

**专利申请**

**技术交底书**

|  |  |
| --- | --- |
| **交底书名称**  **必须填写** | 一种SDN 控制器中采用分层拓扑模型实现SR路径计算和SIDList生成的方法 |
| **技术联系人姓名**  **及其电话、email**  **必须填写** | 卫岳民 18069762058 [ymwei@utstar.com](mailto:ymwei@utstar.com) |
| **经办人姓名**  **及其电话、email**  **必须填写**  **如不填写，则默认技术联系人为负责人** | 卫岳民 18069762058 [ymwei@utstar.com](mailto:ymwei@utstar.com) |
| **备注**  **事项** | 1. 是否仅用于申请政府高新资质或政府项目 2. 是否将来申请国外专利   **一份好的技术交底书有助于代理人撰写高质量的专利申请文件，从而获得更好的授权前景和保护范围。** |

|  |
| --- |
| **注意事项**  1.技术联系人应为深入了解本申请技术方案的技术人员，如交底书撰写人，负责向专利审核人员和代理人解释技术细节、修改交底书、审核申请文件等工作,请务必填写技术联系人的姓名、E-mail、手机。  2.经办人应为管理本案知识产权的人，例如指示本案的申请人是谁，本案的发明人是谁，是否申请费用减缓，确认本案是否可以提交等。请务必填写经办人的姓名、E-mail、手机。  3.专利申请不要求已具体实现或实施，形成完整的技术方案即可提交申请，特别是需要向合作方公开、向标准提案或以其他方式公开的重要技术构思应在公开前尽早申请。 |

一、与本专利最接近的现有技术：

1、现有技术的方案简述（只需要描述与本专利相关的内容）：

要求提供现有技术的技术方案图，针对图示做详细的文字说明，以使代理人了解本发明（实用新型）的技术背景。

专业缩写词列表：

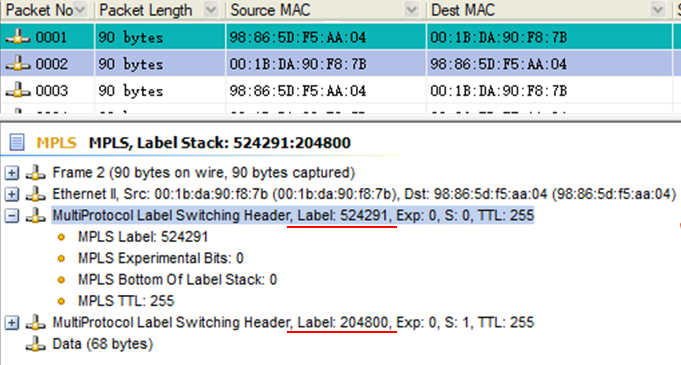
|  |  |  |
| --- | --- | --- |
| 专业缩写词 | 英文解释 | 中文解释 |
| SDN | Software Defined Network | 软件定义网络 |
| PCE | Path Computation Element | 路径计算单元 |
| IGP | Interior Gateway Protocol | 内部网关协议 |
| TED | Traffic Engineer Database | 流量工程数据库 |
| FIB | Forwarding Information Base | 转发信息表 |
| SR | Segment Routing | 分段路由 |
| SRTE | SR Traffic Engine | 分段路由流量工程 |
| SPF | Shortest Path First | 最短路径优先 |
| BFS | Breadth First Search | 广度优先算法 |
| CSPF | Constrained Shortest Path First | 约束式最短路径优先 |
| SID | Segment ID | 段标识 |
| Node SID | Node SID | 节点标识 |
| AdjSID | Adjacency SID | 邻接标识 |
| SIDList | SID List | SID列表 |
| SPT | Shortest Path Tree | 最短路径生成树 |
| MPLS | Multi-Protocol Label Switching | 多协议标记交换 |
| RSVP-TE | Resource Reservation Protocol | 资源预留协议 |
| L3 Topo | Layer3 Topology | 三层链路拓扑 |
| IP SPT Topo | IP SPT Topology | IP可达链路拓扑 |
| SRLG | Shared Risk Link Group | 风险共享链路组 |
|  |  |  |

