H3C

智广域

向未来

路由器SD-WAN技术专刊

广域网中低端路由器全家福

网络位置 – 总部

SR6608

SR6602-IE

MSR5680

MSR3640-X1-HI

★★ 推荐型号 ★★

大中型总部	中小型总部	公有云、私有云
SR6604	MSR5620	VSR1000
SR6608	MSR3640-X1-HI	
SR6616		
SR6602-I		
SR6602-IE		
MSR 56-60		
MSR 56-80		

网络位置 – 分支

MSR3640-X1

MSR3620-X1

MSR3610E-X1

MSR3600-28[51]-X1

MSR2630E-X1

MSR2600-15-X1

MSR1000系列

MSR810系列

★★ 推荐型号 ★★

大中型分支		中小型分支		微型分支	
MSR3640-X1	MSR3610-X1	MSR2630E-X1	MSR2600-15-X1	MSR810	MSR810-CNDE-SJK
MSR3620-X1	MSR3620-G	MSR3600-51-X1	MSR2600-6-X1	MSR810-W	MSR810-LM-CNDE-SJK
MSR3620-DP	MSR3610-G	MSR3600-28-X1	MSR1008	MSR810-W-LM	MSR830-10HI-GL
MSR3610E-X1		MSR3600-51-G-DP	MSR1004S-5G	MSR810-LM	
		MSR3600-28-G-DP	VSR1000		

目录

第 1 章 AD-WAN 分支解决方案 · · · · · 1-1

1.1 当前广域网面临的挑战 · · · · · 1-1

1.2 AD-WAN 分支解决方案介绍 · · · · · 1-2

1.3 客户价值 · · · · · 1-4

极简开局，随需而动 · · · · · 1-4

智能选路，灵活调配 · · · · · 1-6

极致优化，安全加固 · · · · · 1-7

智能分析，可视运维 · · · · · 1-10

1.4 关键技术 · · · · · 1-12

零配置开局 · · · · · 1-12

灵活拓扑技术 · · · · · 1-18

NAT 会话穿越 · · · · · 1-25

智能选路 · · · · · 1-30

DPI 应用识别 · · · · · 1-34

SaaS 路径优化 · · · · · 1-39

TCP 拥塞控制 · · · · · 1-46

自适应音视频保障 · · · · · 1-53

多路径包复制 · · · · · 1-56

Web 应用缓存 · · · · · 1-60

智能网络质量分析 · · · · · 1-61

音视频质量分析 · · · · · 1-66

1.5 典型组网 · · · · · 1-69

总部+分支（Hub-Spoke 组网） · · · · · 1-69

分支入云（Hub-Spoke 组网） · · · · · 1-70
总部+分支（多总部组网） · · · · · 1-70
总部+分支（Full-mesh 组网） · · · · · 1-71
总部+汇聚+分支（三级分层组网） · · · · · 1-72
多租户 POP 组网（MSP 运营组网） · · · · · 1-73
1.6 成功实践 · · · · · 1-74
大型集团 · · · · · 1-74
保险行业 · · · · · 1-76
银行系统 · · · · · 1-78
零售行业 · · · · · 1-79
MSP (Managed Service Provider) · · · · · 1-81
第 2 章 云简广域网解决方案 · · · · · 2-1
2.1 传统 VPN 方案存在的问题 · · · · · 2-1
2.2 云简广域网解决方案介绍 · · · · · 2-2
2.3 客户价值 · · · · · 2-3
多——线路多样，特性丰富 · · · · · 2-3
快——集中部署，快速上线 · · · · · 2-4
好——分权分域，整网可视 · · · · · 2-6
省——云管服务，智能运维 · · · · · 2-8
2.4 典型组网 · · · · · 2-9
单数据中心分支单线路组网 · · · · · 2-9
双数据中心分支双线路组网 · · · · · 2-10
2.5 成功实践 · · · · · 2-12
大型饮品公司 · · · · · 2-12

第 1 章
AD-WAN 分支解决方案

摘要

本章主要介绍当前广域网面临的挑战、AD-WAN 分支解决方案概述、客户价值、关键技术、典型组网和 AD-WAN 分支解决方案在百行百业的成功实践。

1.1 当前广域网面临的挑战

当今我们所处的时代，被赋予了众多信息技术变迁的标签，互联网时代、自媒体时代、大数据时代、云计算时代、元宇宙元年等，这一切的背后归根结底都是数据。以大量的数据作为依托，运用不同的技术手段，在数据的采集、处理、运用过程中不断地创新与深度发掘，成就了一个又一个的时代标签，而广域网正是这些数据在不同地区、省市、国家分支机构之间传输的基础。

随着云计算技术的不断深入人心，企业业务不断向公有云、私有云搬迁，业务数据对作为传输管道的广域网提出了如下图 1-1 所示的新的挑战。

图1-1 当前广域网面临的挑战

网络弹性差

云计算的快速发展，企业对于公有云、私有云、SaaS 云、边缘云多云之间的互联需求日趋增加，对于业务上线、变更耗时更加敏感。然而传统广域网封闭的管理控制方式，无法灵活适应业务的快速变化和增长。并且由于传统的广域网存在分布广的特点，这一问题将尤为突出。

业务爆炸式增长

伴随着业务的爆炸式增长，传统广域网面临如下两个问题：

- 不同的业务对网络质量的要求不尽相同。然而传统广域网分支的网络模型是基于最短路径进行转发，没有兼顾实际的业务需求，因此无法达到网络资源最大化的合理使用。
- 业务需求的快速迭代，使得需求永远走在带宽的前面，因此需要一种优化机制，能够在不断增长的带宽需求与网络建设的进度之间起到缓和的作用，缓解企业的 IT 建设压力。

安全形式严峻

万千业务应用上云的同时，海量级数据搬迁到云上，然而传统广域网分支存在数据传输过程中安全保障能力弱，无法提供基于全网维度的安全审计管理能力等安全隐患。

运维日趋复杂

传统广域网网络管理手段有限，以手工为主，对 IT 维护人员的技能要求较高。同时流量和业务无可视化呈现，造成故障无法快速识别和定位，运维难度大。

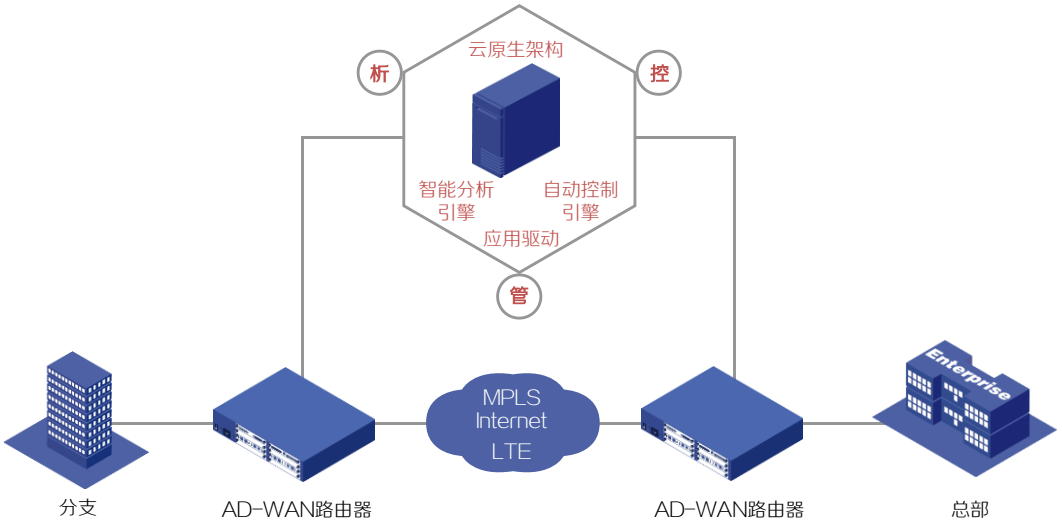
1.2 AD-WAN分支解决方案介绍

为了解决广域网面临的新时代挑战，H3C 提出了 AD-WAN（Application-driven Wide Area Network，应用驱动广域网）分支解决方案。

作为对应行业通用叫法 SD-WAN（Software Defined Wide Area Network，软件定义广域网）的产品形态呈现，AD-WAN 分支解决方案遵循业内 SD-WAN 标准，采用 SDN 网络可编程、逻辑集中控制的核心思想，统一融合智能网络管理、智能控制、智能分析，实现“管”、“控”、“析”三维一体的融合网络控制中枢和智能大脑，网络全域覆盖，实现端到端的网络和业务自动化、可视化、精细化的网络管理，如下图 1-2 所示。

AD-WAN 分支解决方案不仅兼具 SD-WAN 的种种优势，同时结合广域网分布广、接入形态多样、网络质量变化频繁等特点，采用设备自主智能选路，全网链路拓扑、质量集中呈现等多种形式，实现了 SDN 在广域网的最佳实践。

图1-2 AD-WAN 分支解决方案



AD-WAN 分支解决方案通过如下表 1-1 所示的技术，解决了当前广域网面临的网络弹性差、业务爆炸式增长、安全形式严峻、运维日趋复杂的问题。

表1-1 AD-WAN 分支解决方案技术

当前广域网面临的挑战	AD-WAN 分支解决方案技术
网络弹性差	**极简开局**：网络管理员只需在远程服务器上进行基础配置，无需在开局现场配置设备即可完成网络开局，大大降低了开局成本。 **网络随需而动**：一方面企业用户可以灵活选择Internet、4G/5G等线路无差别接入，不再依赖于传统昂贵的运营商专线，同时支持接入多种云服务商及混合云，为客户最大限度降低成本，保护客户投资；另一方面提供超强的业务弹性扩容能力，动态调整网络规模，降低了用户对网络变化的敏感度。
业务爆炸式增长	**应用智能选路**：基于链路类型、链路带宽、链路质量和时间段等丰富的选路和调度策略进行精细化的选路，保障应用的网络需求和用户的最佳网络体验。 **全面WAN加速**：通过TFO、BBRv2、DRE、LZ、FEC、包复制、Web Cache等多种加速手段，提高应用流量的传输速度和传输质量。
安全形式严峻	**本地安全**：支持IPS、状态防火墙、URL过滤、防病毒等多种安全业务部署，对外提供丰富的安全保障能力，实现安全保障。 **云防御**：支持SASE云防御部署，基于云服务提供的全方位安全服务，可以实现随时随地按需部署，降低分支对安全硬件的依赖，减轻分支网络维护复杂性，为用户提供网络和云安全融合的一站式解决方案。
运维日趋复杂	**智能网络分析**：基于Netstream技术从全网、站点等不同角度对应用流量实时采集，快速感知网络变化，同时采用大数据分析技术和AI算法，实现全网状态智能分析。 **可视化运维**：提供基于全网的多维度可视化能力，实现故障先知先决，极大提升快速定位的效率，变被动运维为主动运维。

1.3 客户价值

极简开局，随需而动

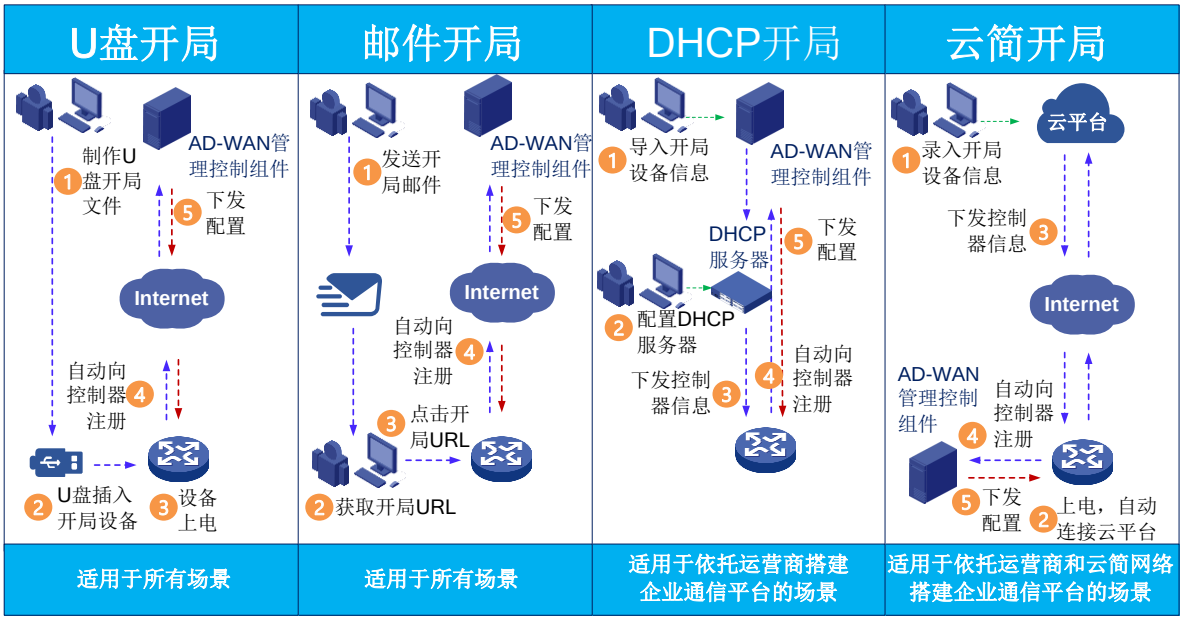
极简开局

传统开局需要工作人员到客户现场对设备进行完整的配置，这种开局方式有着相对较高的时间成本和金钱成本，同时对配置人员有一定门槛要求。

AD-WAN 分支解决方案推出的零配置（ZTP，Zero Touch Provisioning）开局是自动化网络部署方案中的重要技术。网络管理员只需在远程服务器上进行基础配置，无需在开局现场配置设备，即可将新购置的设备接入网络，完成开局。与传统开局相比，零配置开局不需要工作人员到客户现场对设备进行配置，只需要客户端开局人员通过简单的操作即可完成。零配置开局配置门槛低，有着相对较低的时间成本和金钱成本。

AD-WAN 分支解决方案的零配置开局支持如图 1-3 所示的四种开局方式。

图1-3 四种零配置开局的方式



网络随需而动

灵活入云

根据 Gartner 咨询公司对 IT 基础设施领域云计算的未来发展趋势的预测，到 2025 年，随着企业业务的增长和云计算技术的不断成熟，将有 85%的企业和组织采用云优先原则。如图 1-4 所示，H3C 为助力企业入云提供的 AD-WAN 分支解决方案，具备如下能力：

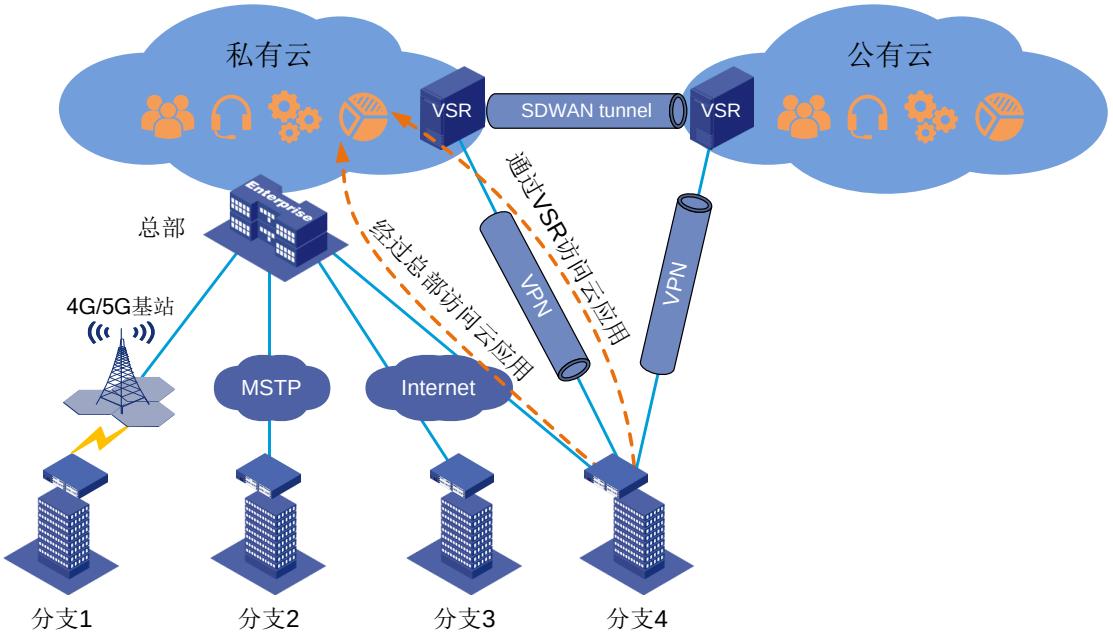
- 支持多种接入方式

提供 Internet、4G/5G 移动通信等接入方式，同时兼容 MSTP 等传统的运营商物理专线，支持企业入云的平滑演进。

- 支持多云互联
  - 通过在公有云服务器上部署 H3C 虚拟网关 VSR，可以为公有云和企业总部/分支之间提供安全访问的 VPN，企业分支既可以经过总部访问云应用，也可以通过 VSR 访问云应用。
  - 企业支持接入多种云（公有云、私有云、混合云），可在不同的云上部署不同的业务应用。
  - 支持 SaaS（Software as a Service，软件即服务）应用的云优化选路。
- 支持多种组网模型入云

  支持 Hub-Spoke、Full-Mesh、二级扁平组网、三级分层组网等组网模型，使企业可以按自身需求选择接入云端方式。

图1-4 灵活入云

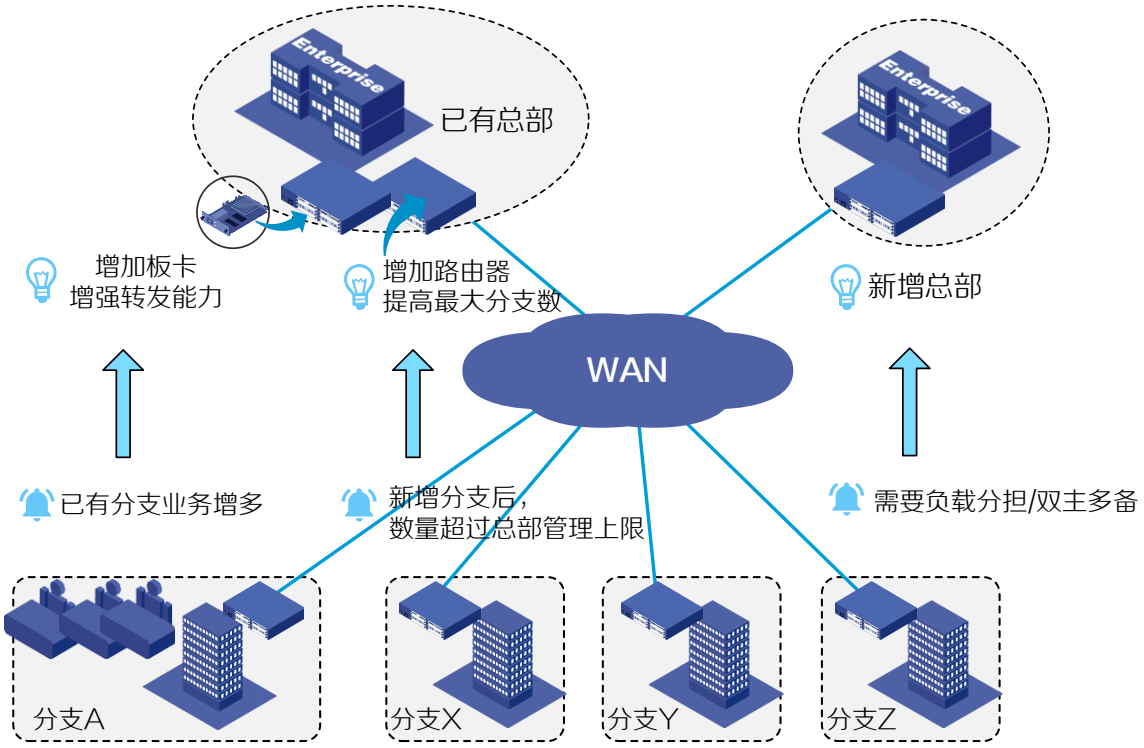


弹性扩容

AD-WAN 分支解决方案可以伴随企业发展，在不影响现有业务策略与网络架构的前提下，不断扩大网络规模。如图 1-5 所示，AD-WAN 分支解决方案具备如下网络扩容能力：

