表,并得到该封装信息,其中,该出口封装表存储有该出口封装表索引和该封装信息的对应关系。

[0040] 本发明实施例提供一种基于ECMP的报文传输方法,图4是根据本发明实施例的另一种基于ECMP的报文传输的方法的流程图,如图4所示,本发明实施例的执行主体可以是报文传输装置,该报文传输装置可以是路由器,本发明对此不作限定,该方法的流程包括:

[0041] S401,报文传输装置接收报文,并获取该报文的哈希值和对应的出口封装表的基址。

[0042] 其中,报文处理设备获取报文目的MAC、源MAC、目的IP、源IP等字段经过哈希得到哈希值。

[0043] S402,报文传输装置从存储的传输链路数目表中获取传输链路的数目,并将该哈希值对该传输链路的数目取模得到模值。

[0044] 例如,本发明取全部16bit(比特)的哈希值对分担链路数8做模运算(即取余),由于模运算的特性,做模运算后的值不会超过8,避免了所取的部分哈希值不变或变化不均匀。

[0045] S403,报文传输装置根据该分担策略表的基址和该模值得到该分担策略表索引。

[0046] S404,报文传输装置根据该分担策略表索引查询该分担策略表索引对应的分担策略表,并得到该出口封装表的偏移地址。

[0047] 其中,该分担策略表存储有该分担策略表索引和该出口封装表的偏移地址的对应关系。

[0048] S405,报文传输装置根据该出口封装表的基址和该偏移地址得到该出口封装表索引。

[0049] S406,报文传输装置根据该出口封装表索引查询该出口封装表索引对应的出口封装表,并得到该封装信息。

[0050] 其中,该出口封装表存储有该出口封装表索引和该封装信息的对应关系。

[0051] S407,报文传输装置根据该封装信息封装该报文,并发送该报文。

[0052] 需要说明的是,由于不同的ECMP组出口封装信息不一样(即MPLS标签不一样),因此不同的出口封装表对应不同的ECMP组。

[0053] 例如,1000条路由,每条路由从8条分担链路分别学习到,这条路由形成负荷分担,出口有8个,每个出口对应的出向MPLS标签(封装到出去的报文中)不同,即每条路由对应8个出向MPLS标签,1000条路由就对应1000个出口封装表,每个出口封装表中包含8个条目,每个条目中保存一个出向MPLS标签。

[0054] 综上,图5是根据本发明实施例的一种报文传输装置进行报文传输的示意图,如图5所示,图5为本发明实施例中该报文传输装置进行报文传输的示意图,报文处理设备获取报文目的MAC、源MAC、目的IP、源IP等字段经过哈希得到哈希值,并读取传输链路数目表获取传输链路数;将得到的哈希值对传输链路数取模,得到模值,由于模运算的特性,模值不会超过传输链路数,根据分担策略表的基址加上得到的模值(模值即分担策略表偏移)得到分担策略表索引,查到对应的分担策略表表项,得到出口封装表偏移,根据出口封装表基址

加上得到的偏移值得到出口封装表索引,查到对应的出口封装表表项,得到出口封装信息,并根据该封装信息分装报文并发送。

[0055] 在本发明实施例中,在该传输链路故障时,更新该传输链路数目表,并将该分担策略表中该传输链路对应的该偏移地址删除,这样,虽然有多个ECMP组,但切换一个分担策略表即可,实现多到一的切换,避免了对每个出口封装表的切换,显著提高了切换性能。

[0056] 例如,图6是根据本发明实施例另一种报文传输装置进行报文传输的示意图,如图6所示,N个ECMP组共用一个8条链路(分别为链路1至链路8)的接口链路组,并且共用一个分担策略表,此时,该传输链路数目表中记录的值为8,该分担策略表分别依次记载该8条链路的偏移地址,当第3条链路发生故障时,报文传输装置将分担策略表第3条链路对应的偏移地址删除,并将传输链路数目表中记录的数值由8调整为7,这样,只需要切换一个分担策略表即可,从而提高了切换性能。