**源路由技术SegmentRouting简介**

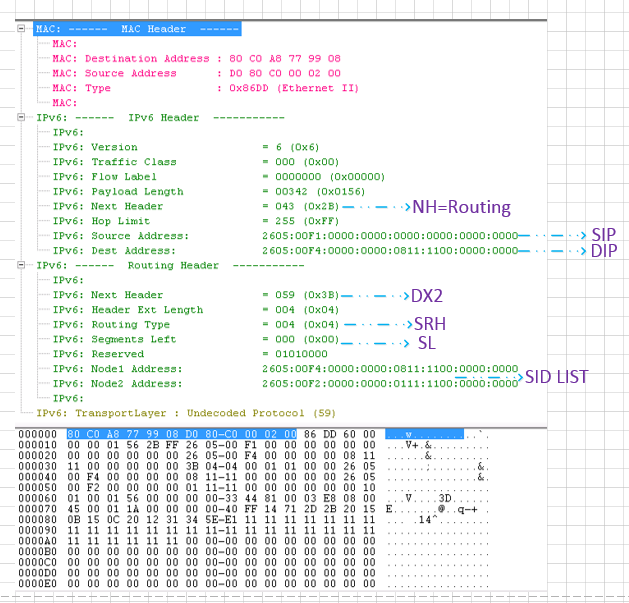
源路由技术指数据包在源节点即确定了转发路径，这种技术为网络应用提供了灵活可扩展的编程能力，不同应用需要业务报文在网络上走不同的路径，这在当今网络需求多样化的情况下尤其重要，比如低时延的大规模工业物联网应用，大带宽的电商大促流量引导，远程手术医疗的实时高清网络保证，等等。SegmentRouting（简称SR）就是一种源路由技术，支持SR的网络设备结合掌控全局网络状态的可编程SDN控制器，是当今各类5G业务传送的一种网络技术手段。

SegmentRouting名称直译就是分段路由，将业务报文的转发路径按需求在网络上做分段传送。分段标识用SegmentID表示，当前SR分段路由的转发面技术架构方案（RFC8402）主要有SR-MPLS（RFC8660）和SRv6（RFC8754），对应的SID值是32bit的MPLS标签，或是128bit的IPv6。SID在网络中可表示节点标识（NodeSID）或接口（AdjSID），将SID组织成SIDList压入发送的报文头部，就确定了该报文在网络中的转发路径。

下图是SR-MPLS报文的样例，可以看到以MPLS标签作为SID，SIDList和传统的IP-MPLS 标签栈没有区别。



下图是SRv6报文的样例，因SRH里的SID以128bit的IPv6格式表示，所以SRv6的SIDList占用更多空间，虽然SRv6的SID结构带来了更多的灵活性，但报文开销也更大。



SR源路由技术突破了传统IP网络的报文转发只能由IGP路由控制，结合SDN控制器可对报文按需做路径调整，为网络可编程提供了技术手段。正如之前说到的，5G时代业务的差异化需求以及网络的动态变化，我们需要网络能按需适时调配业务路径。举个例子，电商大促时有这样的流量需求：DataCenter间需要走足够带宽保证且无拥塞，如下图。