- 当整网业务到达一定规模，总部的处理能力遇到瓶颈时，总部路由器可通过增加可插拔板卡，增加业务转发能力以及支持的分支数量。
- 随着业务增长和分支扩张，当分支数量超出总部可管理的最大数量时，可在该总部内新增路由器来提高支持的最大分支数。
- 基于业务增长、新的业务规划、可靠性保障等新需求，可以新增总部，不但能扩大整网规模，还能实现和原总部之间的负载分担和业务备份。

图1-5 弹性扩容

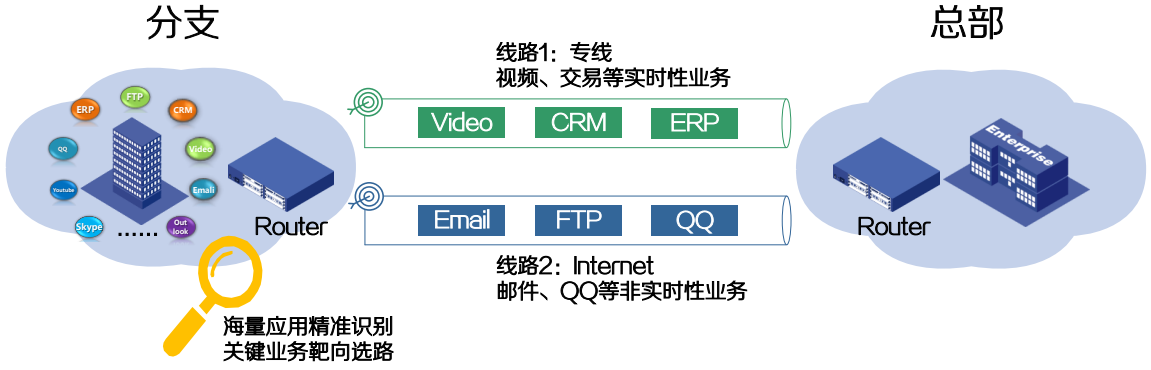


智能选路，灵活调配

企业分支和总部间的应用流量传输常采用基于链路开销或路由策略的选路方式。然而，由于选路机制的单一，流量分布的不均衡等问题，会大大降低应用体验。为此，AD-WAN 分支解决方案提供了可以根据应用类型的智能选路方案。

智能选路利用 iNQA 智能网络分析技术和 DPI 应用识别技术，不仅可以实时探测链路质量，基于链路优先级、链路质量和链路带宽定义链路优劣，还能自动识别应用类型，基于应用类型优选流量传输路径。在企业分支和总部网关设备上配置智能选路，可以避免应用流量分布不均，提高应用流量传输效率，如下图 1-6 所示。

图1-6 应用智能选路



极致优化，安全加固

全面 WAN 加速

企业分支和总部间的应用流量在广域网链路上传输时，常常因为链路时延大、丢包率高和带宽不足等问题，导致业务流量传输速度慢，音视频卡顿等现象。AD-WAN 分支解决方案提供的全面 WAN 加速方案，可以根据使用场景有针对性地优化应用流量，加速应用流量传输，全面提升应用体验。

- 应用流量高速传输的场景中，全面 WAN 加速利用 TFO 和 BBRv2 等传输层流优化技术优化拥塞窗口、DRE 消除冗余技术或 LZ 无损压缩技术压缩应用流量，可以加大网络吞吐量，提高应用流量传输速度，如下图 1-7 、图 1-8 所示。

图1-7 TFO/BBRv2 加速

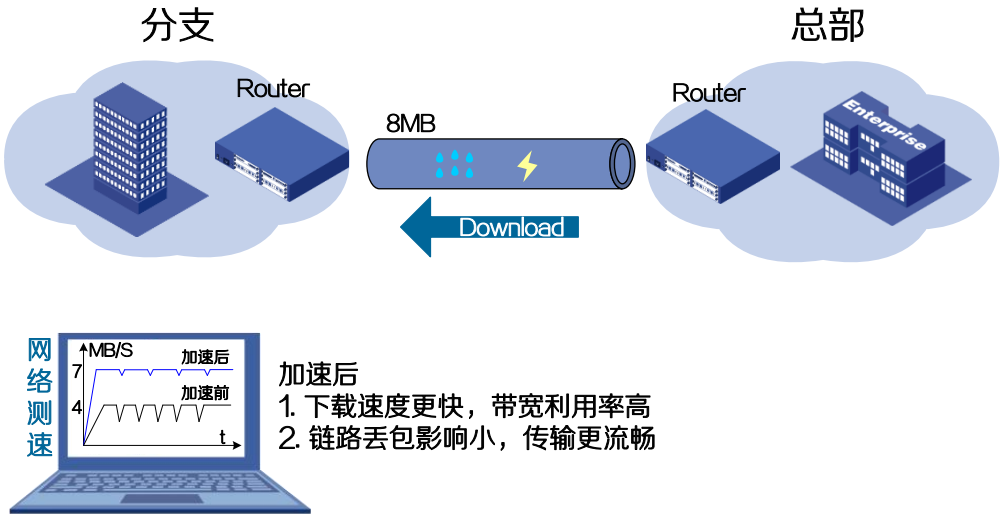
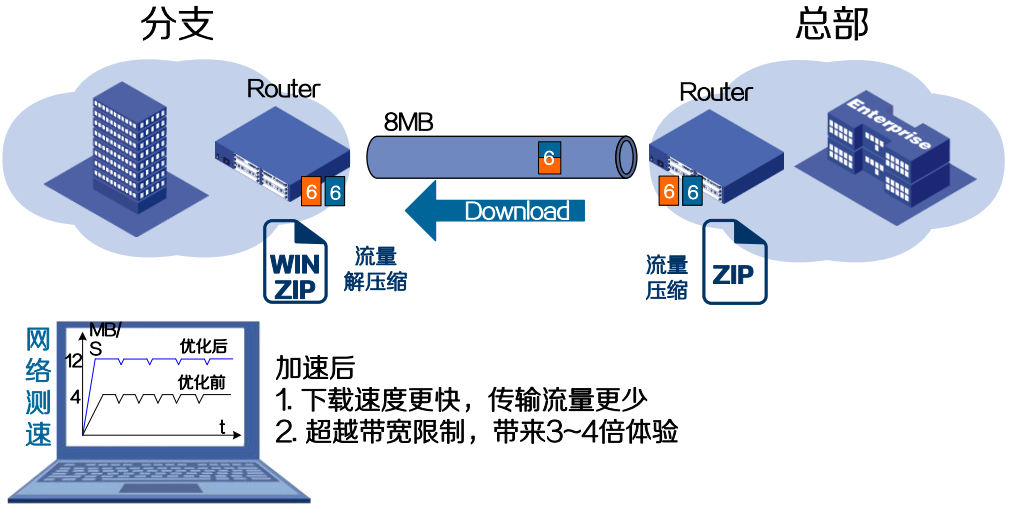
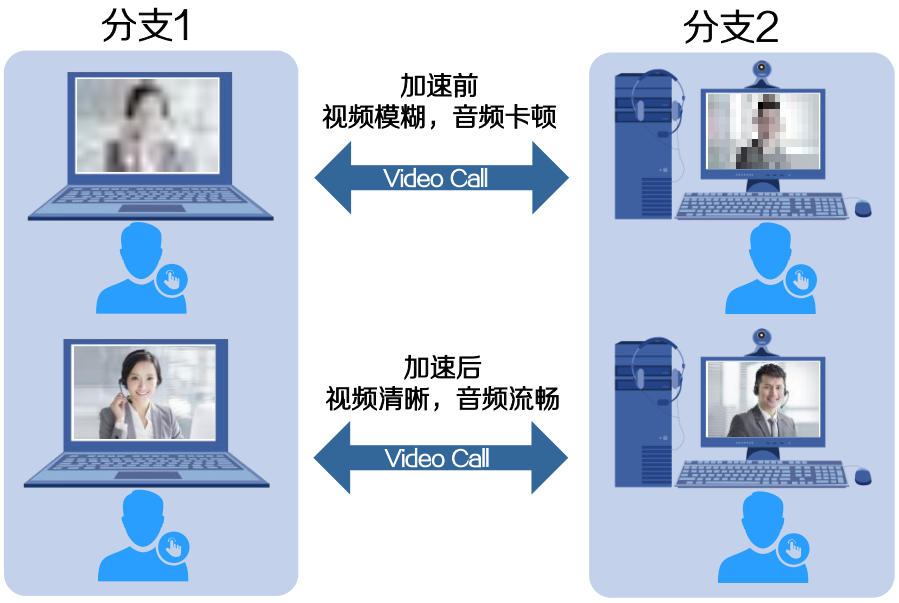


图1-8 DRE/LZ 加速



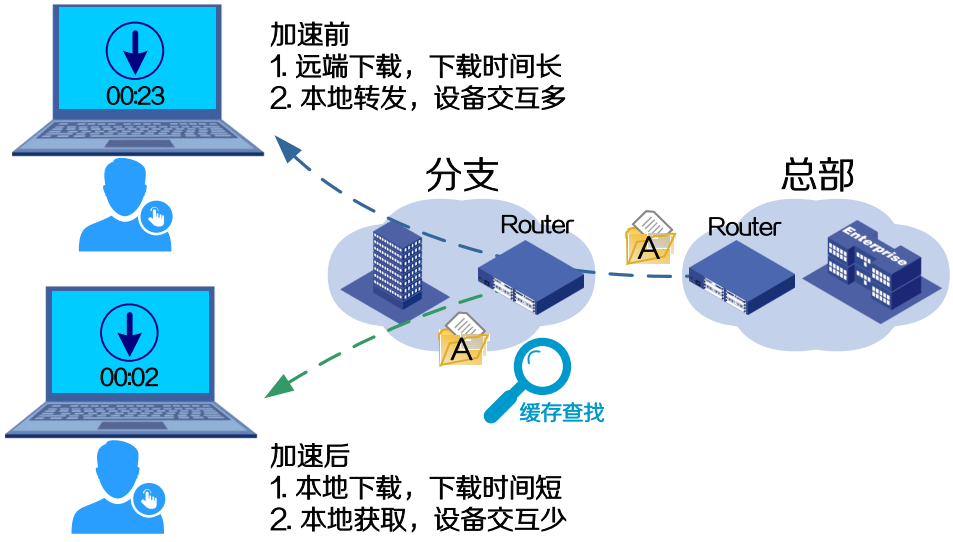
- 应用流量高质量传输的场景中，全面 WAN 加速利用 FEC 前向冗余纠错技术或包复制技术将数据复制传输、去重排序，可以提升应用流量的抗丢包能力和传输质量，如图 1-9 所示。

图1-9 FEC/包复制加速



- 应用数据高频访问的场景中，全面 WAN 加速利用 Web Cache 网站缓存技术在本地缓存应用数据，再次访问时从本地获取应用数据，可以减少设备与服务器间的交互压力，提高应用的访问速度，如图 1-10 所示。

图1-10 Web Cache 加速



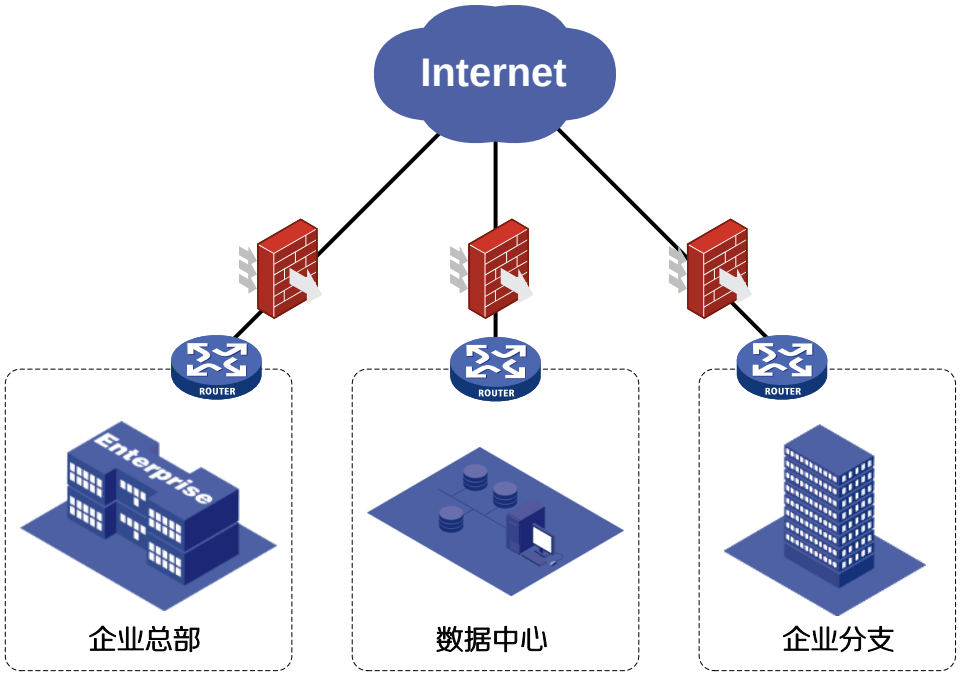
安全

AD-WAN 分支解决方案支持本地安全和云防御两大场景。在同一企业网络中，支持管理员根据企业站点需求不同，按需选择安全方案，实现本地安全和云防御混合部署。

本地安全

如图 1-11 所示，本地安全是指将网络安全边界限定在企业站点或数据中心边缘的硬件设备（比如防火墙等）中，通过在设备上部署 IPS、状态防火墙、URL 过滤、防病毒等多种安全业务，阻断外部网络攻击、防止内部数据泄漏、规范用户上网行为，从而提高网络的安全性。

图1-11 本地安全

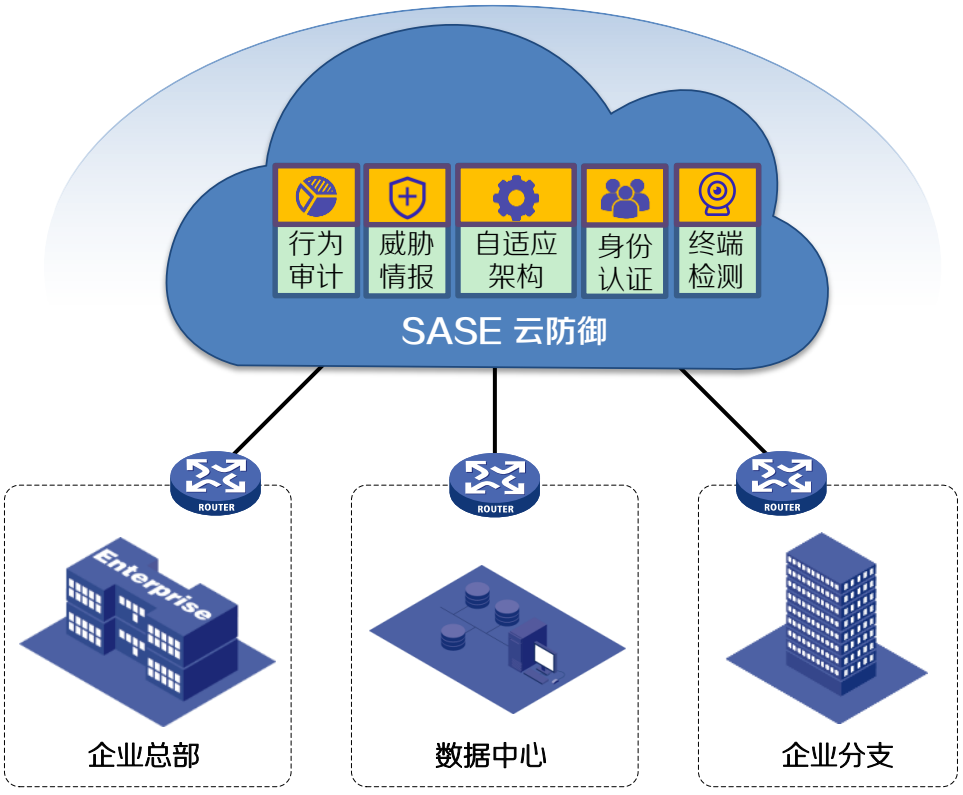


云防御

如图 1-12 所示，SASE（Secure Access Service Edge，安全访问服务边缘）是一种新型的服务架构，通过将广域网和网络安全结合起来，以虚拟云的方式统一交付，从而满足数字化企业动态安全访问的需求。SASE 云防御主要有以下优点：

- 提供更安全的业务：安全业务全方面覆盖，安全策略按需订阅、统一管控。
- 降低运营成本和运维难度：基于云服务交付安全能力，降低分支对安全硬件的依赖，减轻分支网络维护复杂性。
- 简化管理，一站式连接：提供集中式云交付管理控制面板，简化用户的操作；随时随地按需部署，为用户提供网络和云安全融合的一站式解决方案。

图1-12 云防御



智能分析，可视运维

智能网络分析

随着互联网进入大数据时代，网络业务急剧膨胀，网络规模、数据流量都呈爆炸式增长。传统的运维手段在问题发现、定位和解决的过程中逐渐变得力不从心。

AD-WAN 分支解决方案支持故障及业务仿真，音视频质量分析，并通过对设备性能、用户接入、业务流量的实时数据采集和状态感知，采用大数据分析技术和 AI 算法，将网络的运行可视化，主动感知网络的潜在风险并自动预警。例如图 1-13 所示，针对故障采用如下机制进行智能发现、智能定位：