[0057] 采用上述实施方式,解决了现有技术中传输链路无法均匀分担报文传输的问题,从而提高符合分担精度,提高了负荷分担精度,另外,在该传输链路故障时,只切换一个分担策略表即可,实现多到一的切换,避免了对每个出口封装表的切换,显著提高了切换性能。

[0058] 在本实施例中还提供了一种报文传输装置,该装置用于实现上述实施例及优选实施方式,已经进行过说明的不再赘述。如以下所使用的,术语“模块”可以实现预定功能的软件和/或硬件的组合。尽管以下实施例所描述的装置较佳地以软件来实现,但是硬件,或者软件和硬件的组合的实现也是可能并被构想的。

[0059] 图7是根据本发明实施例的一种报文传输装置的结构框图,如图7所示,该装置70包括:

[0060] 获取模块71,用于接收报文,并获得该报文的哈希值和对应的出口封装表的基址,以及获取传输链路的数目;处理模块72,用于将该获取模块71获取的哈希值对该获取模块71获取的传输链路的数目取模得到模值;该获取模块71还用于,根据该获取模块71获取的出口封装表的基址和该处理模块得到的模值从该出口封装表中得到封装信息;封装发送模块73,用于根据该获取模块71获取的封装信息封装该报文,并发送该报文。

[0061] 优选地,该获取模块71,还用于获取分担策略表的基址,其中,该分担策略表包括该传输链路在该出口封装表的偏移地址,并根据该分担策略表的基址和该模值得到该分担策略表中得到该出口封装表的偏移地址,并根据该出口封装表的基址和该偏移地址从该出口封装表中得到封装信息。

[0062] 优选地,该获取模块71,还用于从存储的传输链路数目表中获取该传输链路的数目。

[0063] 优选地,该处理模块72,还用于在该传输链路故障时,更新该传输链路数目表,并将该分担策略表中该传输链路对应的该偏移地址删除。

[0064] 优选地,该获取模块71,还用于根据该分担策略表的基址和该模值得到该分担策略表索引,并根据该分担策略表索引查询该分担策略表索引对应的分担策略表,并得到该出口封装表的偏移地址,其中,该分担策略表存储有该分担策略表索引和该出口封装表的偏移地址的对应关系;

[0065] 该获取模块71,还用于根据该出口封装表的基址和该偏移地址得到该出口封装表

索引,并根据该出口封装表索引查询该出口封装表索引对应的出口封装表,并得到该封装信息,其中,该出口封装表存储有该出口封装表索引和该封装信息的对应关系。

[0066] 采用上述装置,解决了现有技术中传输链路无法均匀分担报文传输的问题,从而提高符合分担精度,提高了负荷分担精度。

[0067] 综上所述,本发明的目的是通过对从哈希模块得到的哈希值取更多的比特位进行模运算,提高负荷分担精度,即使各链路均匀分担;增加分担策略表,将现有技术中出口封装表中公共的部分,即公共的链路独立出来,放到分担策略表中,这样多个ECMP组因为分担链路一样,对应同一个分担策略表,当发生链路故障时,虽然有多个ECMP组,但切换一个分担策略表即可,实现多到一的切换,能显著提高切换性能。

[0068] 显然,本领域的技术人员应该明白,上述的本发明的各模块或各步骤可以用通用的计算装置来实现,它们可以集中在单个的计算装置上,或者分布在多个计算装置所组成的网络上,可选地,它们可以用计算装置可执行的程序代码来实现,从而,可以将它们存储在存储装置中由计算装置来执行,并且在某些情况下,可以以不同于此处的顺序执行所示出或描述的步骤,或者将它们分别制作成各个集成电路模块,或者将它们中的多个模块或步骤制作成单个集成电路模块来实现。这样,本发明不限制于任何特定的硬件和软件结合。