图表 1 多条件下的算路样例

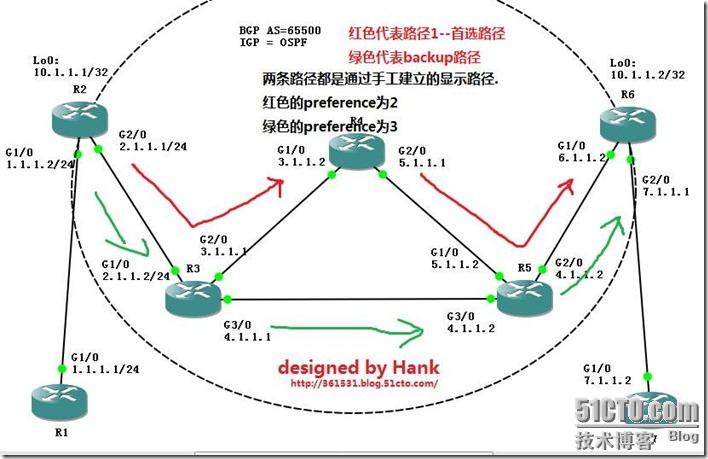
**现有的网络寻路算法介绍**

* + - 1. **IGP路由计算**

IGP路由协议（ISIS，OSPF）通过路由学习得到网络上链路状态信息，由最短生成树算法，计算本节点到网络上其他节点的最短路径。IP路由器中生成最短路径树SPT的基本算法是最短路径算法，即迪杰斯特拉(**Dijkstra**)算法 （当然路由器还用Dijkstra的改进算法，如I-SPF/ Incremental SPF，PRC/ Partial Route Calculation，来提高路由计算效率）。它的主要特点是以起始点为中心向外层层扩展(广度优先搜索思想)，直到扩展到终点为止。

* + - 1. **CSPF在MPLS流量工程应用**

在传统IP-MPLS网络支持MPLS TE的环境中，提供了SPF算法的增强版-CSPF，即带条件的最短路径算法。这些路径条件包括：链路预留带宽，链路属性/亲和属性，权重。 同时MPLS TE对源宿路径的选择也提供显性路径的设置，包括松散路径和严格路径。 如下图网络环境：（参考自https://blog.51cto.com/361531/657115）



下面是路由R2上配置MPLS TE隧道

ip explicit-path name work enable

next-address 2.1.1.2

next-address 3.1.1.2

next-address 5.1.1.2

next-address 6.1.1.2

ip explicit-path name backup enable

next-address 2.1.1.2

next-address 4.1.1.2

next-address 6.1.1.2

interface tunnel0

ip unnumbered loopback0

tunnel destination 10.1.1.2

tunnel model mpls traffic-eng

tunnel mpls traffic-eng autoroute announce

tunnel mpls traffic-eng priority 7 7

tunnel mpls traffic-eng bandwidth 1024

tunnel mpls traffic-eng path-option 2 explicit name work

tunnel mpls traffic-eng path-option 3 explicit name backup

tunnel mpls traffic-eng path-option 10 dynamic

no routing dynamic

从MPLS TE的CSPF算路能力看，它能提供图1 网络环境下的算路需求。

2、现有技术的客观缺点（通过本专利能够解决或改善的一个或多个缺点）：

需结合现有技术方案做推理说明有技术存在的缺点，且该缺点是可以通过本发明（实用新型）中的某个技术手段解决的。

IGP的SPF算法，条件因素过于简单，只能按链路代价计算出最短路径。存在的不足有：

* 相同源宿的业务路径相同，可能导致业务流量一边十分拥挤，而一边路径的空余带宽又无法利用。当网络拥塞发生时，新业务路径也无法避绕带宽瓶颈网段。
* 链路代价仅用一个IGP metric值表示，不能用于对多种链路综合条件的描述，比如链路可用带宽值，链路时延值，亲和属性值，权重，链路带宽利用率， 等等。不能满足不同业务走不同特性路径的要求。

MPLS TE结合CSPF算法，能部分满足图1的寻路要求，但还是存在以下不足：

* 网元实现复杂。MPLS TE网络中的各个网元需要支持RSVP-TE（Resource Reservation Protocol – Traffic Engineering）协议，需支持IGP协议ISIS/OSPF对TE的扩展，并维护庞大的链路信息数据库。故对网元的性能要求高，相应的实现成本也高。
* MPLS TE部署复杂，创建，维护，排障都需要人力参与。 在上述的配置样例图可以看到TE需要在网内每个节点每个端口都要设置调配一堆参数。所以导致现网运维过于繁重。
* 静态配置不能灵活调整，采用RSVP为预留带宽创建的路径不能反映实际使用情况，当网路流量发生拥塞时，新隧道业务带宽只要在静态预留配置带宽范围内还是会走拥塞路径。