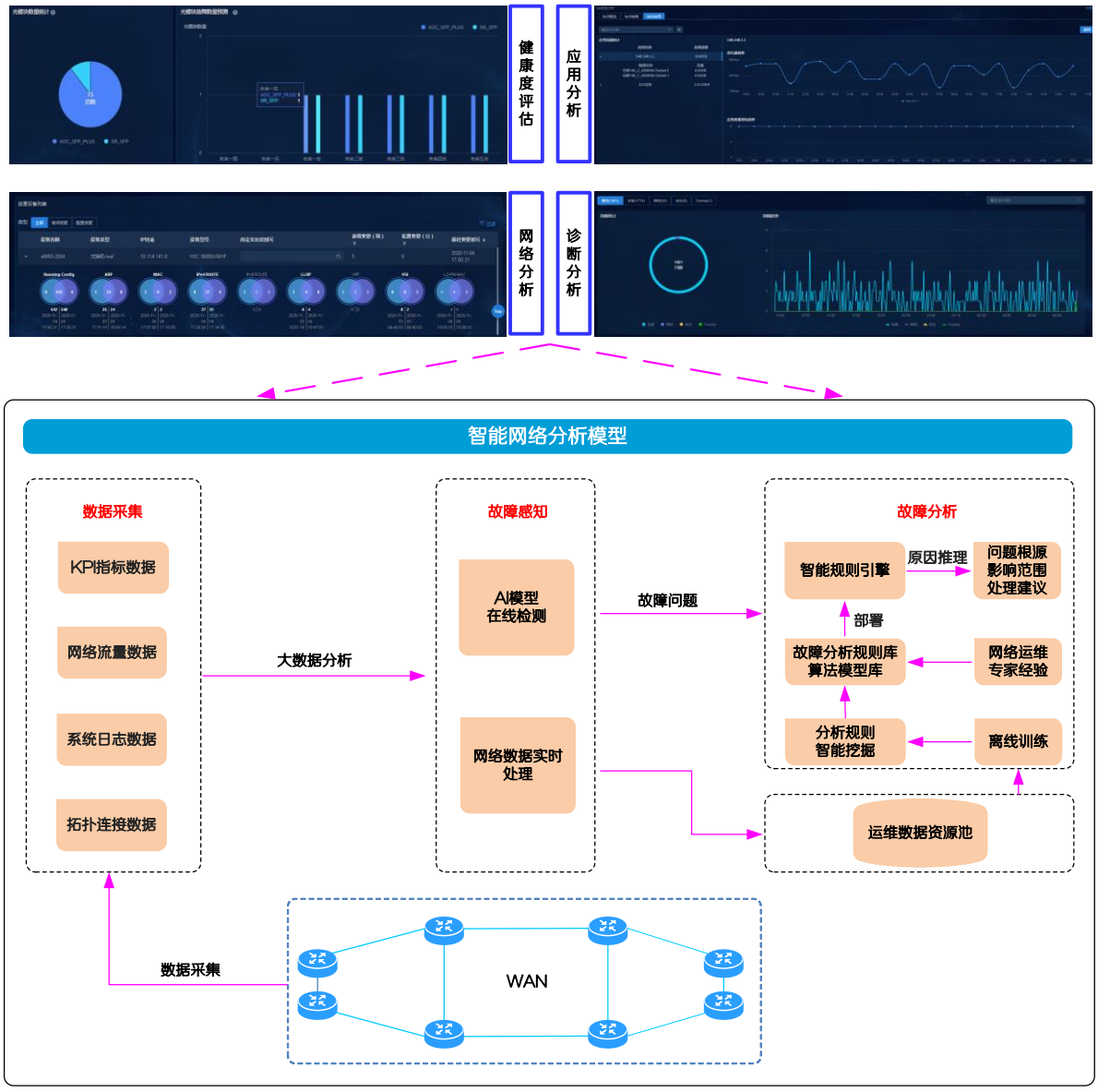
(1) 数据采集：采集网络和流量数据指标，采集粒度支持秒级和分钟级。

(2) 故障感知：基于 AI 模型检测 KPI 指标异常，实现分钟级感知网络故障。

(3) 故障分析：基于 AI 算法和规格模型分析推理故障原因，实现分钟级别的故障定位。

(4) 故障处置：根据故障分析原因和故障 Case，通知按照预案下发处置动作。

图1-13 智能网络分析



智能网络分析主要包括如下几个维度：

- 健康度评估：全网健康可视、主动运维。
- 网络分析：链路带宽/质量预测、网络变更分析，辅助定位。
- 应用分析：Netstream 流分析、SRv6 隧道流分析、IFIT 随流分析，问题定界等。
- 诊断分析：网络/设备/协议/Overlay 四大类常见故障、分钟级发现问题、定位问题根源并给出处理建议。

可视化运维

AD-WAN 分支解决方案能结合相关信息采集技术、流采集技术以及网络质量采集技术，实现基于全局的多维度可视呈现能力，以提供全网多维可视化服务，降低运维难度，减少运维成本，如图 1-14 所示。

图1-14 可视化运维

1.4 关键技术

零配置开局

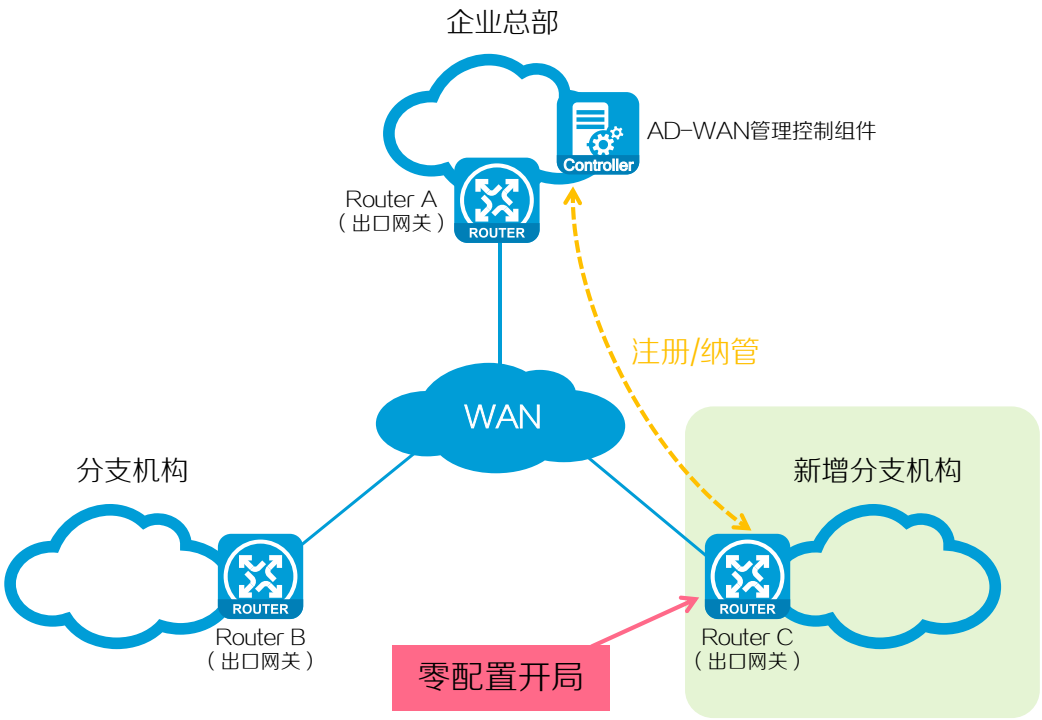
简介

如下图 1-15 所示，零配置（ZTP，Zero Touch Provisioning）开局是 H3C 推出的自动化网络部署方案中的重要技术。使用零配置开局技术，网络管理员只需在远程服务器上进行基础配置，无需在开局现场配置设备，即可将新购置的设备接入网络，完成开局。当局点数量较多、地域分布广泛、开局部署复杂时，使用零配置开局可大大简化开局过程，极大降低开局成本。

不同应用场景，零配置开局的具体要求和配置不同。本文介绍 AD-WAN 分支解决方案的零配置开局技术，它适用于对新增分支机构出口网关进行开局。对于 AD-WAN 分支解决方案的零配置开局技术，只要新增网关可以接入网络，并被 AD-WAN 管理控制组件纳管，则表示开局完成。开

局成功后，管理员可通过 AD-WAN 管理控制组件对网关下发业务配置、进行可视化智能运维，从而降低运维难度和成本，提高运维效率。

图1-15 零配置开局应用场景



技术价值

零配置开局技术价值如图 1-16 所示。

图1-16 零配置开局技术价值

省钱	省时	省力	零配置开局
• 传统开局：差旅加人工，开销达到千元级。 • 零配置开局：百元级。	• 传统开局：开通周期长，业务上线慢（数天到数周）。 • 零配置开局：分支节点分钟级开局。	• 传统开局：需要专业IT人员到现场配置设备。 • 零配置开局：不需要专业IT人员到现场配置设备。	• 网络&业务自动化下发，现场操作“0”门槛。 • 四种开局方式，适配各种应用场景。

四种开局方式

AD-WAN 分支解决方案的零配置开局，其开局流程大致为：

(1) 管理员在 AD-WAN 管理控制组件上导入新网关的信息(如设备名称、序列号和 Router ID 等)，以便后续管理该网关。

(2) 管理员在远程服务器上为该网关准备好基础配置，如 WAN 接口的 IP 地址、拨号上网的账号/密码、AD-WAN 管理控制组件的域名或“IP 地址+端口号”等。

(3) 开局人员 (现场负责安装/检修设备的网络工程师) 只需进行连线、上电等简单操作，无需对网关进行配置。

(4) 网关上电启动时自动加载基础配置，接入网络并自动注册到 AD-WAN 管理控制组件，完成开局。

为适应各种组网环境，零配置开局支持下表 1-2 所示四种开局方式，请根据现网条件选择合适的开局方式。

表1-2 四种开局方式

开局方式	功能概述	使用场景
U盘开局	通过开局U盘下发基础配置，完成开局的方式。	U盘开局适用于所有场景的零配置开局，管理员可通过AD-WAN管理控制组件轻松定制基础配置。 它要求开局人员携带开局U盘至现场，将U盘插入网关。
邮件开局	通过开局邮件下发基础配置，完成开局的方式。	和U盘开局相比，邮件开局要求开局人员准备开局终端 (智能手机、平板电脑、笔记本电脑和PC等) 来接收开局邮件，并将开局终端连接到网关。
DHCP开局	通过DHCP服务器下发基础配置，完成开局的方式。	DHCP开局适用于中小企业依托网络运营商搭建企业通信平台，新网关的基础配置由运营商下发。 它要求管理员具有DHCP服务器的配置权限，可通过DHCP的Option 253下发AD-WAN管理控制组件的地址。
云简开局	通过网络运营商和H3C部署在公有云中的云简管理平台下发基础配置，完成开局的方式。	云简开局适用于中小企业依托网络运营商和H3C云简网络搭建企业通信平台。其中，新网关的上网参数由运营商自动下发，AD-WAN管理控制组件的地址由管理员在云简网络配置后，通过云简网络自动下发。 与DHCP开局相比，它无需修改运营商网络中DHCP服务器的配置，也不用将AD-WAN管理控制组件的地址暴露给第三方。

零配置开局流程

- U 盘开局流程

  U 盘开局流程如下图 1-17 所示：

(1) 管理员在 AD-WAN 管理控制组件上导入新网关的信息，完成该网关的开局配置。管理员利用 AD-WAN 管理控制组件的 U 盘开局功能，将开局配置导出成文件后，保存至开局 U 盘的根目录下。

(2) 开局人员携带开局 U 盘至开局现场，将 U 盘插入网关，给网关连线、上电。

(3) 网关启动时，自动读取 U 盘中携带的开局配置完成初始配置后，自动与 AD-WAN 管理控制组件连接、注册，完成开局。

图1-17 U 盘开局流程

- 邮件开局流程

  邮件开局的流程如下图 1-18 所示：

(1) 制作开局邮件

a.管理员在 AD-WAN 管理控制组件上导入新网关的信息，完成该网关的开局配置。

b.管理员利用 AD-WAN 管理控制组件的邮件开局功能，将开局配置封装成一个特殊格式的 URL（为确保配置安全，封装时可选择对 URL 加密），然后将 URL 作为邮件内容，通过邮件发送给开局人员。

(2) 将开局配置发送给新网关

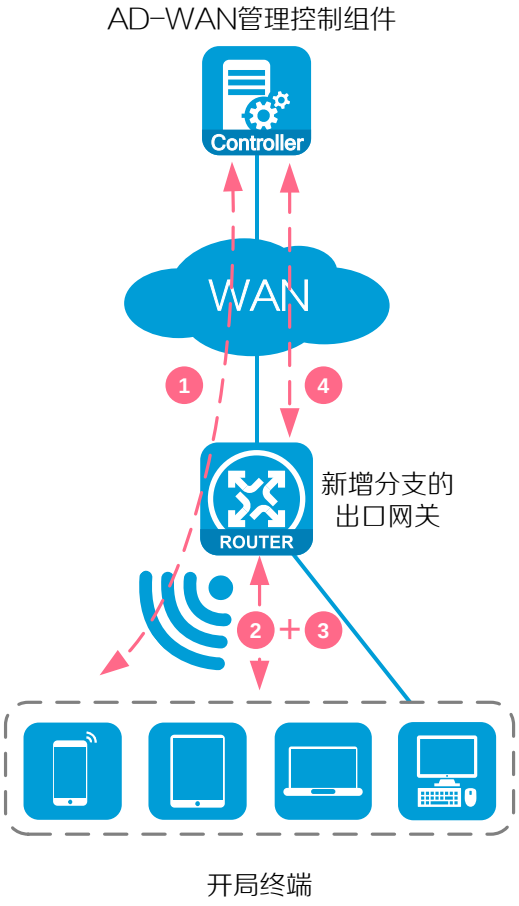
a.开局人员使用开局终端（智能手机、平板电脑、笔记本电脑和 PC 等）成功接收开局邮件后，将开局终端携带至开局现场。

b.开局人员给网关连线、上电，并使用网线将开局终端和网关直连（对于支持无线接入功能的网关，还可以通过网关出厂自带的开局 Wi-Fi 连接网关），然后点击开局邮件中的 URL 链接，将 URL 中携带的开局配置发送给网关。

(3) 网关收到 URL 请求后（对于加密 URL 此处需输入解密密码），自动跳转到开局配置 Web 页面。开局人员在该 Web 页面查看开局配置，并确认执行开局配置。

(4) 网关根据格式约定解析 URL，使用 URL 中携带的开局配置完成初始配置后，自动与 AD-WAN 管理控制组件连接、注册，完成开局。

图1-18 邮件开局流程



- DHCP 开局流程

  DHCP 开局的流程如下图 1-19 所示：

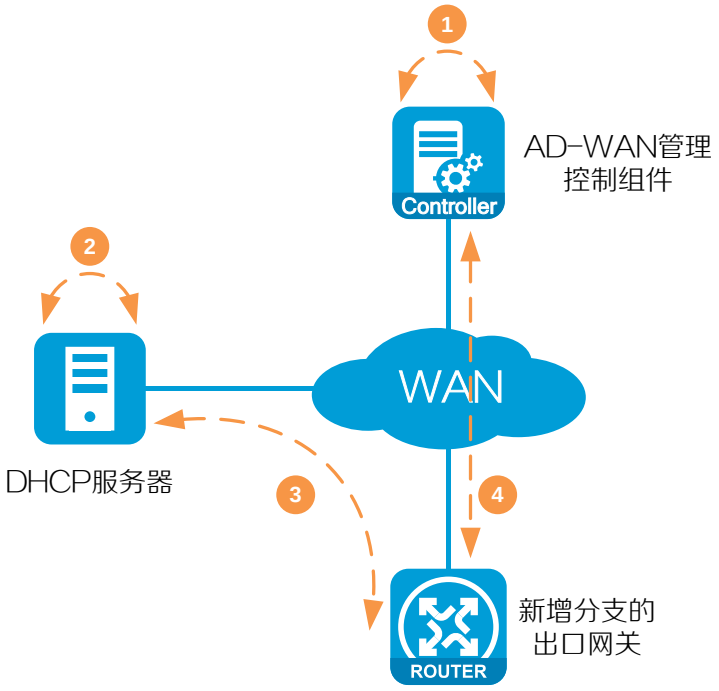
(1) 管理员在 AD-WAN 管理控制组件上导入新网关的信息。

(2) 管理员配置 DHCP 服务器，为新网关分配 WAN 接口地址等上网参数以及 AD-WAN 管理控制组件的域名或 IP 地址/端口号等参数。

(3) 开局人员给网关连线、上电。网关启动时，主动向 DHCP 服务器发起请求，以获取接口 IP 地址。DHCP 服务器在给新网关分配 IP 地址的同时，通过在 DHCP 报文中携带 Option 253 选项，将 AD-WAN 管理控制组件的地址发送给网关。

(4) 网关根据 DHCP 服务器推送的开局配置完成初始配置后，自动与 AD-WAN 管理控制组件连接、注册，完成开局。

图1-19 DHCP 开局流程



- 云简开局流程

  云简开局的流程如下图 1-20 所示：

(1) 管理员在 AD-WAN 管理控制组件和 H3C 云简管理平台分别导入新网关的信息，并在云简网络管理平台为该网关配置 AD-WAN 管理控制组件的域名或 IP 地址/端口号等参数。

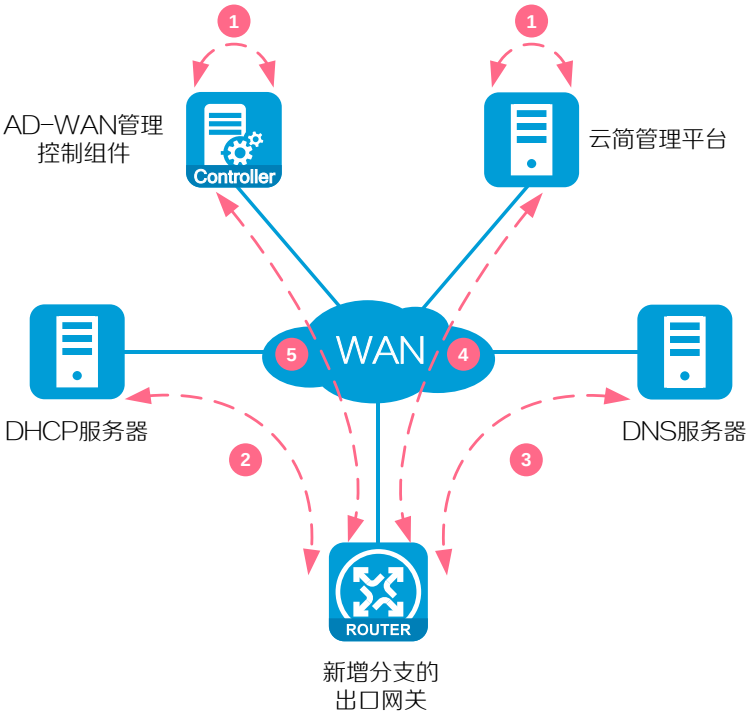
(2) DHCP 服务器为新网关分配 WAN 接口地址、缺省网关和 DNS 服务器地址。开局人员给网关连线、上电。网关启动时，自动通过 DHCP 功能获取 WAN 接口地址、缺省网关和 DNS 服务器地址参数。

(3) 网关读取出厂配置中携带的云简网络管理平台的域名和端口号，并通过 DNS 服务器将云简网络管理平台的域名解析成 IP 地址。（通常情况下，运营商网络提供的 DHCP 和 DNS 服务可满足云简开局需求，无需管理员额外部署）

(4) 网关自动连接到云简网络管理平台。云简网络管理平台将开局配置推送至网关。

(5) 网关根据开局配置完成初始配置后，自动与 AD-WAN 管理控制组件连接、注册，完成开局。

图1-20 云简开局流程

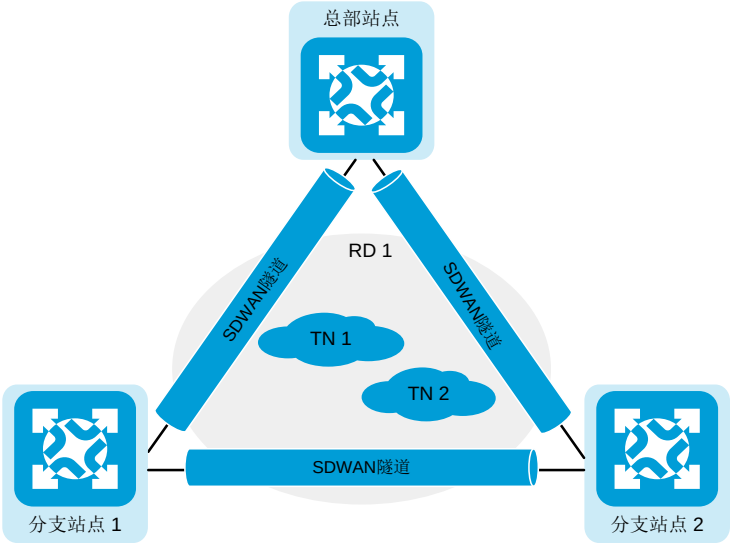


灵活拓扑技术

简介

如下图 1-21 所示，H3C 将支持 SDN 技术的广域网称为应用驱动广域网，即 AD-WAN。在 H3C 提出的 AD-WAN 分支解决方案中，用来实现站点互连的技术称为灵活拓扑技术。灵活拓扑旨在帮助用户降低广域网的开支、提升网络连接的灵活性，为分散在广阔地理范围内的企业网络、数据中心网络等提供安全可靠的互联服务。