[0069] 以上所述仅为本发明的优选实施例而已,并不用于限制本发明,对于本领域的技术人员来说,本发明可以有各种更改和变化。凡在本发明的精神和原则之内,所作的任何修改、等同替换、改进等,均应包含在本发明的保护范围之内。

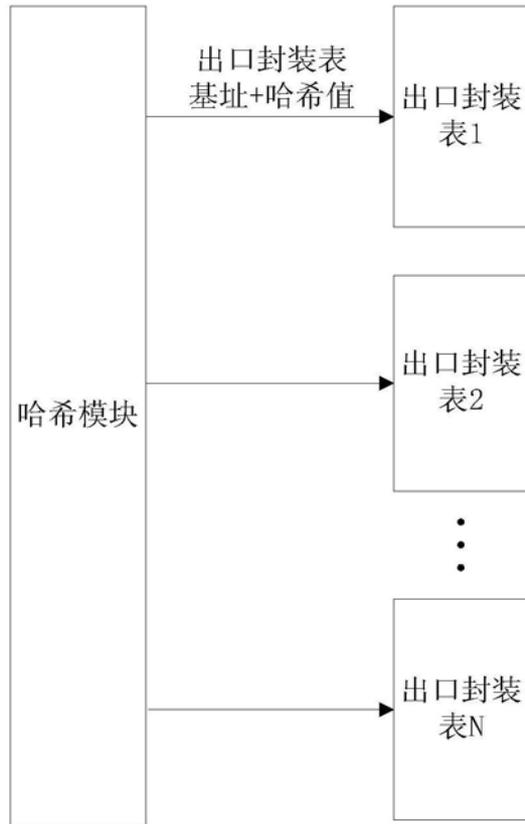


图1

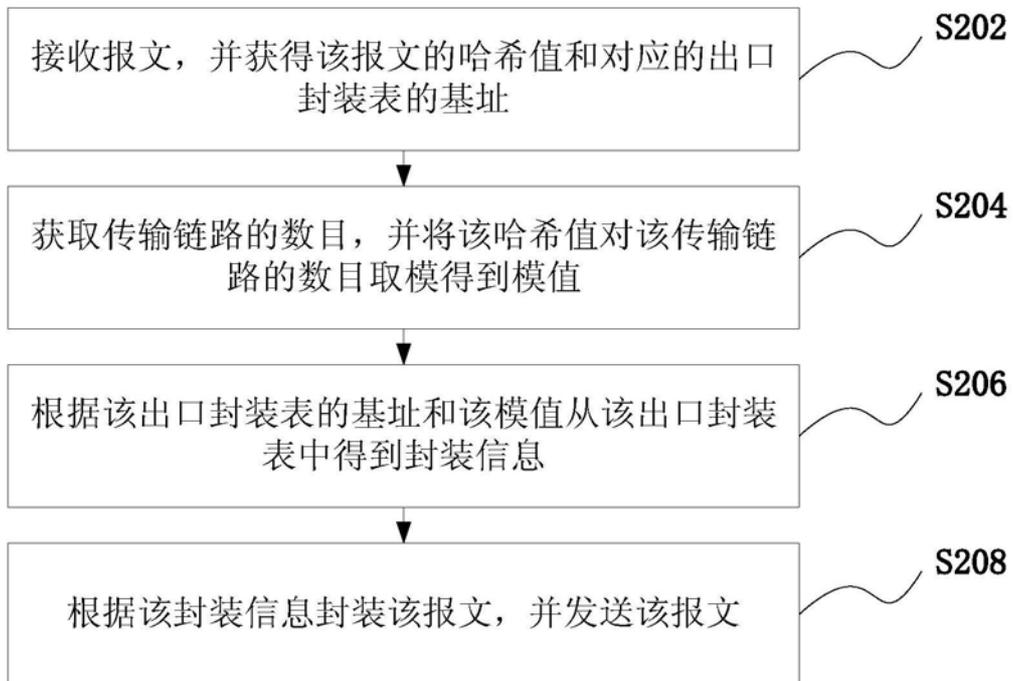


图2

分担策略表表项内容

出口封装表的偏移地址1
出口封装表的偏移地址2
出口封装表的偏移地址3