SR转发带来的问题：

当前，不论是移动的SPN SRTP隧道还是联通的IPRAN SRTE隧道，打SIDList都是基于AdjSID的，即按每跳链路打邻接标签，这样会导致标签过长

在源节点上需将完整路径都表示在SIDList上，SIDList标签栈过长将导致：

* 报文净荷承载效率低
* 网元转发芯片无法支持过长的SIDList，有深度限制

二、本专利的技术：

**（1）具体方案**

（描述的重点：与现有技术对比，本专利的差别以及如何通过这些差别实现客观改进效果）。

如果是计算机软件发明，请附上反映方案流程的流程图。

如果是产品，请附上计算机绘图（应该是线条描绘）以反映产品的各个部件。

请配合附图来描述各部件的连接关系（可结合附图中的数字标记描述），各步骤次序。应当详细到图中有标号的每个部件和每个步骤都要描述到。

上述技术方案，不能只做功能介绍，而应当说明是怎样实现的？技术方案的阐述以本领域技术人员不需付出创造性劳动即可实现为准。

实用新型必须提供结构图，图上需标记部件名称，同时结合图进行文字说明（工艺步骤、结构说明、原理说明、动作关系说明）。发明专利必要时提供图，图的文字说明同实用新型。

SDN架构为复杂的流量工程实现开辟了新的思路，因为SDN控制器掌握了网络的运行状态信息，而且控制器一般部署在比网元硬件强大许多倍的服务器上，不论是对海量网络信息库的存储，还是对付大拓扑数据的路径计算能力，都是单网元所不能匹及的。

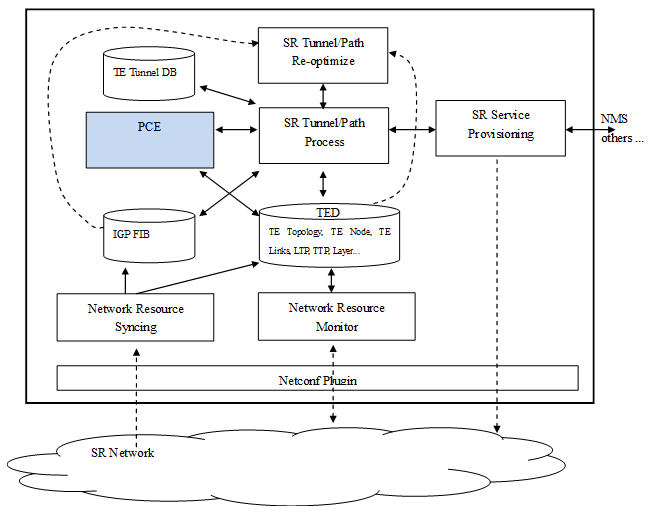
下图是支持的PCE算路功能的SDN控制器架构图，其中的关键模块：

PCE：提供多种路径算法，其中Dijkstra, BFS, K-Shortest, 等

TED：流量工程数据库，包括各类TE拓扑数据，是算路的数据基础

IGP FIB：同步于网元上的IP转发表，使控制器更新最新转发状态

SR Tunnel/Path Process: SRTE隧道业务处理模块



 SegmentRouting网络中报文的转发按SID指令走，而SID又分为NodeSID（对应于节点IP前缀）和AdjSID（对应于本地三层接口IP），所以SR网络的转发实际上还是依赖IP三层转发，段/segment可用理解成IP网络的一段可达链路。

控制器上的分层拓扑模型

底层：L3\_Topo – 三层链路拓扑

由各个网元节点的三层接口互联而形成的拓扑。控制器上的L3 Topo数据来源于网元IGP的邻居表信息。链路表示节点间的IGP邻居连接关系。

上层：IP\_SPT\_Topo - IP可达链路拓扑

由各个网元根据IGP链路状态库LSDB生成的最短路径树SPT拓扑。控制器上的IP\_SPT拓扑数据来源于网元IGP的路由转发表FIB，并和网元做数据同步。IP SPT拓扑的链路表示节点间的IP路由可达关系。