图1-21 灵活拓扑技术



应用灵活拓扑技术的典型组网包含如下部分：

- CPE（Customer Premise Equipment，用户端设备）：站点网络的边缘设备。
- RR（Route Reflector，路由反射器）：用于在 CPE 之间反射路由信息。
- TN（Transport Network，传输网络）：运营商提供的广域接入网络，用来实现分支站点之间的互联，主要包括运营商专线网络和 Internet 公用网络等。传输网络可以通过 TN ID 或传输网络的名称来标识。TN 是构建 Overlay 网络的基础。
- RD（Routing Domain，路由域）：由彼此之间路由可达的多个传输网络构成的区域。属于同一个路由域的 CPE 之间或 CPE 与 RR 之间可以建立 SDWAN 隧道。
- Site ID：站点 ID，是分支站点在网络中的唯一标识，通常用一串数字表示，由 AD-WAN 管理控制组件统一自动分配。
- Device ID：设备 ID，是设备在站点内的唯一标识，由网络管理员统一分配。一个站点通常包含一台或两台 CPE/RR 设备。
- System IP：设备的系统 IP 地址，由网络管理员统一分配。通常采用设备上某个 Loopback 接口的 IP 地址作为 System IP。
- TTE（Transport Tunnel Endpoint，传输隧道端点）：CPE、RR 接入传输网的连接点和 SDWAN 隧道的端点。
- SDWAN 隧道：TTE 之间的点到多点逻辑通道，也叫 TTE 连接。不同站点之间通过 SDWAN 隧道传输数据报文等，以实现不同站点之间的互联。
- Interface ID：SDWAN 隧道接口 ID，由网络管理员统一分配。同一台设备上，不同 SDWAN 隧道接口的接口 ID 不同。

BGP 路由扩展

- BGP IPv4 Tnl-Encap-Ext 地址族

  为了支持灵活拓扑技术，BGP 定义了新的地址族——BGP IPv4 Tnl-Encap-Ext 地址族。通过该地址族交换的路由，称为 IPv4 Tnl-encap-ext 路由。在应用灵活拓扑技术的网络中，IPv4 Tnl-encap-ext 路由主要携带 TTE 信息、SaaS 访问路径质量信息。

  - TTE 信息

    TTE 信息是灵活拓扑技术的重要组成部分，是 TTE 相关信息的集合。TTE 信息通过 TTE ID 唯一标识。TTE ID 由 Site ID、Device ID 和 Interface ID 构成。CPE/RR 通过交互 TTE 信息，完成 SDWAN 隧道的动态建立和维护，实现网络管理的简化。

  - SaaS 访问路径质量信息

    SaaS 访问路径质量信息是指设备访问 SaaS 应用的时延、时延抖动、丢包率等，在 SaaS 路径优化功能中用于 SaaS 路径质量计算。

- EVPN 路由

  在应用灵活拓扑技术的网络中，CPE 将站点的私网路由信息以 EVPN IP 前缀路由的形式发布给其他 CPE。

通道建立

如下表 1-3 所示，在应用灵活拓扑技术的网络中，使用了管理通道、控制通道和数据通道三种通道技术。

表1-3 通道技术

通道类型	通道连接类型	通道所处位置	具体作用
管理通道	NETCONF连接	AD-WAN管理控制组件与CPE/RR之间	• AD-WAN 管理控制组件向CPE/RR 下发配置 • CPE/RR 向 AD-WAN 管理控制组件上报网络运维需要的信息
控制通道	SSL连接	RR与CPE之间	在CPE与RR之间交互TTE信息
	BGP连接	RR与CPE之间	CPE向RR发布TTE信息和私网路由，RR将TTE信息和私网路由反射给其他CPE
数据通道	SDWAN隧道	CPE之间、CPE与RR之间	转发数据报文

- 管理通道

  管理通道是指 AD-WAN 管理控制组件与 CPE/RR 之间的连接，包括 NETCONF 连接等。通过管理通道可以实现如下功能：

  - AD-WAN 管理控制组件通过管理通道（例如 NETCONF 连接）向 CPE/RR 下发配置，主要包括网络基础配置、VPN 业务参数、智能选路和 IPsec 等配置。
  - CPE/RR 等网络设备通过管理通道（例如 NETCONF 连接）向 AD-WAN 管理控制组件上报网络运维需要的信息，主要包括设备的告警、日志以及网络流量的性能采集信息。

- 控制通道

  控制通道是指 RR 和 CPE 之间用于发布 TTE 信息、私网路由的 SSL 连接和 BGP 连接。

  - SSL 连接：用于在 RR 与 CPE 之间发布 TTE 信息，以便在 RR 与 CPE 之间建立 SDWAN 隧道。
  - BGP 连接：用于在 CPE 之间发布 TTE 信息和私网路由。

  控制通道的具体建立过程为：

(1) CPE 与 RR 之间使用网络侧出接口地址建立 SSL 连接，CPE 与 RR 之间通过 SSL 报文交互各自的 TTE 信息。

(2) CPE 与 RR 收到对端发送的 TTE 信息后，比较 TTE 信息中携带的路由域。若路由域相同，则建立到达对端的 SDWAN 隧道；若路由域不同，则不建立 SDWAN 隧道。该 SDWAN 隧道可以用于在 CPE 与 RR 之间转发报文。

(3) 完成 SDWAN 隧道建立后，CPE、RR 自动在本地添加到达对端 Systerm IP 的路由。

(4) CPE、RR 之间基于 Systerm IP 建立 IPv4 Tnl-Encap-Ext 地址族和 EVPN 地址族下的 BGP 连接，通过 BGP 连接发布 TTE 信息和私网路由。

- 数据通道

数据通道是 CPE 之间、CPE 与 RR 之间转发数据报文的 SDWAN 隧道。具体的建立过程为：

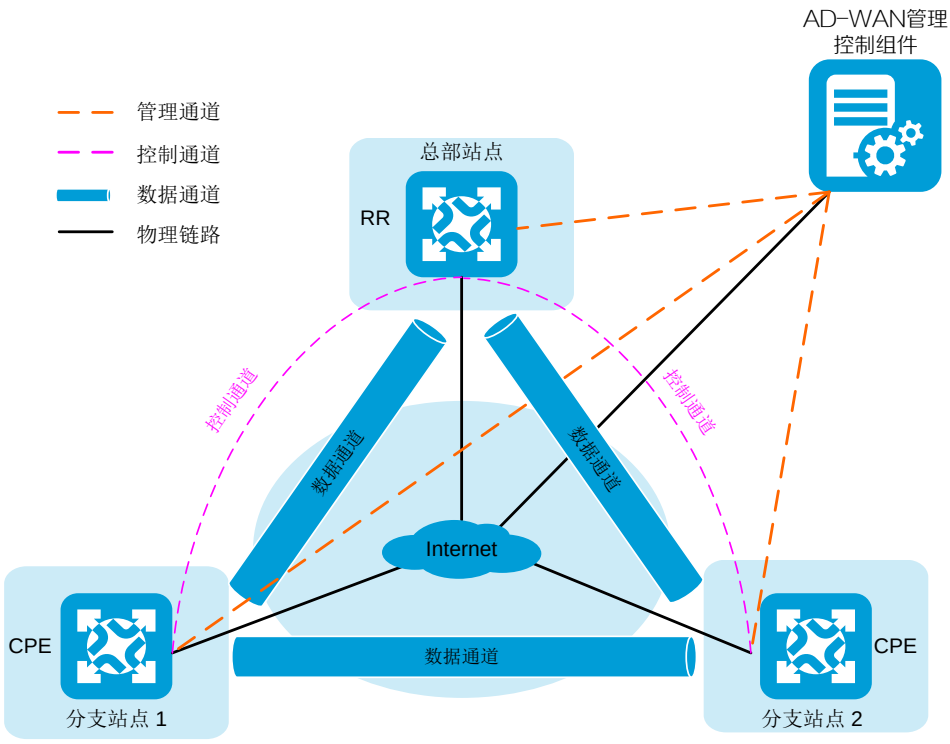
(1) CPE 通过 IPv4 Tnl-encap-ext 路由向 RR 发送 TTE 信息。

(2) RR 将 TTE 信息反射给其他 CPE。

(3) CPE 收到由 RR 反射的 TTE 信息后，比较 TTE 信息中携带的路由域。若路由域相同，则建立到达对端 CPE 的 SDWAN 隧道；若路由域不同，则不建立 SDWAN 隧道。

(4) 完成 SDWAN 隧道建立后，CPE 自动在本地添加到达对端 Systerm IP 的路由。

图1-22 通道示意图



私网路由发布

如下图 1-23 所示，在应用灵活拓扑技术的网络中，站点间的私网路由发布过程包含三部分：本地站点向 CPE 1 发布私网路由、CPE 1 与 CPE 2 之间通过 IP 前缀路由发布私网路由、CPE 2

向本地站点发布私网路由。完成上述路由发布过程后，本地站点的路由将发布到远端站点，以实现站点之间路由可达。

- 本地站点向 CPE 1 发布私网路由

  本地站点向 CPE 1 发布私网路由：本地站点 1 使用静态路由、RIP、OSPF、IS-IS、EBGP 或 IBGP，将本站点的 VPN 路由通过 IPv4 路由发布给本端 CPE。

- CPE 1 与 CPE 2 之间通过 IP 前缀路由发布私网路由

  CPE 1 与 CPE 2 之间通过 IP 前缀路由发布私网路由具体过程为：

(1) CPE 1 从本地站点 1 学到 VPN 路由信息后，将其存放到 VPN 实例的路由表中。

(2) CPE 1 将 VPN 路由信息封装到 IP 前缀路由中发布给 RR， IP 前缀路由携带 RD 和 RT，且该路由的下一跳地址为本端 CPE 1 的 System IP。

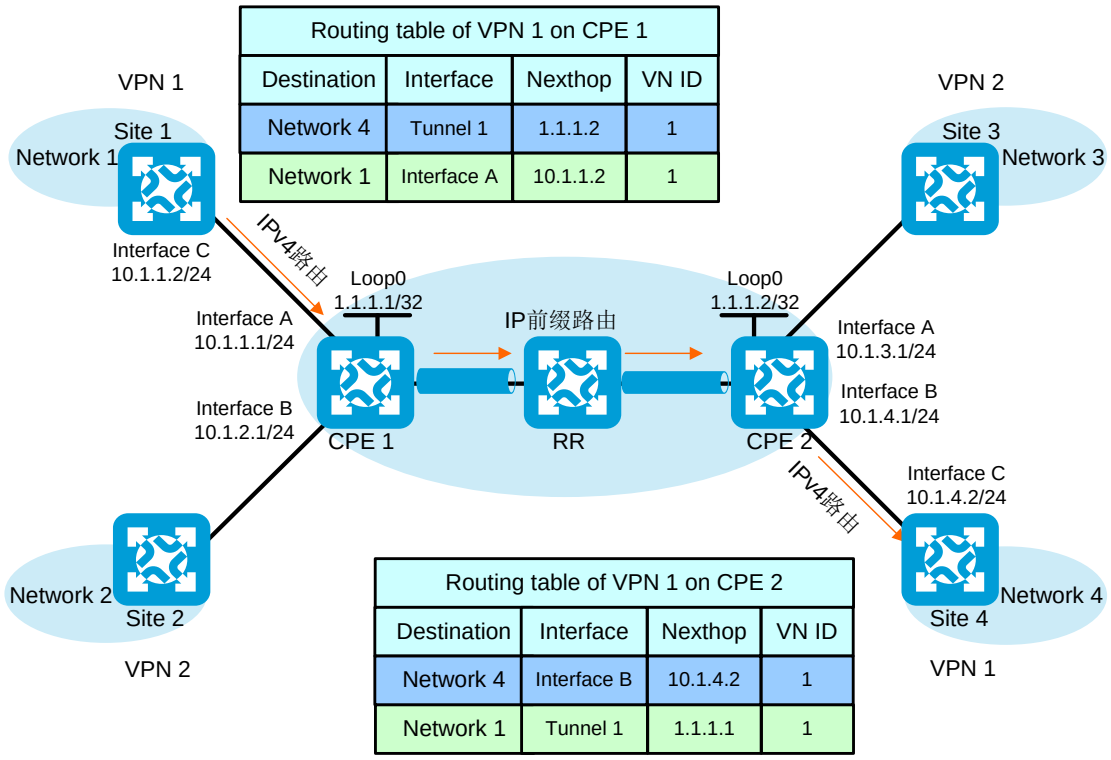
(3) RR 将收到的 IP 前缀路由反射给 CPE 2。

(4) CPE 2 收到 RR 反射的 IP 前缀路由后，对比路由中携带的 RT 属性和本地的 RT 属性，将 RT 匹配的 IP 前缀路由添加到 VPN 路由表中，其中路由的下一跳为 CPE 1 的 System IP，出接口为 SDWAN 隧道接口（该接口上建立了到达 CPE 1 的 System IP 的 TTE 连接）。

- CPE 2 向本地站点发布私网路由

  CPE 2 使用使用静态路由、RIP、OSPF、IS-IS、EBGP 或 IBGP，将 VPN 路由通过 IPv4 路由发布给本端 CPE 2。

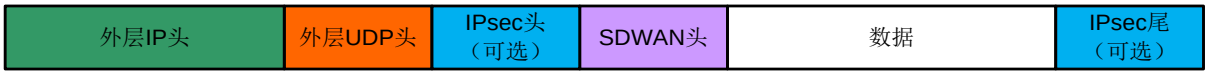
图1-23 私网路由发布



报文转发

如下图 1-24 所示，SDWAN 数据报文封装格式为：在原始数据报文外添加 12 字节 SDWAN 头、8 字节外层 UDP 头和 20 字节外层 IP 头；如需对报文进行安全保护，则还会封装 IPsec 头。其中 SDWAN 报文头中会携带标识数据报文所属的 VPN 实例的 VN ID。

图1-24 SDWAN 报文头



以下图 1-25 为例，数据报文的转发过程为：

(1) Device 1 发送的目的地址为 10.1.4.3 的 IP 报文到 CPE 1。

(2) CPE 1 根据报文到达的接口及目的地址查找对应 VPN 实例的路由表，找到路由的出接口为 SDWAN 隧道接口，下一跳为 CPE 2 的 Systerm IP。

(3) CPE 1 对报文进行 SDWAN 封装，然后沿着 SDWAN 隧道的出接口转发该报文。具体的封装信息如下：

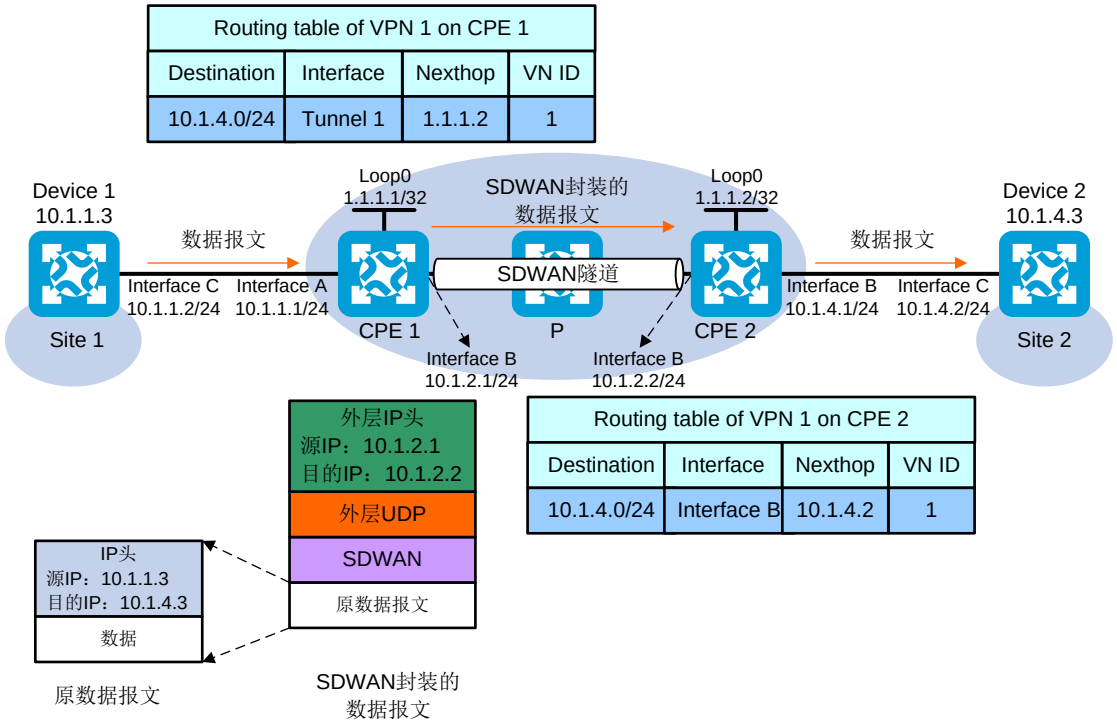
- SDWAN 头中的 VN ID 为 CPE 1 接收该报文的接口绑定的 VPN 实例的 VN ID。
- 外层 UDP 头中的源端口号为配置的 SDWAN 报文采用 UDP 封装时的源 UDP 端口号；
- 外层 IP 头的源/目的 IP 地址分别为 TTE 连接信息中的 Source IP 和 Destination IP，Source IP 即 CPE 1 上该 SDWAN 隧道的物理出接口的 IP 地址，Destination IP 即 CPE 2 上接收 SDWAN 报文的物理接口的 IP 地址；

(4) P 设备根据报文的目的 IP 地址将其转发到 CPE 2。

(5) CPE 2 对报文进行解封装，根据 SDWAN 报文头中的 VN ID 确定报文所属的 VPN 实例，通过查找该 VPN 实例的路由表，确定报文的出接口，并将解封装后的报文转发至 Device 2。

(6) Device 2 按照正常的 IP 转发过程将报文转发给目的主机。

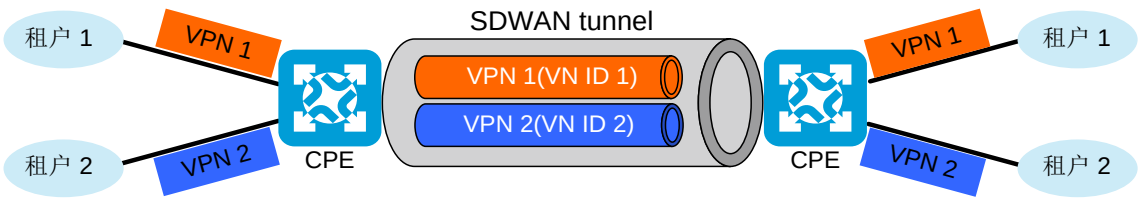
图1-25 数据报文转发过程



租户隔离

如下图 1-26 所示，在应用灵活拓扑技术的组网中，可以使用 VPN 实例为租户提供路由表、数据报文的隔离功能。VPN 实例的数据报文通过 VN ID 标识，当不同 VPN 实例的数据报文可以共享同一条 SDWAN 隧道进行传输，通过 VN ID 还可以减少设备间隧道数量，减少网络资源的消耗。