...
出口封装表的偏移地址M

M为传输链路数目

图3a

传输链路数目表表项内容

传输链路数目

图3b

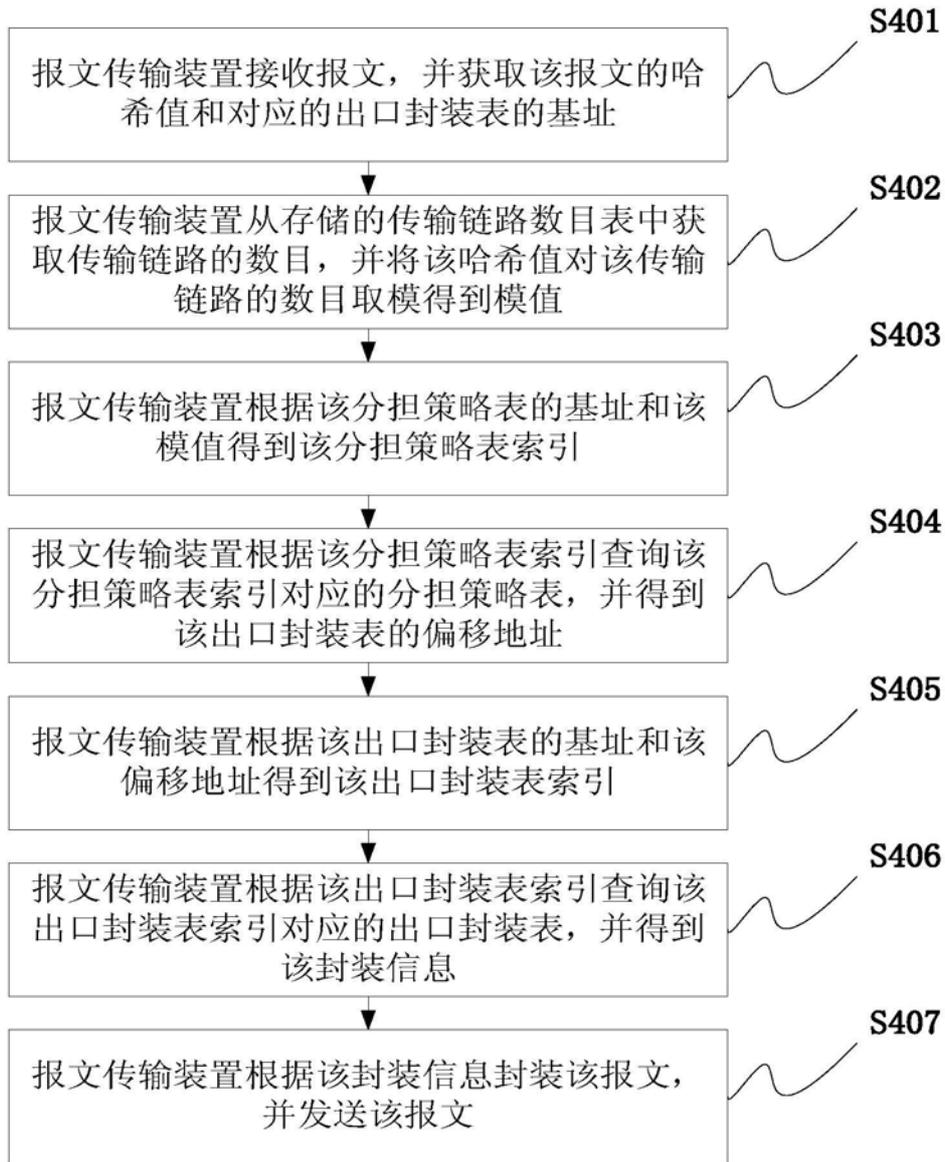


图4

