

Segment Routing : 大规模SDN部署必备技术

2016年12月1日

序号	问题	回答
1	P13,如果中间有节点故障了,只有一条路的时候,咋办?	会选择备用路径,只有一条路径的时候,就会只走唯一的这条路径
2	每个Segment的可达性,怎么保证呢?	通过IGP/BGP学习得到,也可以由控制器静态配置
3	这个备用路径是SR开始就算好的,还是当故障出现后,SR才会再计算出一个路径	开始就计算好的,并且是故障后的最优路径
4	ECMP是等值负载均衡吗,SR只支持这个吗	ECMP(Equal Cost Multipath)即等价路径负载均衡.SRTE还支持按权重(Weight)的负载均衡
5	只有一条路,是否可以segment不通了,就切换到IGP?	一般而言IGP通,SR就通了
6	计算,谁负责?控制器?意思是SR不能离开控制器使用?	SR可以由路由器自己计算,和RSVP-TE类似,此时不需使用控制器.但SR结合控制器会更灵活,更可扩展
7	感觉SR和MPLS的概念有点类似	SR和MPLS主要区别有4点: 1)SR是源路由 2)SR可以运行于MPLS数据平面也可运行于Native IPv6数据平面 3)SR不依赖于额外的协议如LDP/RSVP实现标签分配 4)SR实现流量工程时中间节点无须维持路径信息

8	segment list 和tag 差不多	SR使用MPLS数据平面时,Segment List在形式上和MPLS标签一样(重用),但语义上有很大差别,Segment本质是代表指令,因此Segment List本质是指令集合:例如Prefix-SID标签代表把报文按IGP最短路径送达产生此Prefix-SID的节点,Adj-SID标签代表把报文从某个IGP邻居链路发送出去,L2-Adj-SID代表把报文从某个二层链路发送出去等
9	跨域咋办呢?	可以用BGP来分segment ,即SR BGP.当然也可以采用IGP/BGP双向带Prefix-SID重分发的方式
10	ASR9K可以支持几层Segment ?	支持10层Segment,并且不会有任何性能损失
11	控制器与router之间通过什么协议互通?	PCEP/BGP/Netconf均可
12	选完路径后,传输过程中QoS如何能保证?网络可以随时变化	SR在使用MPLS数据平面时, QoS和传统MPLS网络是一样的, 包头有EXP
13	使用segment之后,设备少了swap的环节,硬件处理要求下降?简单一点?	还是swap的, 只是很多时候 是swap到一个相同的label, 对硬件的要求其实更高, 因为在做复杂的流量工程的时候, 要处理的标签栈的深度会更深
14	SR技术是cisco独有吗?如果不是,其他竞争对手也使用这种技术吗?	SR是Cisco 3年前提出的,现已成为IETF标准, Cisco是SR的领导者,但SR并非Cisco独有。事实上业界主流厂商已经支持SR, 开源SDN控制器ODL从锂版本开始就支持SR, 同时很多第三方机构及用户已经进行过互通测试甚至是实际部署
15	大型OTT/运营商,现在有哪些案例呢?	北美Tier-1电信运营商,北美Tier-1有线运营商, 北美Tier-1 OTT运营商,中国Tier-1运营商,中国Tier-1 OTT运营商等
16	BGP怎么实现负载分担	采用BGP multi-path

17	在银行业有实际应用吗？	有， 高盛
18	MVPN，使用SR的话，有相应的解决方案？	SR主要解决LSP建立的问题，不影响上层的service
19	segment 和 SDN结合，是控制器维护SR还是数据平面支持	SR无论是否和控制器结合,设备的数据平面均需支持SR. SR在和控制器结合时,有两种做法:可以由控制器生成所有的SR标签,然后静态配置给各个设备,设备此时无需启动SR的控制平面;然而更普遍的做法是设备启用SR控制平面,然后控制器按需下发路径.
20	SR主要好处就是支持基于源地址来做路由转发吗	这是其中之一.SR的最大好处其实是实现了应用驱动网络
21	prefix两个属性分别是啥功能	全局可见,全局有效
22	sr是如何选择一条时延最小路径？怎么判断的？	可以用TE metric代表链路的时延，然后根据TE metric计算最优路径
23	对应用层的程序有什么要求吗？	TE path的建立和应用层没有关系
24	option，epe分别作用？	BGP EPE用于实现出向跨域流量的调度
25	对控制器有什么要求吗？	可以不要任何控制器，SR也能跑起来，要做跨域的TE，可以利用PCEP从控制器获得计算好的TE policy Cisco WAE控制器支持SR
26	SR不需要控制器，但SDN需要SR，是这样吗？	一定程度上可以这样说.但SR没有和控制器结合的话,无法发挥其最大作用.
27	这技术主要用于电信商层面还是企业？还是都可以？	都可以，主要还是运营商，但是很多企业也肯感兴趣，因为企业以前不喜欢MPLS就是觉得MPLS的控制平面太复杂，而SR简化了MPLS