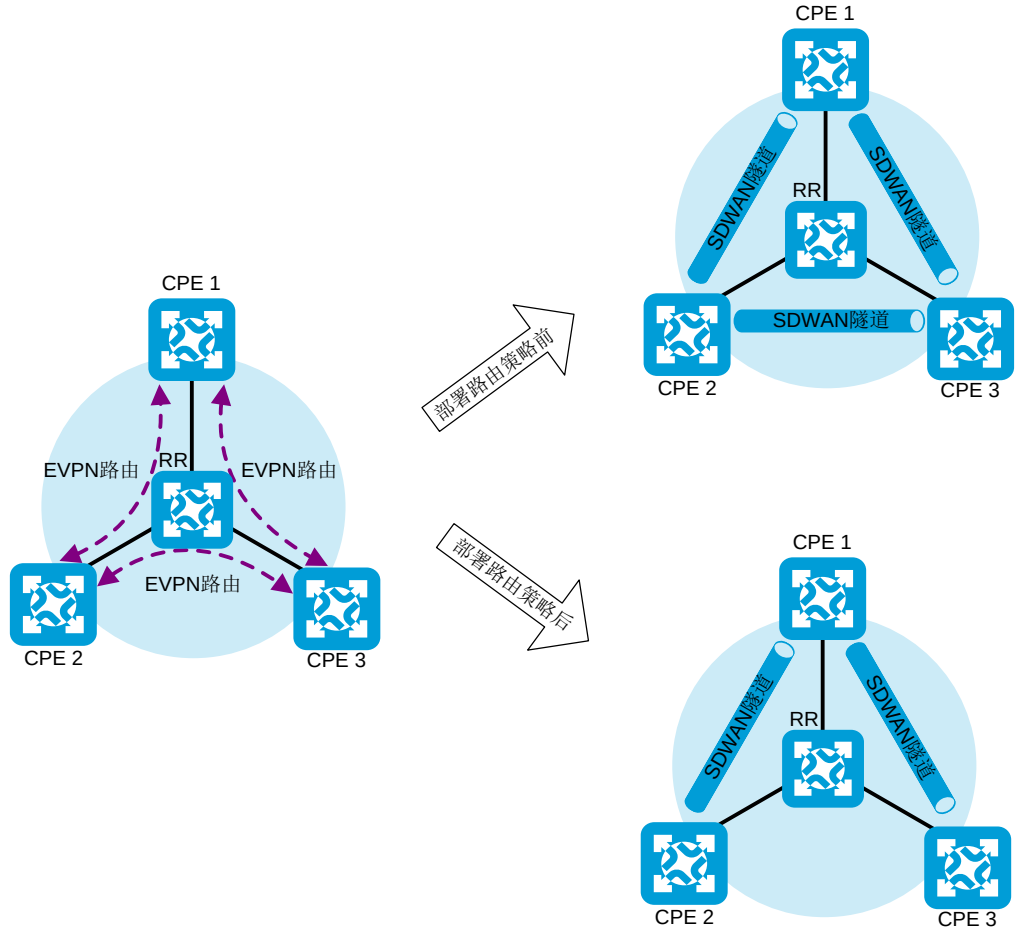
图1-26 租户隔离



灵活控制网络拓扑

如下图 1-27 所示，在 Underlay 网络互连的基础上，通过在 RR 上配置路由策略灵活地控制 Overlay 网络的拓扑结构。

图1-27 灵活控制网络拓扑



在部署路由策略前，CPE 2 和 CPE 3 之间存在 SDWAN 隧道，CPE 2 和 CPE 3 之间的流量直接通过两者之间的 SDWAN 隧道进行转发。由于租户业务的变更，需要 CPE 2 和 CPE 3 之间的流量绕行到 CPE 1 进行转发。此时，可以在 RR 上部署路由策略，将 CPE 2 和 CPE 3 发送的 EVPN 路由的下一跳地址修改为 CPE 1 的地址，从而灵活改变 Overlay 网络的拓扑。

NAT 会话穿越

产生背景

P2P（Point to Point，点到点）应用为互联网的主要应用之一。通常情况下，P2P 应用服务器位于公网，P2P 客户端位于私网。为了私网中的 P2P 客户端能够访问公网 P2P 应用服务器，需要部署 NAT 设备将私网 IP 地址转换为公网 IP 地址，并转换端口号。不同网络环境中部署的 NAT 设备，其地址映射行为和报文过滤行为不尽相同。因此部署 NAT 设备后可能会带来如下问题：

- 不是所有的私网的 P2P 客户端之间均能直接建立 P2P 连接。
- 私网的 P2P 客户端间建立 P2P 连接的交互过程会有所不同。

STUN（Session Traversal Utilities for NAT，NAT 环境下的会话传输）是一种 C/S 架构的网络协议，用于帮助各类应用协议穿越 NAT。STUN 协议根据 NAT 设备的地址映射行为和报文过滤行为这两类特征，定义了不同的 NAT 类型。在 P2P 组网中，STUN 协议使得 P2P 客户端具备了探测所在网络中的 NAT 设备和 NAT 类型，以及发现自己公网 NAT 地址的能力，继而 P2P 客户端可以根据对端 NAT 设备的 NAT 类型选择合适的交互过程建立穿越 NAT 设备的 P2P 连接。

四种 NAT 类型

STUN 协议工作的核心是，根据 NAT 设备的映射和过滤行为判断出 NAT 类型。各种 NAT 类型的特征如下表 1-4 所示。

表1-4 各种 NAT 类型的特征

NAT 类型	映射行为	过滤行为
完全锥型NAT	• 与私网发往公网的报文的目的 IP 和端口号无关 • 所有从同一个私网 IP 和端口号发送到任意公网 IP 和端口号的报文，地址转换后的结果相同	所有来自公网的报文，均可通过过滤
限制锥型NAT		仅来自公网指定IP的报文，才可以通过过滤
端口限制锥型NAT		仅来自公网指定IP且使用指定端口号的报文，才可以通过过滤
对称NAT	• 与私网发往公网的报文的目的 IP 和端口号有关 • 从同一个私网 IP 和端口号发送到不同公网 IP 和端口号的报文，地址转换后的结果不同	

NAT 类型对 P2P 连接的影响

不同 NAT 类型的映射行为和过滤行为对建立 P2P 连接的影响如下：

- 映射行为： P2P 客户端经过 NAT 设备主动访问某个外部节点，NAT 设备会建立地址映射表项，相当于为该外部节点访问 P2P 客户端创建了一个通道。不同的映射行为决定了 P2P 客户端访问不同的外部节点时，是否会创建不同的通道。
- 过滤行为：NAT 设备会对公网主动发往私网的报文进行过滤。对于不同的过滤行为，外部节点通过上述通道与私网中的 P2P 客户端建立连接时需要满足的条件不同。

由于上述影响，对于不同私网中的 P2P 客户端，只有各自所在网络中 NAT 设备的 NAT 类型为下表 1-5 所示的组合时，P2P 客户端之间才能通过一定的交互方式直接建立 P2P 连接。对于 P2P 客户端之间无法直接建立 P2P 连接的情况，需要网络管理员部署中转设备转发 P2P 客户端之间的数据报文。

表1-5 NAT 类型对 P2P 连接影响

本端 NAT 设备的 NAT 类型	对端 NAT 设备的 NAT 类型	私网 P2P 客户端之间能否直接建立 P2P 连接
完全锥型NAT	完全锥型NAT	√
完全锥型NAT	端口限制锥型NAT/ 限制锥型NAT	√

本端 NAT 设备的 NAT 类型	对端 NAT 设备的 NAT 类型	私网 P2P 客户端之间能否直接建立 P2P 连接
完全锥型NAT	对称NAT	√
端口限制锥型NAT/ 限制锥型NAT	端口限制锥型NAT/ 限制锥型NAT	√

STUN 探测 NAT 类型

- STUN 的通信模式

私网 P2P 客户端判断自己所在网络中 NAT 设备的 NAT 类型时，需要在该客户端，以及客户端经过 NAT 设备后路由可达的某个公网节点上开启 STUN 功能。STUN 功能采用客户端/服务器通信模式，STUN 服务器（STUN Server）和 STUN 客户端（STUN Client）之间的基本工作流程如下：

(1) STUN 客户端主动向 STUN 服务器发送探测请求报文。

(2) STUN 服务器将 STUN 客户端地址转换后的公网 IP 和端口号填充到探测响应报文中。

(3) STUN 客户端从探测响应报文中获取并记录自己地址转换后的公网 IP 和端口号，并通过比较不同的探测响应报文中的公网 IP 和端口号来判断 NAT 类型。

- STUN 探测机制

STUN 客户端通过 STUN 服务器探测本端 NAT 设备的 NAT 类型时，STUN 服务器需要拥有两个公网 IP 地址，以及两个监听探测请求的端口号。具体的探测机制如下图 1-28 、图 1-29 、图 1-30 和图 1-31 所示。

图1-28 判断是否存在 NAT 设备

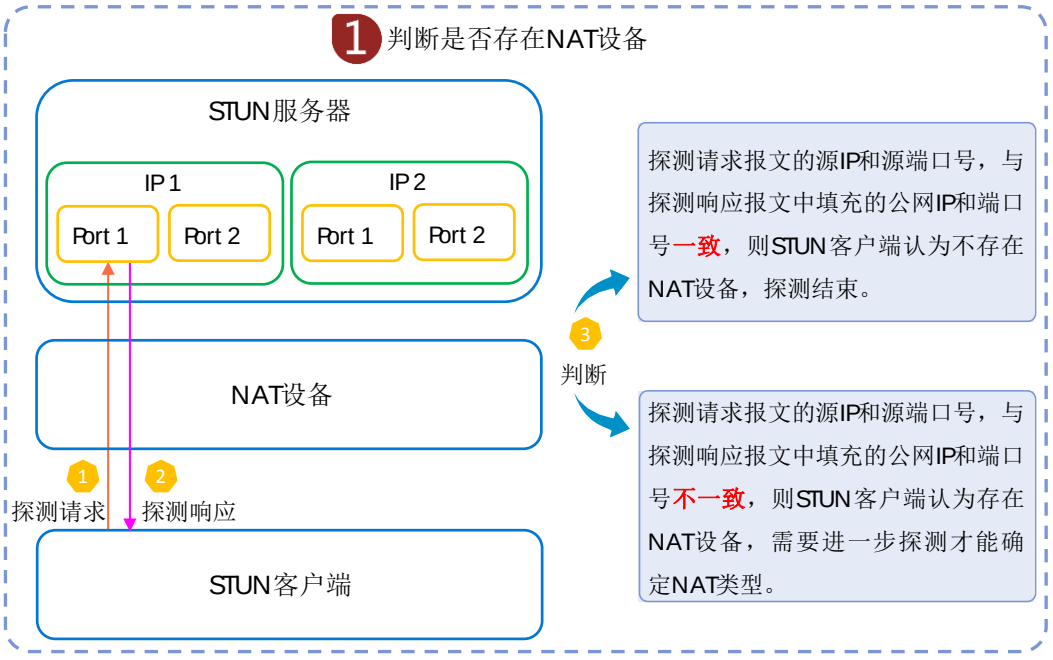


图1-29 判断是否为完全锥形 NAT

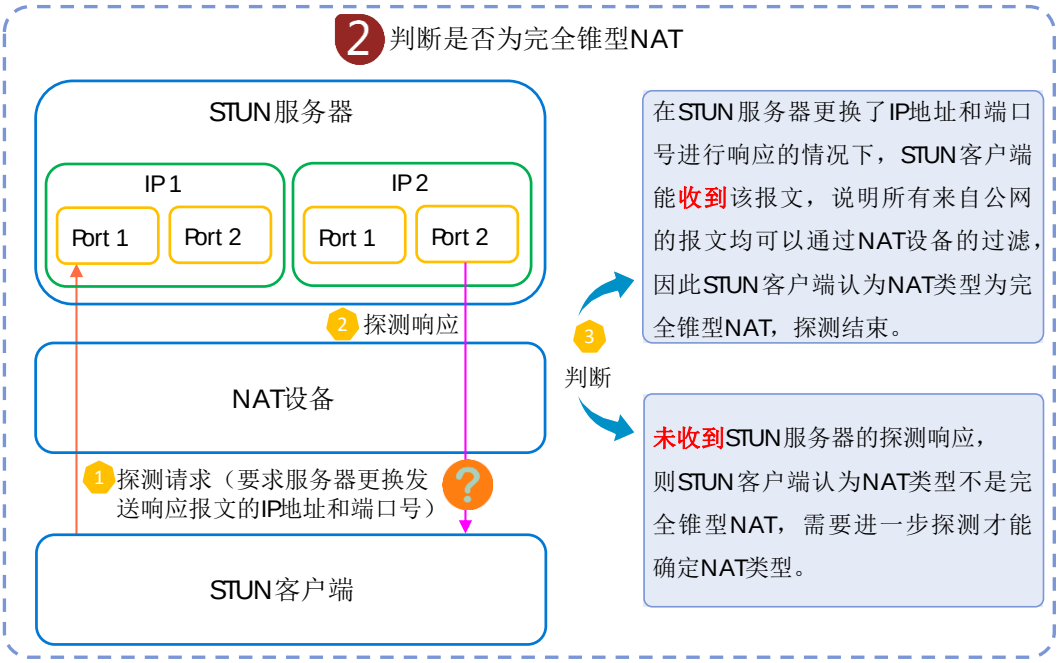


图1-30 判断是否为对称 NAT

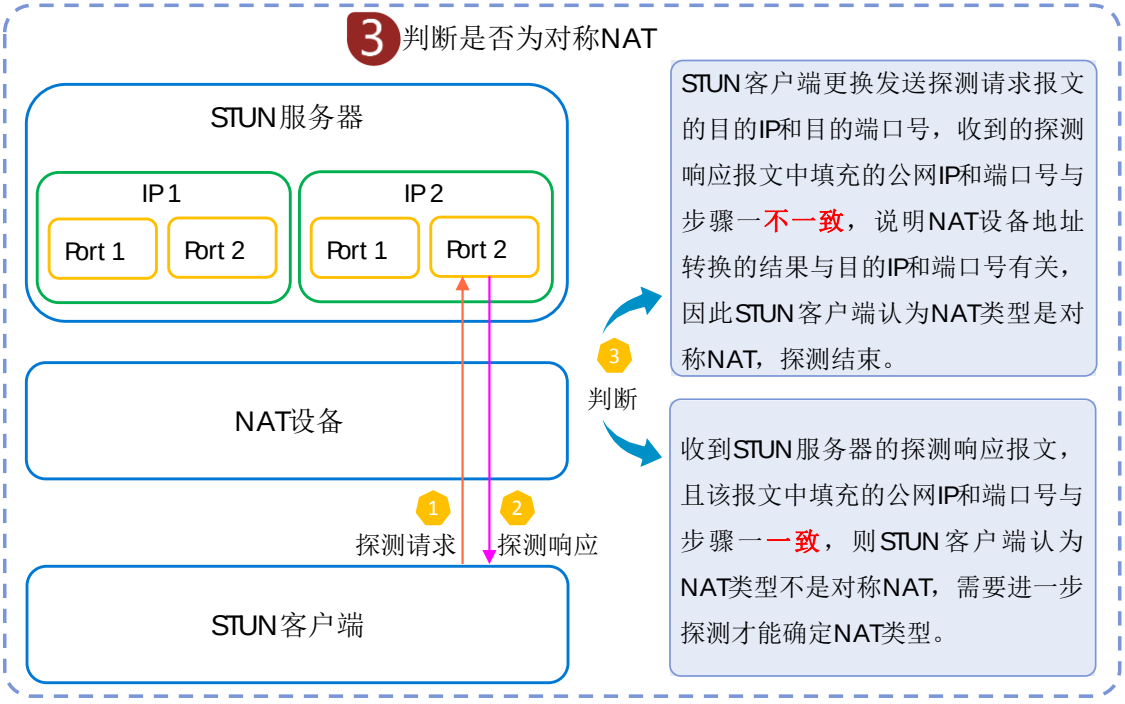
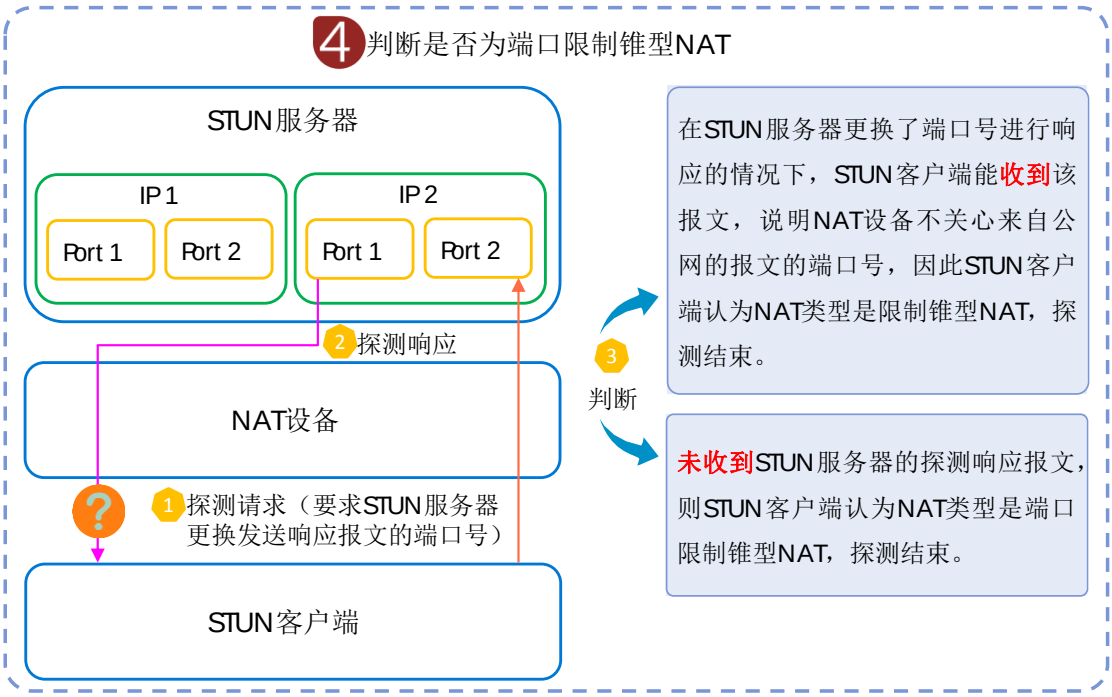


图1-31 判断是否为端口限制锥形 NAT



典型应用

在 AD-WAN 网络中，CPE（Customer Provided Edge，用户提供的网络边缘）之间需要建立数据通道，以便进行用户数据报文转发。CPE 之间建立的数据通道就属于 P2P 类型的连接。

如下图 1-32 所示，不同分支站点的 CPE 穿越 NAT 设备建立 P2P 连接时，需要在 CPE 1、CPE 2 和 RR（Route Reflector，路由反射器）上开启 STUN 功能。开启 STUN 功能后，CPE 之间建立数据通道的步骤如下：

(1) CPE 1 和 CPE 2 作为 STUN 客户端主动向作为 STUN 服务器的 RR 发送探测请求报文。

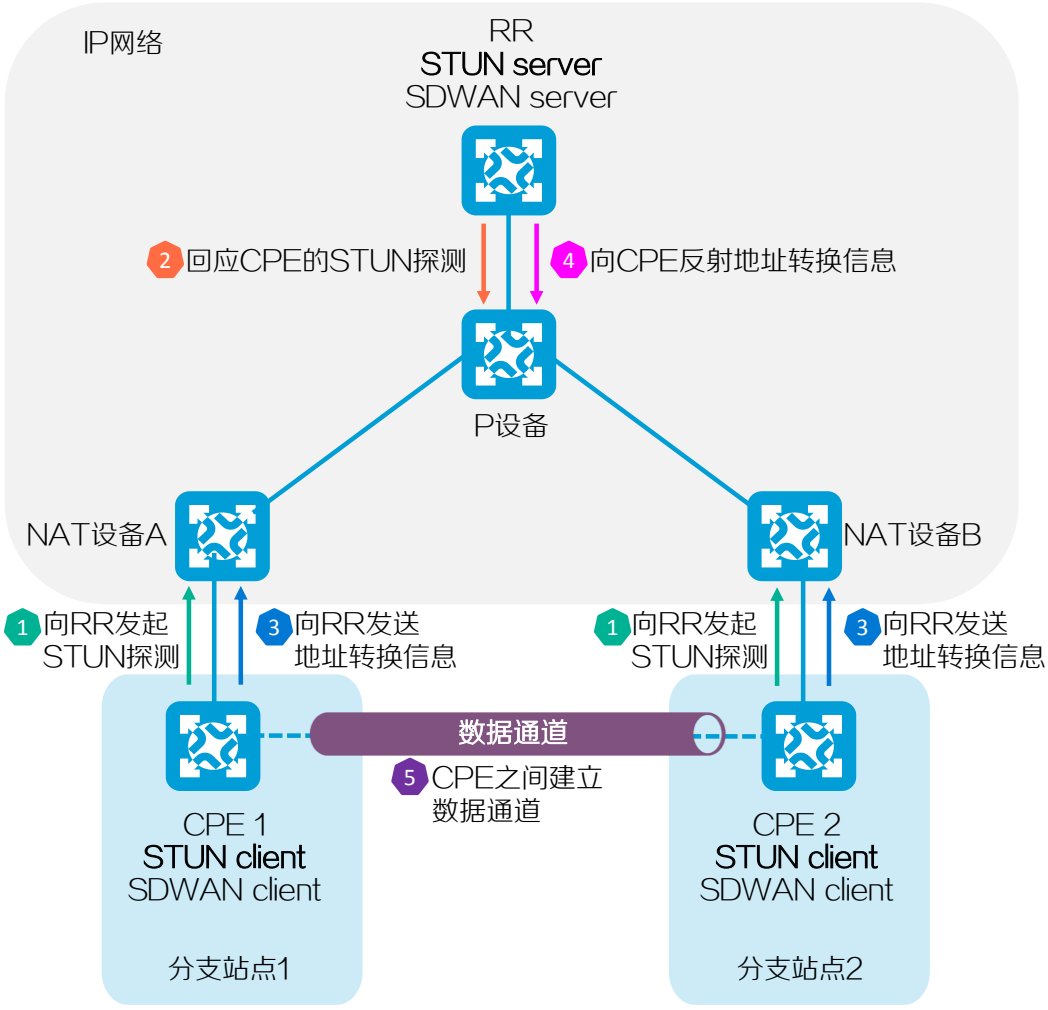
(2) CPE 1 和 CPE 2 从 RR 回复的响应报文中获取到自己地址转换后的公网 IP 和端口号，并通过与 STUN 服务器多次交互协议报文确定 NAT 类型。