对网元来讲，只需简单地维护本网元SPT拓扑即可，即IP报文转发按传统IGP路由转发；而控制器上维护的L3\_Topo可以赋予丰富的流量工程属性，可以设置三层链路带宽配置参数，反映带宽流量的实时状态参数，设置链路的TE Metric值以及设置链路风险共享链路组SRLG，链路亲和属性，时延等参数；同时，控制器可以将L3拓扑链路上的这些属性也反应到IP SPT拓扑链路上。 这样控制器可以直接在IP STP拓扑上做路径计算，得到SR分段路径结果——最短深度的SIDList。



分层模型UML图，如下：



解释：

L3\_Topo和IP\_SPT\_Topo有相同的拓扑对象结构

L3\_Topo的link和IP\_SPT\_Topo的link 也有相同的对象结构（注：左下面Link简化了）

IP\_SPT\_Topo的under\_layer\_topo\_id指向下层的L3\_Topo

L3\_Topo的upper\_layer\_topo\_id指向上层的IP\_SPT\_Topo

L3\_Topo的相关Link有基础链路Cost

L3\_Topo的上层链路关联到IP\_SPT\_Topo链路

IP\_SPT\_Topo的下层链路关联到L3\_Topo链路

IP\_SPT\_Topo的链路Delay, bandwidth, cost可以由下层L3\_Topo链路做相关计算得到

Segment Routing技术在IETF已有RFC标准化，SR-TE隧道技术也已经部署在国内大运营商网络，如移动SPN网络或联通IPRAN网络。如前所述，SR网络中控制报文转发是通过SID标签栈来实现的，那么如何计算出即符合业务需求，又打最少层数的SIDList，是运维配置SR-TE隧道的关键要求。

下面我们分场景描述如何利用IP\_SPT拓扑模型来计算所需的SIDList

* 场景一: 当源宿节点间的业务报文转发路径需**直接走IGP最短路径**，且源宿间**存在**可达路由。

算法的输入条件：（源宿节点是必输条件，下面省略）

选路方式=IGP最短路径

算法的计算过程：

根据IP\_SPT数据，只要存在源宿节点间链路，即有IGP可达路径

算法的输出：

宿节点的NodeSID

实际应用：

SRBE（Segment Routing Best Effort的简称）就是直接走IGP最短路径。报文头部标签打入对端的NodeSID即可。

* 场景二：当源宿节点间的业务报文转发路径需**直接走IGP最短路径时**，源宿节点间**不存在直接**可达路由。

实际情况分析：导致域内路由不通的原因就2条，要么配置错误，要么网络故障。 但这两种原因都和算路条件无关，是需人力介入做问题排查的。所以要求走IGP最短路径但网元转发表中又没有路由的情况下，没有选路结果输出是合理的。

* 场景三：当源宿节点间的业务流需要走指定路径，指定路径的设定条件有：

必经节点，必经链路，必绕节点，必绕链路。

算法的输入条件：

选路方式=指定路径，输入必经节点或链路列表，必绕节点或链路列表

算法的计算过程：

必经路径要求是有序的。在SR网络中，必经节点可以对应于该节点的NodeSID，必经链路可以对应于该链路发端的AdjSID。所以必经路径的表示就是NodeSID和AdjSID的SIDList序列。



优化：L3 TE路径本来就是包含必经节点或链路的情况下，无需打NodeSID或AdjSID。



有必绕路径的节点和链路条件输入情况下，可考虑以下两种计算方法：

方法一、基于底层L3 拓扑数据算路，并将结果和IP SPT拓扑做掩映

需先复制L3 Topo数据，并将该拓扑数据中减去需必绕的节点和链路，再在删减后的拓扑上计算出源宿节点间的最短路径，将该路径和IP SPT拓扑上的路径做最长匹配，匹配的段用NodeSID表示，不能匹配的节点间直连链路用AdjSID表示。