28	没有其他路由器协议，直接静态配置SR，能转发？直接成为类似MPLS CORE的设备使用，能行？	可以的,支持静态SR
29	RSVP 到SR需要配置吗	请关注后续思科技术更新
30	都已经源路由了，还要IGP么？慢慢所有的路径都是控制器计算了，路由器都不需要路由功能了，应该叫做转发器了	完全用控制器来计算路由，在简单地拓扑没问题，比如DC，但是在复杂的网络里面，扩展性和HA都有问题，所以不能完全依赖控制器，尤其是故障检测到故障收敛这段时间容易造成路由黑洞
31	SR会转成转控分离方向？	可以，但是没有必要，SR做的TI-FRR（50ms 保护所需要的per prefix的LFA计算），用controller就做不了
32	SR是为了提升网络效率，利用SR实现SDN在大型网络中的部署。可以怎么理解么？	可以，Segment routing可以被看做MPLS之后routing领域的又一波革命性的变化，它本身是在infrastructure层面的创新,但其简化和可扩展的特点使其可以支持与应用的紧密结合,实现应用驱动网络
33	现在数据中心的协议是全BGP？	不一定，但是很多大型DC用BGP作为IGP，即在每台交换机间建立单跳的BGP会话来实现DC路由通告
34	请问一下如果将SR和SDN控制器结合的话，本身SR在中间节点是无状态的。如何解决故障场景下的快速收敛问题？	SR支持TI-LFA,即SR会先算好备份路径,并且此备份路径是故障收敛后的最优路径
35	部署SDN后，建议原路由协议还是继续保留使用？	IGP还是需要，MBGP传VPNv4也还是要的啊，只是LDP/RSVP不用了
36	现在支持全程带宽预留的是SR吗？	SR本身不预留任何带宽

37	另外SR对硬件支持有要求吗？	对于支持MPLS转发的设备，都支持SR转发，只要升级软件应该都支持SR，但是支持的label深度可能会带来一定的限制。对于SRv6,由于SRH由多个IPv6地址组成,处理这个可变长的SRH可能会需要某些硬件升级
38	我觉得如果结合全程带宽预留的话，适用场景会更多些	全程带宽预留是RSVP试图做的事情，但实际上全程预留带宽的做法很难扩展,部署很少.SR目前不支持带宽预留，但后续可能会结合控制器实现一定程度的带宽预留，请留意后续思科技术更新
39	比如我现在的N7K的硬件，是否需要升级什么引擎或则线卡。还是直接升级软件就可以在设备使用SR	请联系与您接口的思科销售团队,以核实您的软硬件版本,看具体如何升级
40	另外，SR下的FRR怎么实现的？	SR设备支持TI-LFA,基本原理是SR设备预先计算出来备份路径,而且此备份路径是故障收敛后的最优路径.即使P空间和Q空间无交集,SR也可以压入标签使得报文可以从P空间进入Q空间,也即可以支持任何拓扑 具体可以看 https://xrdocs.github.io/segment-routing/tutorials/
41	SR如何去探测网络中路径的带宽和网络延迟信息？	SR提供确定性路径检测能力,即SR使得探测报文可以通过网络的指定路径.因此通过构建不同的Segment List,探测代理将可以主动向网络的所有路径发出探测报文,然后将所有探测报文的的结果加以关联分析,即可得到全网的带宽/时延/丢包等信息
42	SR与SDN结合有好多仿真平台么？	Cisco dCloud(dcloud.cisco.com)上有可以马上使用的沙盒.也可以使用Cisco CML仿真建模工具加上ODL来搭建自己的试验平台
43	结合SDN后，做带宽预留就简单了。全程带宽预留可以做，可以做	controller去预留好之后，可以告诉SR的头端，怎么走，那没问题
44	kernel的标签深度有限制么？	请关注后续思科技术更新

45	采用SR之后，流量采集有什么手段？	流量采集和原来一样.流量采集未来发展方向是Telemetry,采用发布/订阅的方式而不是今天的轮询方式来采集,可以做到秒级的采集间隔和非常大的采集量,这都是为了支持应用级别的采集和后续的调度
46	Segment Routing: Foundation for Application Engineered Routing (Networking Technology) 这本书么？□	Kindle版本链接: https://www.amazon.com/Segment-Routing-Part-Clarence-Filsfils-ebook/dp/B01I58LSUO/ref=sr_1_1?s=books&ie=UTF8&qid=1480860167&sr=1-1&keywords=segment+routing
47	SR端到端后，怎么解决标签栈深度问题？	可以分段来分标签，用Binding SID接续起来
48	请问Segment routing和Network Slicing如何结合起来？谢谢！	可以为每个Slice分配一个Anycast-SID,然后由控制器压入不同的Anycast-SID实现Slicing
49	全局标签的路径一定会走负载分担吗？	是的，根据IGP SPF算法或者BGP选路规则，支持ECMP
50	这是属于哪个认证的部分？	CCIE Service Provider
51	AS跨域时，标签是谁来发布的？	BGP可以发，也可以是controller
52	如果是BGP发布的标签，意味着这个标签还需要再和Controller交互，告知Controller这个标签？	是的
53	用SR替代LDP最大的好处是什么？	避免了LDP和IGP同步黑洞问题,简化了整个网络的控制平面
54	是否可以disable SR的ECMP？	要Adj-SID指定一条路径就可以不走ECMP
55	SR 去掉RSVP 怎么保障QoS?	只是没有带宽预留,仍然可以采用通常的Diffserv QoS机制

56	ECMP这个是默认的？	是的
57	遇到node/link failure时, 是这个failure node/link的所有相邻结点都会自动rerouting吗? 对于出发端向SR压栈时, 响应failure的速度是快吗?	对的, 头端不用重新压栈, 只是临接failure的node重新压标签, 转到备用路径去
58	LDP向SR迁移难度和风险如何?	LDP向SR迁移完全是平滑的,支持make before break,而且可以逐节点迁移,有问题可以马上回退
59	这个感知还是利用bfd之类的技术吗?	故障检测机制和今天采用的机制基本一致,LF/RF/LOS/BFD等.特别地,为了提高效率,SR建议采用SBFD