(3) CPE 1 和 CPE 2 将自己地址转换后的公网 IP 和端口号、NAT 类型发送给 RR。

(4) RR 将 CPE 1 和 CPE 2 的公网 IP 和端口号、NAT 类型反射给对端 CPE。

(5) CPE 1 和 CPE 2 通过对端的公网 IP 和端口号、NAT 类型确定是否可以直接建立数据通道，以及建立数据通道的交互过程。对于 CPE 之间无法直接建立数据通道的情况，需要网络管理员部署 NAT transfer 设备转发 CPE 之间的数据报文。

图1-32 STUN 典型应用



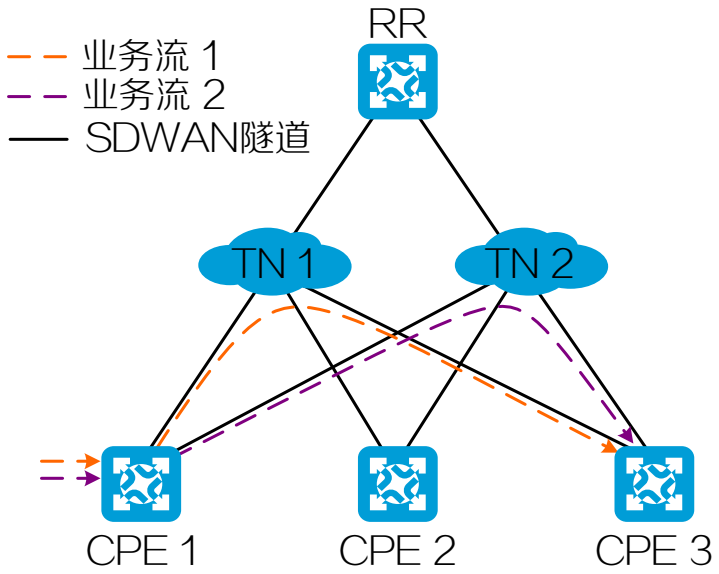
智能选路

简介

传统的链路优选一般基于链路开销或路由策略，无法根据实际的业务需求选择最适合的链路。RIR（Resilient Intelligent Routing，智能选路）可以根据不同业务流量的链路需求，如链路质量、链路带宽等，为其选择最适合的链路。如果业务流量当前选择的链路由于链路状态变化而不再符合要求，智能选路还可以自动将流量 CPE 切换到另一条符合要求的链路上。

如下图 1-33 所示，将 RIR 应用于 AD-WAN 网络，根据链路属性，如链路优先级、链路质量和链路带宽等，在 CPE、RR 设备之间为不同业务流量选择不同的 SDWAN 隧道进行传输。

图1-33 智能选路



- CPE（Customer Premise Equipment，用户端设备）：用户网络的边缘设备。
- RR（Route Reflector，路由反射器）：用于在 CPE 之间反射 TTE（Transport Tunnel Endpoint，传输隧道端点）信息和私网路由等。
- TN（Transport Network，传输网络）：运营商提供的广域接入网络，用来实现分支站点之间的互联，主要包括运营商专线网络和 Internet 公用网络等。传输网络可以通过 TN ID 或传输网络的名称来标识。TN 是构建 AD-WAN Overlay 网络的基础。
- SDWAN 隧道：AD-WAN 设备之间的点到多点逻辑通道。不同站点之间通过 SDWAN 隧道传输数据报文等，实现不同站点之间的互联。
- 业务流：具有相同特征（例如，MAC 地址、目的 IP 地址）的业务流量称为同一类业务流。设备为同一类业务流采用相同的选路策略，基于相同的选路策略为同一类业务流内的不同会话分别进行智能选路，从而实现更精细地链路管理。通过五元组（源 IP 地址、目的 IP 地址、源端口、目的端口以及传输层协议）可唯一定义一个会话。

链路属性

如下表 1-6 是链路属性的相关概念和说明。

表1-6 链路属性

链路属性	相关概念	说明
链路优先级	• 用户可以根据链路特征、业务需求等因素定义链路优先级，如，根据链路价格 • 同一业务流模板下的不同链路可以具有相同的优先级	链路优先级数值越小优先级越高
链路质量	质量探测： • 链路通断探测：设备通过发送 Keepalive 报文等方式探测每条链路的通断情况 • 链路质量探测：设备通过 iNQA（Intelligent Network Quality Analyzer，智能网络质量分析）功能探测每条链路的时延、抖动和丢包率 基于质量探测结果进行质量评估： • 设备基于 SLA（Service Level Agreement，服务等级协议）协议对链路质量进行评估，包括链路延迟、抖动、丢包率等 • 设备通过 CQI（Comprehensive Quality Indicator，综合质量指标）算法及各链路延迟、抖动、丢包率等数据，计算出链路的综合质量 CQI 值。取综合质量 CQI 值可以被 5 整除的最大数作为综合质量近似 CQI 值（例如，如果综合质量 CQI 值为 82.5，则综合质量近似 CQI 值为 80）	• 若链路的综合质量近似 CQI 值＜100，则认为该链路不符合业务质量要求 • 若链路的综合质量近似 CQI 值＝100，则认为该链路符合业务质量要求
链路带宽	通过计算带宽使用率，来判断链路的繁忙程度，从而确定链路是否满足业务的带宽需求。带宽使用率＝（链路或所属物理接口的已使用带宽+会话预计使用的带宽）/链路或所属物理接口的总带宽	同时满足以下条件则认为符合带宽要求 • 待选链路所属物理接口的带宽使用率＜80% • 待选链路的带宽使用率＜80%

智能选路机制

如下图 1-34 和表 1-7 所示是智能选路的机制。

图1-34 智能选路机制

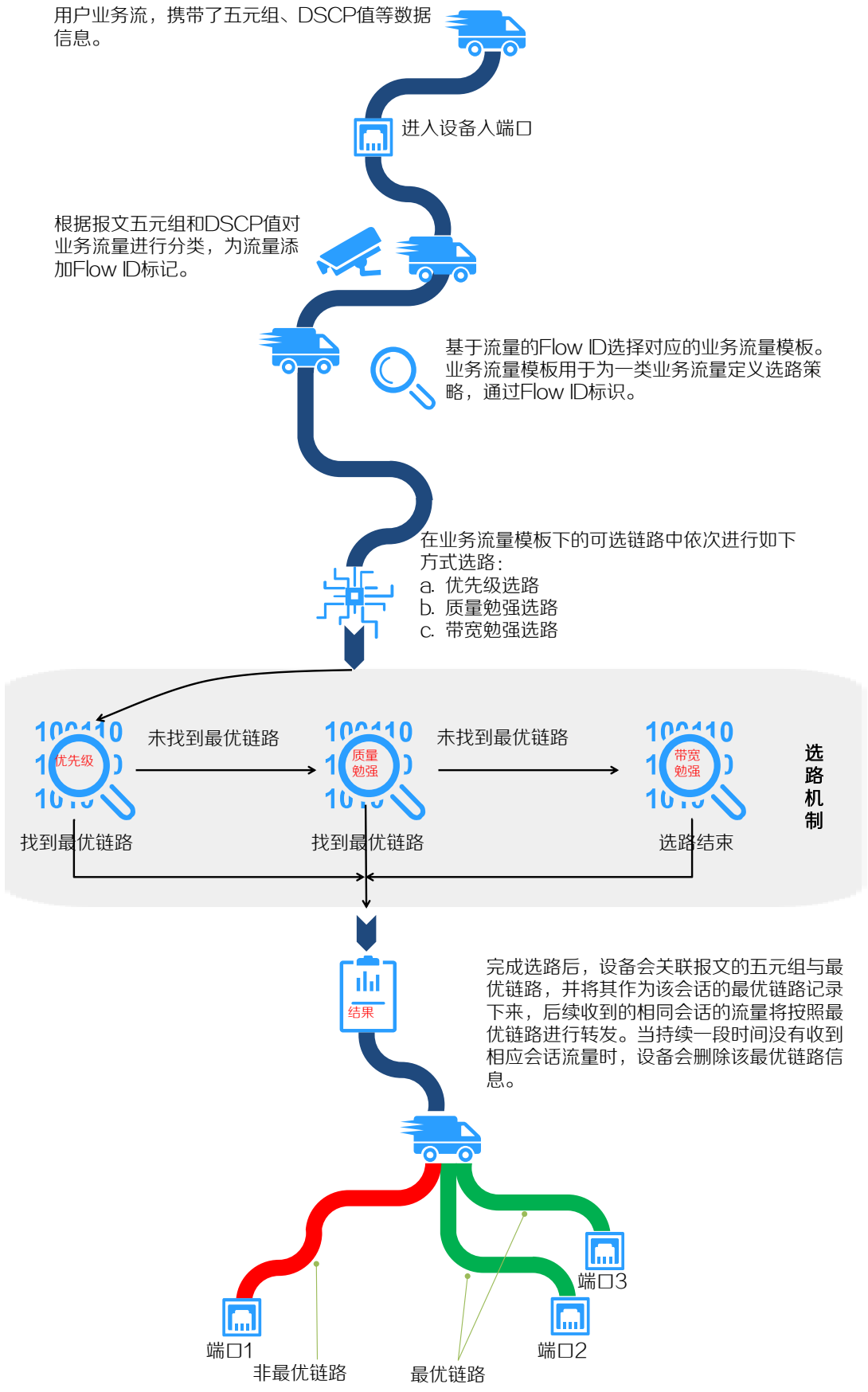

用户业务流，携带了五元组、DSCP值等数据信息。
进入设备入端口
根据报文五元组和DSCP值对业务流量进行分类，为流量添加Flow ID标记。
基于流量的Flow ID选择对应的业务流量模板。业务流量模板用于为一类业务流量定义选路策略，通过Flow ID标识。
在业务流量模板下的可选链路中依次进行如下方式选路：
a. 优先级选路
b. 质量勉强选路
c. 带宽勉强选路
优先级
未找到最优链路
质量勉强
未找到最优链路
带宽勉强
选路机制
找到最优链路
找到最优链路
选路结束
结果
完成选路后，设备会关联报文的五元组与最优链路，并将其作为该会话的最优链路记录下来，后续收到的相同会话的流量将按照最优链路进行转发。当持续一段时间没有收到相应会话流量时，设备会删除该最优链路信息。
端口3
端口1
端口2
非最优链路
最优链路

表1-7 选路方式

选路方式	选路规则	选路结果
优先级选路	在优先级选路过程中，按照链路优先级从高到低的顺序，依次判断链路是否符合业务质量、带宽要求： • 若链路符合要求，则结束选路 • 若链路不符合要求，则继续选择次优先级的链路进行选路	• 若选出最优链路，则通过最优链路转发报文 • 若未找到最优链路，则继续进行质量勉强选路
质量勉强选路	在质量勉强选路过程中，按照链路的综合质量近似CQI值从高到底的顺序，依次判断链路是否符合业务带宽要求： • 若链路符合要求，则结束选路 • 若链路不符合要求，则继续选择次综合质量近似 CQI 值的链路进行选路	• 若选出最优链路，则通过最优链路转发报文 • 若未找到最优链路，则继续进行带宽勉强选路
带宽勉强选路	判断链路的带宽占用率是否＜100%： • 若链路带宽占用率＜100%，则为最优链路 • 若链路带宽占用率≥100%，则为非最优链路	• 若选出最优链路，则通过最优链路转发报文 • 若未找到最优链路，则按原路由表项进行普通转发

选路延迟和选路抑制

为提高报文转发效率，业务流量完成第一次智能选路后，后续相同业务的流量均按照第一次选路结果进行转发。当业务流量模板中的任一链路发生如下任一变化时，设备会重新选路：

- 链路质量在满足业务质量要求、不满足业务质量要求之间切换。
- 链路带宽使用率≥90%，或者链路所属的物理接口的带宽使用率≥90%。

为避免链路震荡时设备频繁选路，智能选路定义了选路延迟时间和选路抑制周期。设备执行一次选路后，若配置了选路抑制周期，则会进入选路抑制周期。

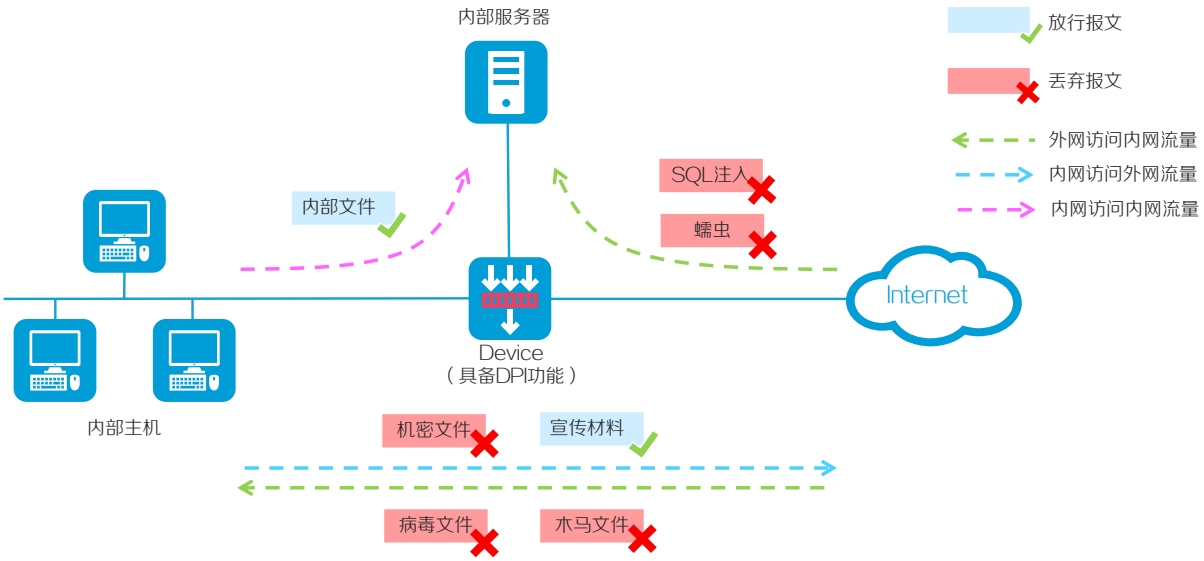
- 在选路抑制周期内：
  - 不会重新选路。
  - 会一直维护链路状态数据。
- 选路抑制周期结束后：
  - 选路延迟时间内一直满足重新选路条件，则在选路延迟时间超时后重新选路。
  - 选路延迟时间未一直满足重新选路条件，则在选路延迟时间超时后不会重新选路。

DPI 应用识别

什么是 DPI 应用识别

如下图 1-35 所示，DPI（Deep Packet Inspection，深度报文检测）是一种基于应用层信息对流量进行检测和控制的安全功能。DPI 支持丰富的业务类型（如防病毒、应用审计与管理等），可以阻断外部攻击、防止内部数据泄漏、规范用户上网行为，极大地提高了网络的安全性。

图1-35 DPI 简介



深度检测“深”在哪里

如下图 1-36 所示，普通四层检测，仅对报文的二层到四层内容进行识别。DPI 深度检测不仅可以检测二层到四层信息，还可以对应用层信息（如 HTTP 数据）进行检测。DPI 深度检测的结果更加具体、精确，更能满足用户需求。

图1-36 普通四层检测与 DPI 深度检测

如何实现深度检测

DPI 通过应用层检测引擎对报文进行特征匹配，来实现深度检测。

- 应用层检测引擎是 DPI 功能的核心处理单元。它用来对报文进行协议解析、特征匹配，以及向各 DPI 业务模块通知检测结果。
- 特征是由专业的攻防团队对各类业务报文进行分析而归纳出的、具有特定格式的业务特征数据。

- 特征库是各业务的特征集合。从官方路径获取各业务特征库后，将其加载到设备，以便应用层检测引擎调用。

DPI 深度检测流程如下图 1-37 所示：

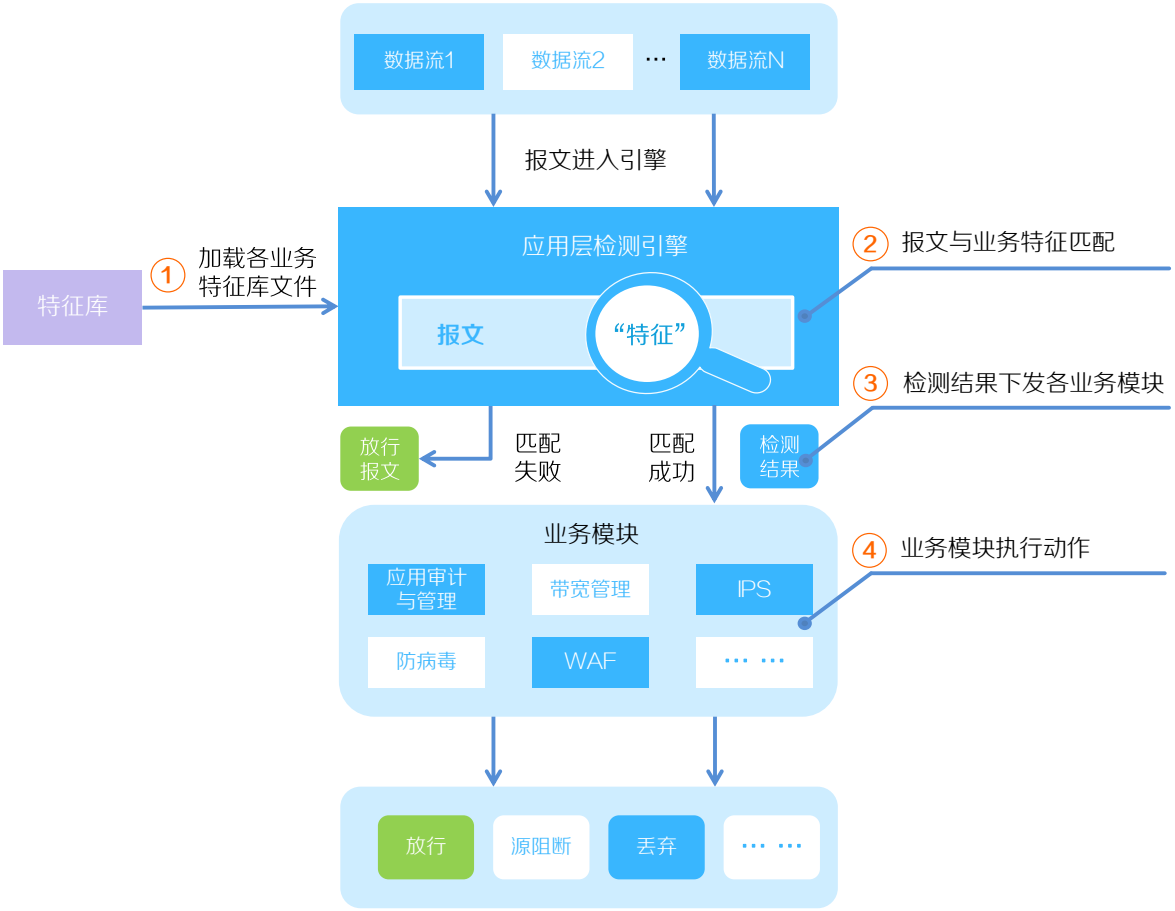
(1) 设备加载各个业务的特征库文件（例如防病毒特征库、IPS 特征库等），为引擎提供丰富的特征资源。