比如下例的SIDList在完成上述匹配流程后，得到结果是：

AdJSID1，AdjSID2，…，NodeSID-Z，…



SIDList的深度取决于必绕节点必绕链路在L3拓扑中的依赖程度（相关链路的IGP Metric值很小，则IP流量都会受牵引），如果一个必绕的节点是L3拓扑中都要走的最短路径，那么被移除该节点后，实际需打的SIDList就比较长了，如下图。 实际上这种情况在现网中是不常见的，重要节点一般是有备用节点的，如果要必绕重要节点做流量牵引时，一般会走备用路径，也就是备用路径节点也是在IP SPT路径上的，这样就没有过多的AdjSID要打。



方法二、直接基于修剪过的IP SPT拓扑，做最短路径计算：

这里以必绕链路条件为例，做算路说明（必绕节点方法类似）：在L3拓扑中找出被删除链路，再找出被删除链路的上层链路upper\_layer\_link（即相关IP SPT链路），将这样的IP SPT拓扑链路排除后，剩下的拓扑就是可以走的IP可达拓扑，再对该拓扑做最短路径计算，即可得到符合条件的最短路径及其SIDList。 以下图做解释：

注：这里的最短路径包括了最短的路径权重之和，以及最短跳数。



源节点3，宿节点4，必绕链路：(3,4)

L3拓扑链路(3,4)的upper\_layer\_link是SPT拓扑上的链路(33,34)，链路(33,35)

复制IP\_SPT\_Topo，并将上述链路移除，得到新的IP SPT拓扑，并做最短路径计算。

得到最短路径是{(33,31),(31,34)}或{(33,32),(32,34)}，SIDList只要打{Node1,Node4}或{Node2,Node4}

下面例子是AdjSID在算路中的应用场景：

如图，要求是必绕(1,2)间的L3链路。

在IP\_SPT\_Topo上移除相关链路后会发现没有了从节点31出发到其他节点的路了。所以，对IP\_SPT\_Topo的模型数据还需对两点间不存在IP SPT链路，但存在L3链路的，将这样的L3链路数据补充到IP SPT算路拓扑数据上。当路径经过这样的链路时，就需要打AdJSID。比如：IP SPT上的路径(31,34)走 31-33-34，那么31上打的SIDList就是{Link-31-33-AdjSID, Node34-NodeSID}



总结，可以看到方法二，逻辑简单，是优选。

* 场景四：按非IGP链路Metric选的最短路径，比如按时延最短路径，按TE Metric最短

路径计算和SIDList生成的分析：

IP SPT拓扑链路的底层路径（under\_layer\_links）就是L3拓扑链路构成的。设置在L3拓扑链路上的时延Delay值，TE Metric值可以做累加计算到IP SPT拓扑的链路上。 这样直接基于IP SPT拓扑的链路做最短路径算法即可得到最低时延或最小TE Metric路径，然后按路径上的途径点打NodeSID或AdJSID，形成SIDList。

* 场景五：按流量负载均衡策略选路

控制器可以监控网络上的实时流量，记录每条L3拓扑链路的剩余可用带宽和带宽使用率，对新上的业务或对老业务调整，都可以按网络链路的带宽实际使用情况做负载均衡。 如按走最大剩余带宽链路的方式来做负载均衡为例，控制器将对IP SPT链路设置可用带宽=Min(available bandwidth of under\_layer\_links)，即选途径L3拓扑链路上的可用带宽值为最小那段链路的带宽作为IP SPT连接链路的可用带宽。然后按高带宽低权重的方式将可用带宽折算为链路权重，这样在IP SPT上做出的最短路径就是符合负载均衡要求的路径。