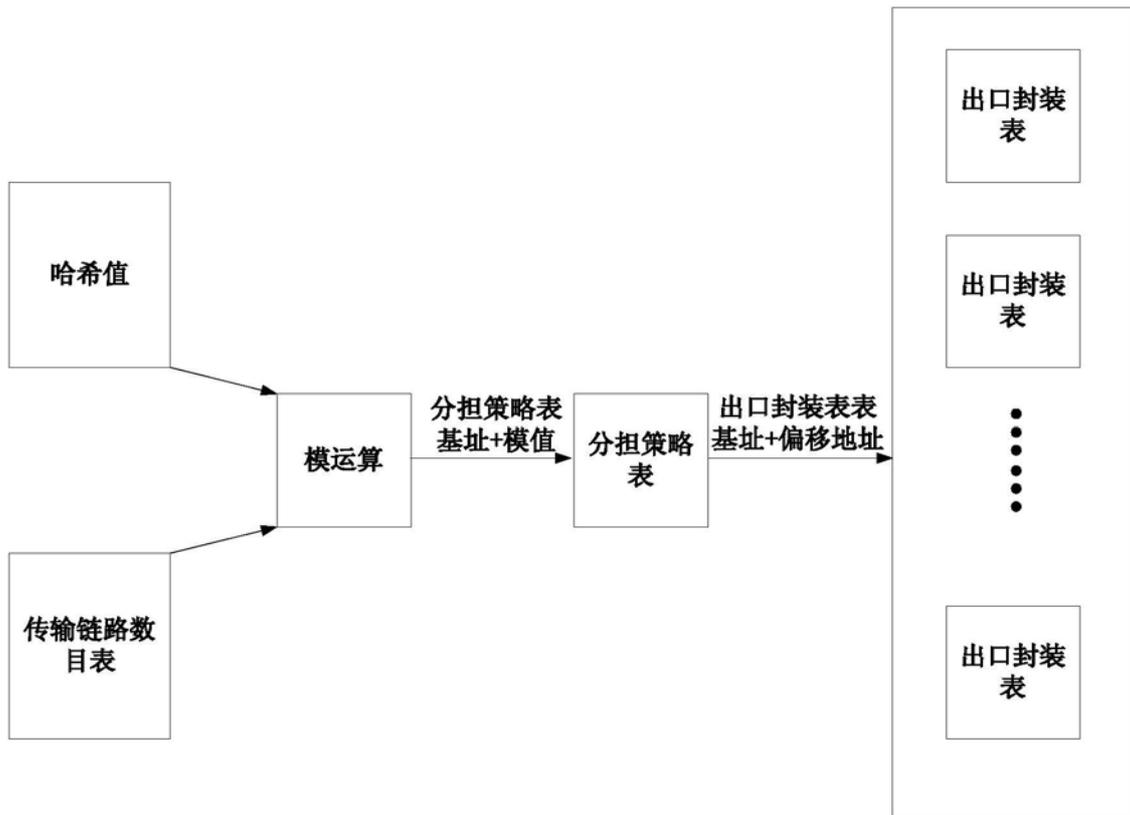


图5

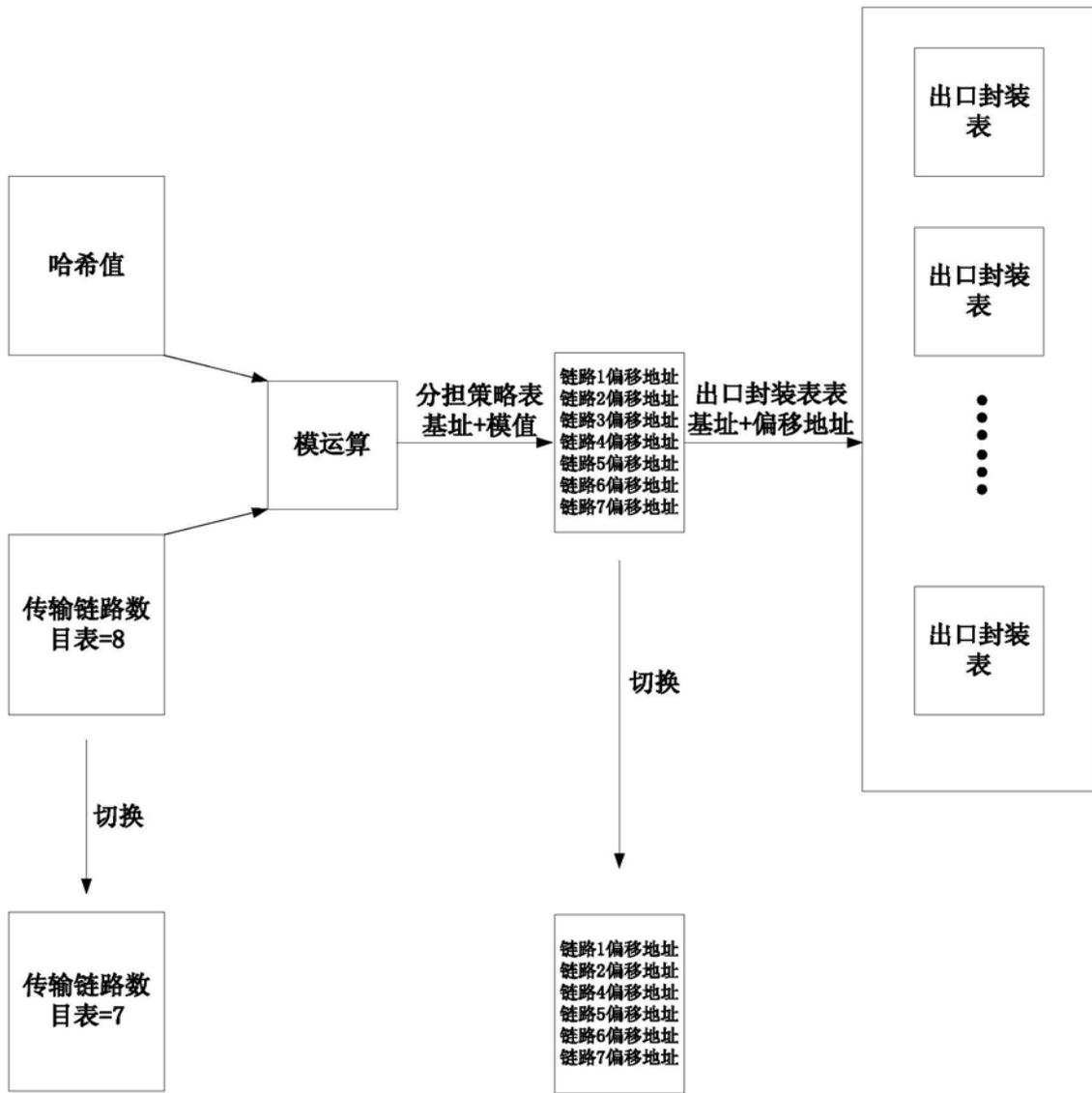


图6

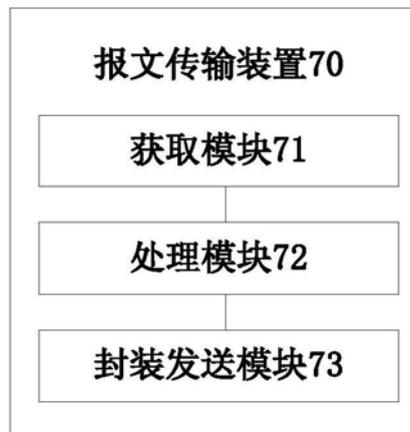


图7