(2) 引擎通过将报文与特征匹配，对报文内容进行识别。

(3) 引擎对检测结果进行处理：如果报文与特征匹配成功，设备会将检测结果发送给相应的 DPI 业务模块；如果报文与特征匹配失败，则直接允许报文通过，不进行任何 DPI 业务处理。

(4) 各业务模块根据引擎检测结果对报文进行相应的处理。

图1-37 DPI 深度检测流程图



DPI 业务有哪些

- 应用识别

应用识别是 DPI 的基础业务，可基于不同应用的特征信息，识别出具体的应用。其他业务可使用应用识别结果做进一步处理。

图1-38 应用识别

- 应用审计与管理

  应用审计与管理是在应用识别的基础上进一步识别出应用的具体行为和行为内容，据此对用户的上网行为进行审计和记录。

图1-39 上网行为应用审计与管理

- 带宽管理

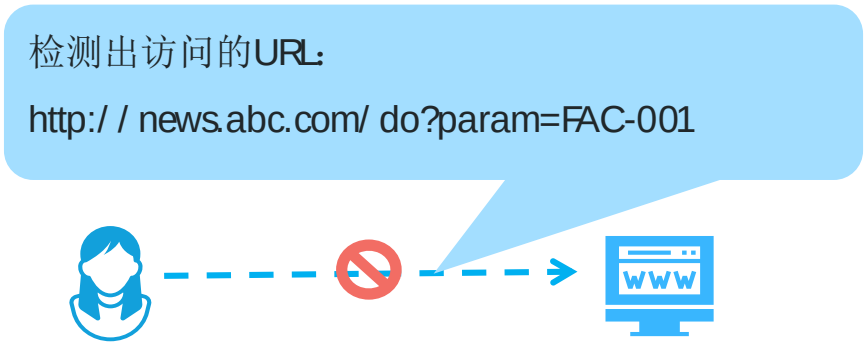
  带宽管理可基于安全域、用户、应用（通过应用识别业务检测出的应用）和时间段等限制条件，对通过设备的流量进行精细化的带宽管理和控制。

图1-40 设备带宽管理

- URL 过滤

  URL 过滤功能可对用户访问的 URL 进行识别和限制，即允许或禁止访问某个 URL，达到规范用户上网行为的目的。

图1-41 URL 过滤



- 数据过滤

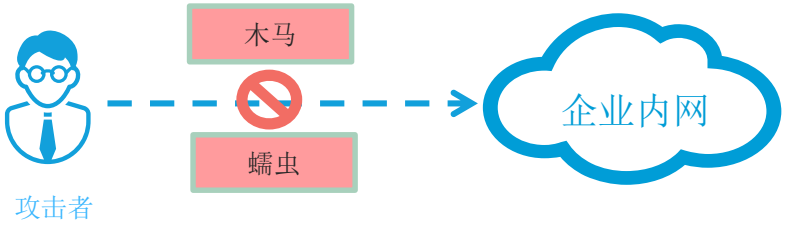
  数据过滤功能可对应用层协议报文中携带的内容进行过滤，以防止企业机密信息泄露、阻止非法和敏感信息的传播。

图1-42 数据过滤

- IPS

  IPS（Intrusion Prevention System，入侵防御系统）是一种对应用层攻击进行检测并防御的安全防御技术。IPS 通过分析网络流量来实时检测入侵行为，并通过响应动作来阻断入侵行为，确保企业信息系统和网络免受攻击。

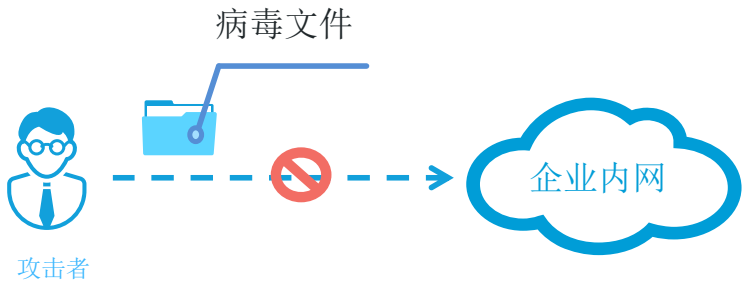
图1-43 IPS



- 防病毒

  防病毒功能是一种通过对报文应用层信息进行检测来识别和处理病毒报文的安全机制。防病毒功能凭借庞大且不断更新的病毒特征库可有效保护网络安全，防止病毒在网络中的传播。

图1-44 防病毒功能

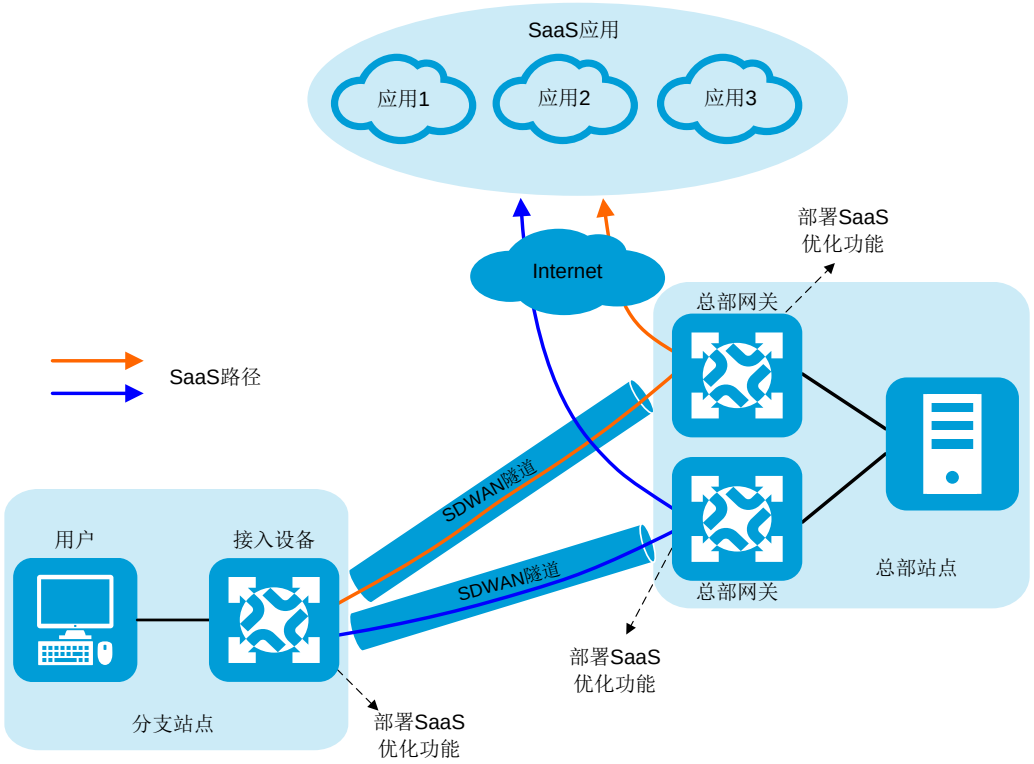


SaaS 路径优化

简介

智能选路可以为 AD-WAN 网络中不同站点之间的业务选择最优路径。但当用户访问部署在公网中的 SaaS 应用时，智能选路技术不能综合考虑用户与 SaaS 应用之间的整条链路（SDWAN 隧道+IP 网络），来进行智能选路，无法为用户提供访问 SaaS 应用的最优路径。SaaS 路径优化功能通过综合考虑设备访问 SaaS 应用的整条路径的质量，使用户通过最优路径访问 SaaS 应用，减少访问时间，提升用户的上网体验，如下图 1-45 所示。

图1-45 SaaS 路径优化简介



访问方式

用户网络支持的公网接入方式不同时，用户访问 SaaS 应用的方式也会有所不同。SaaS 应用的访问方式包括：

- 直接访问方式：当分支网关可以直接接入公网时（无需借助总部网关接入公网），用户通过分支网关设备的接口直接访问 SaaS 应用。
- 间接访问方式：当用户必须借助总部网关接入公网时（本地网络不支持直接接入公网），需要先通过接入设备到达网关设备，再通过网关设备间接访问 SaaS 应用。
- 直接/间接混合方式：用户既可以通过分支网关直接访问 SaaS 应用，也可以通过总部网关间接访问 SaaS 应用。

不同访问方式下，SaaS 应用的部署位置、路径探测和评估方法等，详见下文介绍。

SaaS 路径质量探测

- 健康监测 URL

  一个 SaaS 应用通常通过多个 URL 为用户提供服务。这些 URL 对应的服务器常常位于相近的网络位置（例如同一个数据中心），因此访问不同 URL 的路径具有相同的路径质量。获取访问 SaaS 应用的路径（即 SaaS 路径）的质量时，如果针对每一个 URL 都进行一次探测，则耗时较长，且网络资源占用较多。

  为了提高 SaaS 路径质量探测效率，可为每个 SaaS 应用指定一个健康监测 URL。设备到健康监测 URL 的路径质量，就能够反映出设备到 SaaS 应用中所有 URL 的路径质量。因此只需检测设备到健康监测 URL 的路径质量，即可获取到 SaaS 路径的质量，从而减少进行质量探测的路径数量。

  可以直接采用为用户提供 SaaS 应用服务的多个 URL 中的一个作为健康监测 URL 。

- 路径质量探测方法

  对于不同类型的路径，SaaS 路径质量探测的方法不同：

  - 对于 SDWAN 隧道，设备采用 iNQA（Intelligent Network Quality Analyzer，智能网络质量分析），来探测 SDWAN 隧道质量。
  - 对于 IP 路径，设备每隔 30 秒向健康监测 URL 发送 Ping 报文，以便获取设备到健康监测 URL 的路径质量（时延、时延抖动、丢包率等）。

SaaS 路径质量评估和优选机制

通过 CQI（Comprehensive Quality Indicator，综合质量指标）算法，可对 SaaS 路径的质量进行评估，从而计算出综合质量近似 CQI 值，并根据综合质量近似 CQI 值优选出 SaaS 路径。

- 根据 CQI 算法计算综合质量近似 CQI 值

  为避免链路频繁切换，设备使用综合质量近似 CQI 值评估链路优劣。SaaS 路径质量包括时延、时延抖动和丢包率三项指标。针对每一项指标，都存在如下两个值：

- 路径质量期望值。
- 路径质量探测值。

CQI 算法根据路径质量的期望值和探测值，计算出综合质量 CQI 值。该值越大，表示 SaaS 路径质量越好。SaaS 路径 CQI 算法的计算方法为：

(1) 当 SaaS 路径的单项质量（时延、时延抖动或丢包率）探测值≤SaaS 路径质量期望值时，认为该 SaaS 路径的单项 CQI 值为 100。

(2) 当单项质量探测值＞SaaS 路径质量期望值时，单项 CQI 值=（单项期望值 * 100）/单项质量探测值。

(3) 综合质量 CQI 值=（x ·Ds+ y · Js + z ·Ls）/（x + y + z）。其中 x、y、z 分别为时延、时延抖动和丢包率的权重（取值范围为 0～10，且不可以全为 0）；Ds 、Js、Ls 分别为时延 CQI 值、时延抖动 CQI 值和丢包率 CQI 值 。

(4) 综合质量近似 CQI 值为不大于综合质量 CQI 值的最大的 5 的倍数。例如，设备间接访问 SaaS 应用的路径综合质量 CQI 值为 82.5，则综合质量近似 CQI 值为 80。

- 根据综合质量近似 CQI 值优选 SaaS 路径

如下图 1-46 所示，设备根据综合质量近似 CQI 值，将 SaaS 路径加入到对应的色区：

图1-46 根据综合质量近似 CQI 值优选 SaaS 路径

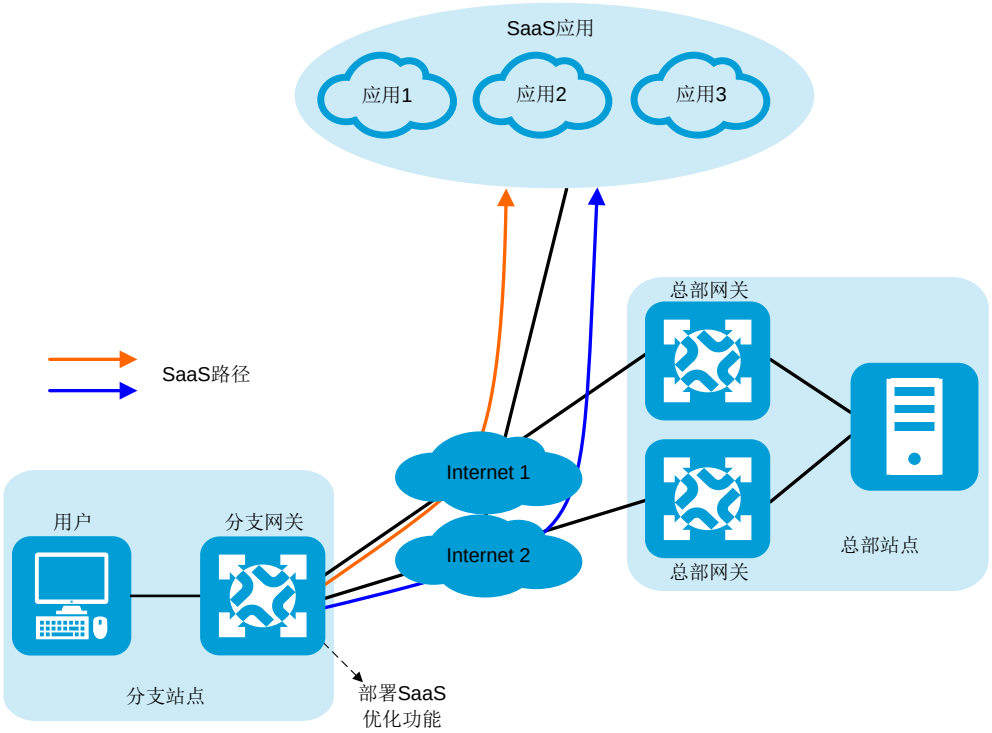
设备为访问 SaaS 应用的流量选择最优路径时，首先按照绿、黄、红的顺序选择色区，然后在当前色区中选择最优路径。如果当前色区中没有任何 SaaS 路径，则继续选择下一个色区；如果色区中只有一条路径，则选择该路径作为最优路径；如果色区中有多条路径，则随机选择一条路径作为最优路径。

直接访问方式

- SaaS 路径优化部署位置

如下图 1-47 所示，采用直接访问方式时，通过在分支网关设备上部署 SaaS 路径优化功能，可以实现对分支网关设备与 SaaS 应用之间的路径进行优选。

图1-47 SaaS 路径优化部署位置（直接访问方式）



- SaaS 路径质量探测和评估

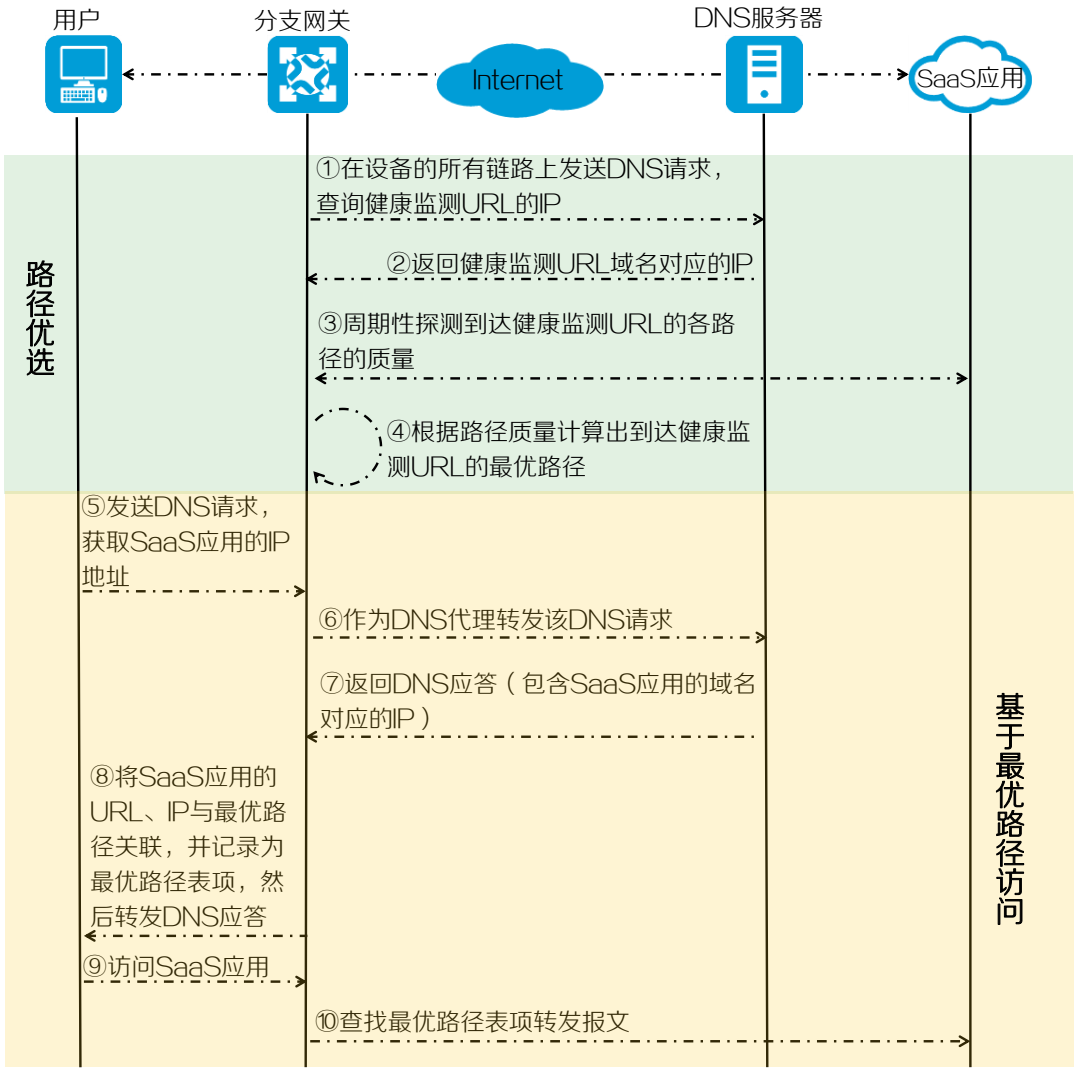
  在直接访问方式中，SaaS 路径是指分支网关与 SaaS 应用之间的路径。

  - 质量探测：分支网关在所有 SaaS 路径上向健康监测 URL 发送 Ping 报文，探测每条 SaaS 路径的质量。
  - 质量评估：每条 SaaS 路径质量 = 分支网关设备与健康监测 URL 之间每条路径的综合质量近似 CQI 值。完成质量评估后，分支网关根据各条 SaaS 路径的综合质量近似 CQI 值，选择最优路径