* 场景六：按链路亲和属性Affinity选路

亲和属性提供了一种链路的分类方式，不同业务可以限制走到具有不同亲和属性的链路上。在SDN控制器的支持下，即使网元不提供链路亲和属性功能，也能完成MPLS TE的亲和属性寻路规则。通过在控制器上对L3拓扑的链路设置亲和属性值，这些亲和值会同步到IP SPT链路上，即SPT链路的亲和值是途径L3链路的亲和值的交集。 这样，当用户要求按某亲和值选路时，控制器就先按条件过滤出SPT拓扑，然后再按最短路径算法计算路径。

**（2）技术效果**

给出本发明（实用新型）与现有技术相比存在的优点、所产生的效果，此处需要结合技术方案来描述，以推理方式说明，做到有理有据，切忌只下结论，而无分析。

传统的IGP路由技术只按最短路径的方式做流量引导，在网络业务量繁忙时，很容易导致局部路径过度拥塞，而又有其他空闲路径的存在，网络带宽资源利用不均衡等问题的存在。 MPLS TE流量工程提供了不同于默认最短路径的流量调整方式，但网元需支持相关协议RSVP-TE，且配置和维护相对复杂。Segment Routing技术与SDN控制器系统的结合为网络业务的多样性需求提供了灵活的支持手段，但到目前为止的SR部署要求来看，也基本上是对SRTE/SRTP隧道源宿节点间的每段连接都打AdjSID的方式做报文转发，这样导致报文净荷低效。

本专利中提供的技术方案既能实现对流量的灵活按需引导，又能提供最精简的SR报文转发标签头。SDN技术的本质是软硬分离，为网络业务发展提供更多可能性。网元不需要做MPLS TE那样的复杂功能，只要提供传统的IGP路由功能和SegmentRouting路由转发功能，复杂的业务路径计算由控制器来实现，网元简单化，控制器集中化，是网络智能化的一个发展方向。本专利通过分层的拓扑数据模型，L3拓扑反应的是IGP网络上节点邻居关系的拓扑，IP SPT拓扑反应的是网元IP可达链路，利用这两层拓扑数据来作为SR网络业务路径计算的拓扑数据源，结合通用的最短路径算法，可以方便地得到最小深度的SIDList，为智能地引导各种需求的业务流提供了有效方案。

三、上述技术方案是否有替代方案

如果有替代方案，请详尽写明，该内容的提供可以扩大专利的保护范围，防止他人绕过本技术去实现同样的发明（实用新型）目的；

可以是部分结构、器件、方法步骤的替代，也可以是完整的技术方案的替代，例如：两个部件的连接为卡式连接，但铰链连接也可能实现本发明（实用新型），因此铰链连接即为替代方案。

本发明中，重点在分层拓扑数据的构造，是从符合SegmentRouting分段网络的理念出发来考虑的，目前没有替代方案可寻。

四、本发明（实用新型）的关键技术点（即最想保护和强调的技术手段）

此处需提炼出技术方案的关键创新点，列出1、2、3……以提醒代理人注意，便于专利代理人撰写权利要求书。具体可以结合本发明（实用新型）带来有益效果的关键技术点；请标出创新点的重要性顺序，无需写成权利要求的形式。

1. 一种SR网络中按业务要求计算路径后生成SIDList的方法，其特征在于：**按SR分段网络的特性，构造L3和IP SPT拓扑数据模型，并基于此拓扑模型数据做路径计算和SIDList生成**。所述拓扑模型的定义以及各种场景下的算路方案描述，为SR网络的PCE计算功能实现和业务部署应用提供了是实际参考意义。
2. 一种构造IP SPT拓扑数据的方法，其特征在于：**控制器的IP SPT拓扑数据同步了网元的IGP SPT DB，同时又综合了L3邻居链路数据**。所述的分层拓扑数据模型，L3拓扑数据模型提供传统IP网络基本的链路Metric，并且在控制器的支持下，还可以扩充网元可能不支持一些链路特征值，比如TE metric，Delay time, available rate, Affinity, SRLG等；IP SPT拓扑数据表示了IP可达性，并综合了底层L3拓扑数据的特征值，结合节点的NodeSID以及链路的AdJSID。正是有这样充分必要的数据源支持，使SR隧道路径的SIDList生成十分方便。