- SaaS 路径路径优选过程

  SaaS 路径路径优选过程如下图 1-48 所示。

图1-48 SaaS 路径路径优选过程（直接访问方式）

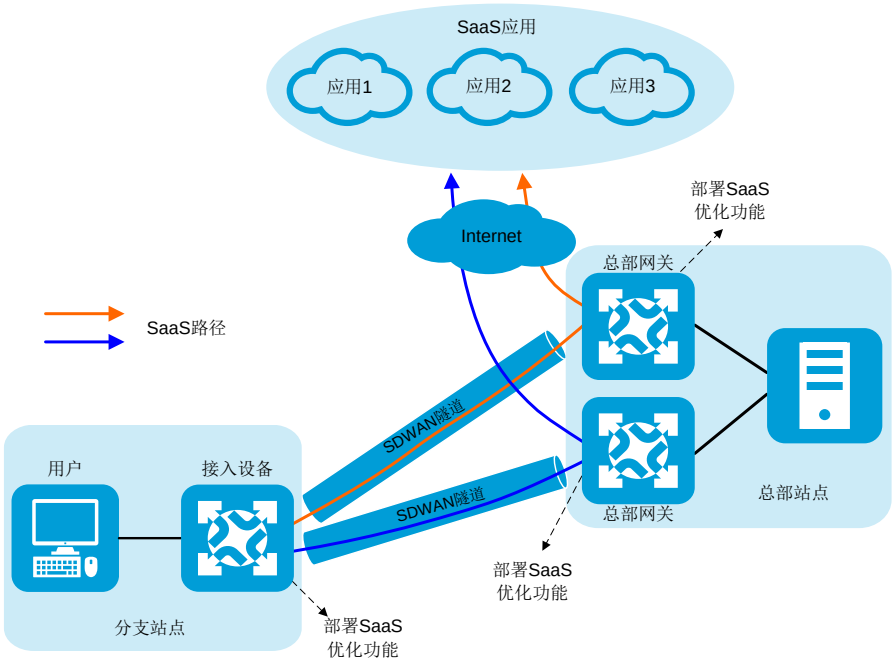


间接访问方式

- SaaS 路径优化部署位置

  如下图 1-49 所示，采用间接访问方式时，需要在本地接入设备与总部网关设备上同时部署 SaaS 路径优化功能，对本地接入设备到 SaaS 应用的整条路径进行优选。

图1-49 SaaS 路径优化部署位置（间接访问方式）



- SaaS 路径质量探测和评估

  在间接访问方式中，SaaS 路径由本地接入设备与总部网关之间的 SDWAN 隧道、总部网关与健康监测 URL 之间的路径两部分组成。

  - 质量探测

    a. 本地接入设备通过 iNQA 探测每条 SDWAN 隧道的质量。

    b. 总部网关设备通过在所有路径上 Ping 健康监测 URL，探测总部网关与健康监测 URL 之间的所有路径的质量。

    c. 总部网关设备通过 BGP EVPN 路由将探测到的路径质量通告给本地接入设备。

  - 质量评估

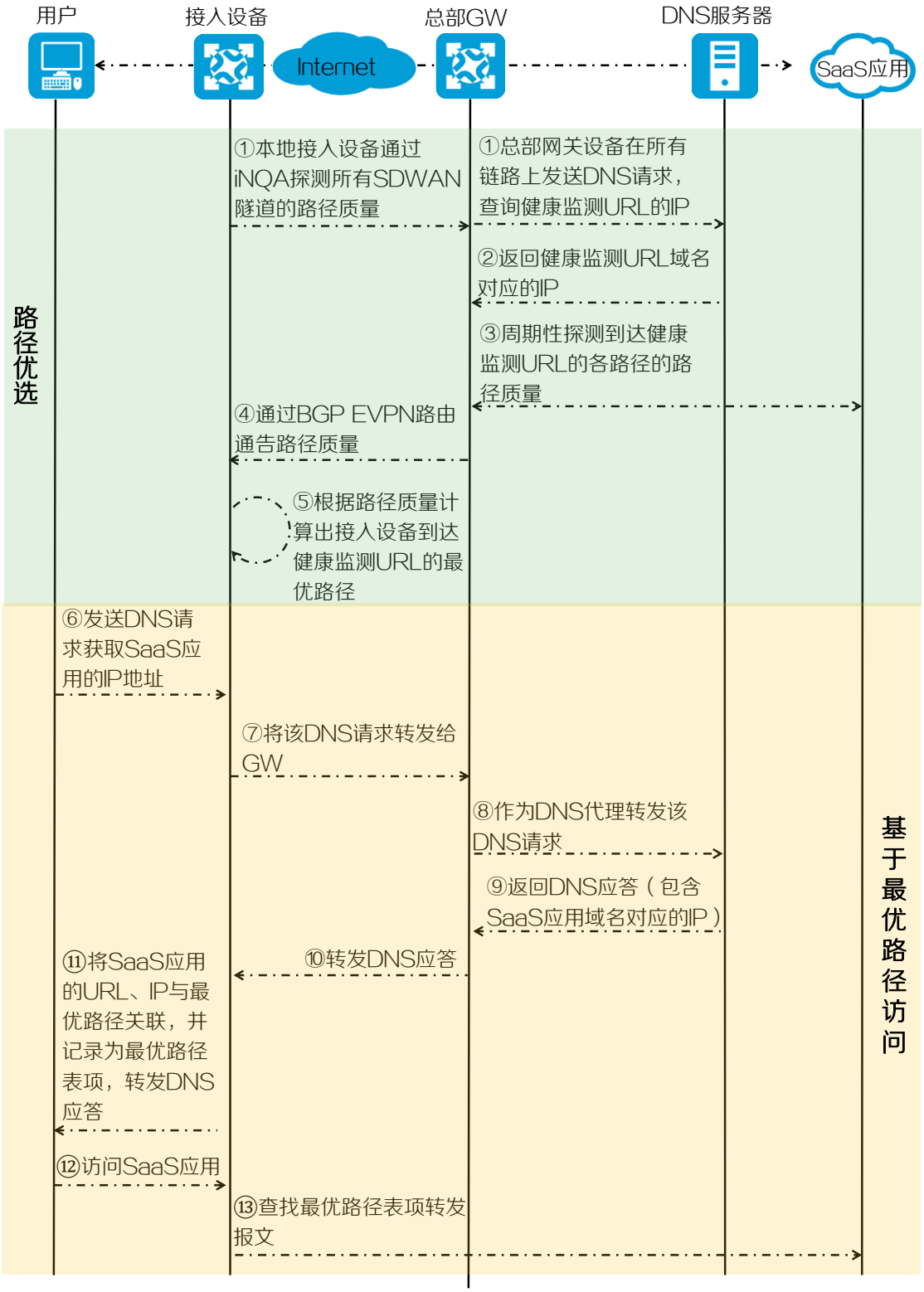
    本地接入设备根据如下算法计算 SaaS 路径的质量，完成质量评估后，本地接入设备根据各条 SaaS 路径的综合质量近似 CQI 值，选择最优路径。

每条 SaaS 路径质量 = 本地接入设备访问健康监测 URL 的每条路径的综合质量近似 CQI 值 =（本地接入设备与总部网关之间的 SDWAN 隧道的综合质量近似 CQI 值+总部网关与健康监测 URL 之间路径的综合质量近似 CQI 值）/2

- SaaS 路径优选过程

  SaaS 路径路径优选过程如下图 1-50 所示。

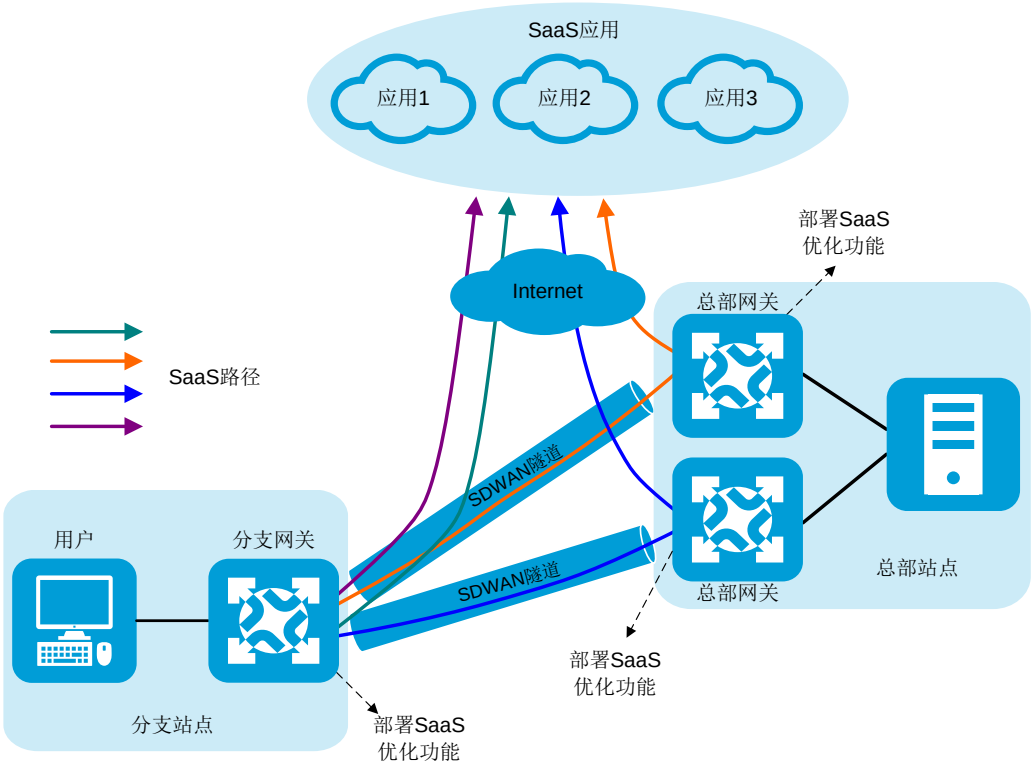
图1-50 SaaS 路径路径优选过程（间接访问方式）



混合访问方式

如下图 1-51 所示，采用直接/间接混合方式时，需要在分支网关与总部网关设备上同时部署 SaaS 路径优化功能。对于每条直接访问路径和间接访问路径，分别进行质量探测，评估路径的质量，从中选择一条最优路径。

图1-51 直接/间接混合方式部署 SaaS 路径优化功能



TCP 拥塞控制

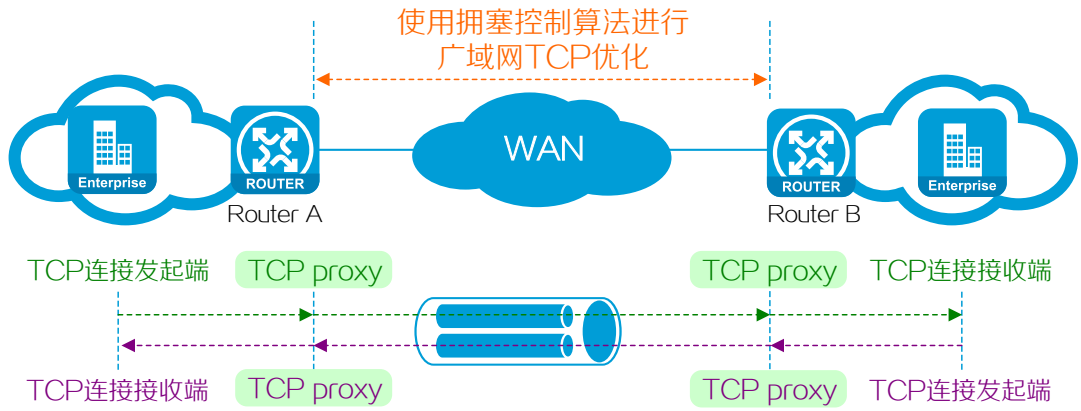
TCP 拥塞控制技术简介

TCP 协议的确认重传机制可为应用提供可靠性传输。但如果网络拥塞，尤其在丢包率和传输时延双高的广域网，确认等待和报文重传会大大降低网络吞吐量和传输效率，加剧网络拥塞，甚至导致传输中断。

如下图 1-52 所示，在广域网接入设备上部署拥塞控制算法，接入设备上会运行 TCP proxy。对于匹配条件的 TCP 连接，发起侧 TCP proxy 和接收侧 TCP proxy 间会建立代理 TCP 连接，并用代理 TCP 连接传输原始 TCP 报文。拥塞控制算法对代理 TCP 连接进行优化，从而实现：不管 TCP 会话的发起端是否支持拥塞控制算法，TCP 代理均可确保原始 TCP 报文在穿越广域网期间，能尽可能利用广域网带宽，又不会造成网络拥塞，快速到达接收端。

Reno、BIC、BBR 为 H3C 设备支持的、三种主流 TCP 拥塞控制算法。

图1-52 使用拥塞控制算法进行广域网 TCP 优化



TCP 拥塞控制技术指标

计算机网络环境复杂，瞬息万变， TCP 拥塞控制算法也在不断地演进，来寻求带宽、丢包、时延的更优解决方案。衡量拥塞控制算法的优劣，有两个重要指标：

- 带宽利用率：即当网络空闲时，算法能否尽快找到最佳 CWND，最大化的利用网络带宽，提升传输速率，提高通信效率。
- 公平性：即算法能否在各 TCP 连接间公平地分配带宽。例如，当新会话加入、争抢网络带宽时，算法应能使现有的各 TCP 连接让出部分带宽，让新连接可以和旧连接公平地分享带宽。

带宽利用率高、公平性好的算法更优。

本文仅从理论上阐述了 Reno、BIC 和 BBR 算法的基本原理机制，及优缺点，来帮助用户选择合适的拥塞控制算法。实际应用中，不同厂商会对算法的参数取值进行调整，请以实际情况为准。

Reno 和 BIC 算法原理

Reno 和 BIC 算法基于丢包反馈，让发起端被动调整 CWND（Congestion Window，拥塞窗口）的大小来控制本端可发送的 MSS（Maximum Segment Size，最大报文段长度）的数量，从而提高带宽利用率，避免网络拥塞。这两种算法均包含慢启动、快速重传、快速恢复和拥塞避免四个阶段。

(1) 慢启动阶段

TCP 连接建立后，算法启动并进入慢启动阶段。在慢启动阶段：

- TCP 进程以较低速率开始发包，然后每经过一个 RTT（Round-Trip Time，往返时延），CWND 加倍，以便快速探测到最佳 CWND。

- 同时，算法设置了慢开始门限 ssthresh（初始值为 256 个 MSS）。ssthresh 用于防止 CWND 无限制过快增长引起网络拥塞，如果 CWND≥ ssthresh，网络仍无丢包，则直接进入拥塞避免阶段，开始慢增长 CWND。

(2) 快速重传阶段

随着 CWND 不断增加，势必会引起网络拥塞导致丢包。如果发送端连续收到三个重复 ACK 报文，TCP 进程会认为当前网络拥塞导致丢包。算法进入快速重传阶段，并立即重传对方尚未收到的报文，而不用等到 ACK 报文超时再重传。

(3) 快速恢复阶段

快速重传后，算法进入快速恢复阶段。在快速恢复阶段：

- 算法将 CWND 减小到丢包时 CWND 的一半，来快速解除网络拥塞，恢复网络畅通。
- 同时，算法将 ssthresh 设置为丢包时 CWND 的一半，以便进入拥塞避免阶段。

(4) 拥塞避免阶段

在拥塞避免阶段，ssthresh 保持不变， CWND≥ ssthresh，若网络无丢包，则 CWND 缓慢增加，用于探测链路中是否存在富余带宽；若网络出现丢包，Reno 和 BIC 均会再次进入快速重传、快速恢复和拥塞避免阶段。

Reno 和 BIC 算法最根本的差异在于：拥塞避免阶段探测最佳 CWND 时，CWND 的增长幅度不同。

- Reno 采用线性加一法，收敛慢。
  - 无丢包情况下， Reno 每经过一个 RTT，CWND 加 1，来线性探测更大的 CWND。
  - 若不断增加 CWND 值后，引起了网络丢包，此时会再次进入快速重传、快速恢复和拥塞避免阶段。
- BIC 采用二分搜索法和乘性增加法，收敛快。
  - BIC 先采用二分搜索法，探测当前最佳带宽。BIC 每经过一个 RTT，CWDN 的增加值依次减半（CWND 的增加值依次为：丢包折半后的 CWND 值/2、丢包折半后的 CWND 值/4、丢包折半后的 CWND 值/8），直到出现如下两种情况中的一种时二份搜索法结束：

    a. 采用二份搜索法不断增加 CWND 值后，引起了网络丢包，此时会再次进入快速重传、快速恢复和拥塞避免阶段。

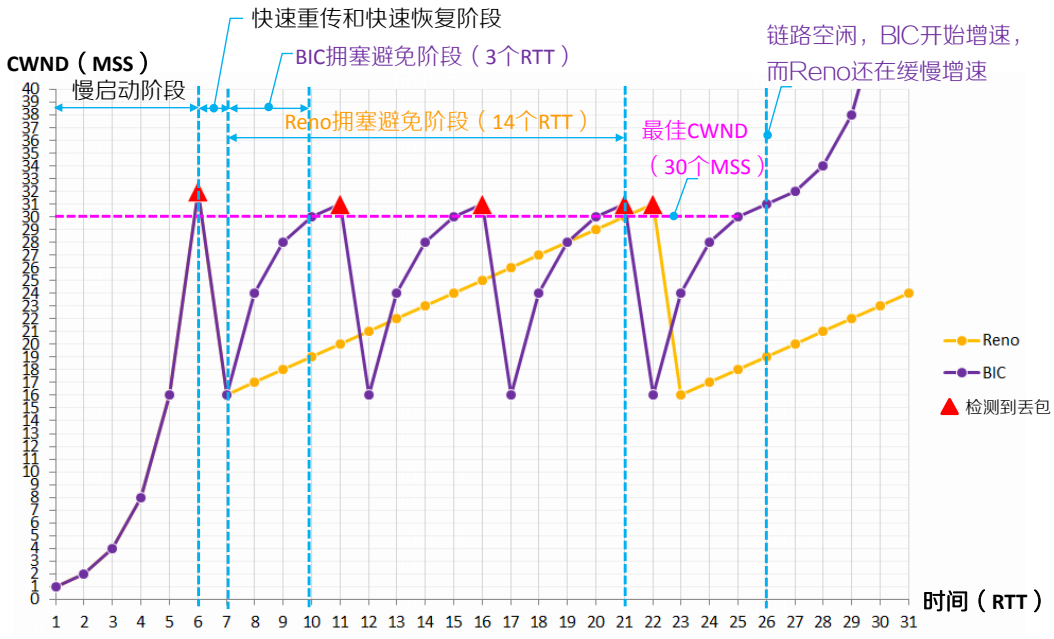
    b. CWDN 的增加值多次折半取整后（只保留整数位）等于 1，网络仍未丢包，则说明网络带宽存在空闲，需要通过更高效的方式继续探测当前最佳带宽。
  - 若通过二分搜索法还未找到当前最佳带宽，则需要进一步采用 CWND 的增加值加倍的方式来快速探测出当前最佳带宽。CWND 的增加值依次为：1、2、 4、 8 、 …… ，直到出现丢包或者达到算法设定的增长门限。

如下图所示：

- Reno 的 CWND 从 16 增长到 30，需要 14 个 RTT 。

- BIC 的 CWND 从 16 增长到 30，需要 3 个 RTT（比 Reno 节省 11 个 RTT），3 个探测点对应的 CWND 为 24（16+8）、28（16+8+4）、30（16+8+4+2）。
- 当链路空闲时，BIC 能迅速利用带宽，抢占能力明显高于 Reno。

图1-53 拥塞避免阶段



Reno 和 BIC 算法总结

当网络带宽越大，或者 RTT 越大，BIC 的优势比 Reno 越明显。所以，BIC 比 Reno 更适合计算机网络，BIC 一经推出，即替代 Reno 作为 Linux 的缺省拥塞控制算法。基于 Reno 和 BIC 的原理，我们可以发现 Reno 和 BIC 均存在以下使用约束：

- 算法会将传输过程中产生的错包、随机丢包误认为是拥塞丢包，从而大幅减小 CWND，降低发送速率，浪费了网络带宽。
- 为最大化利用带宽，只要未丢包，Reno 和 BIC 就会持续增大 CWND，并尽力发送报文，这一方面会导致报文突发，增加丢包的风险；另一方面会导致来不及处理的报文堆积在网络设备的接收缓冲区，使得传输时延增大。

BBR 算法原理

BBR（Bottleneck Bandwidth and RTT，瓶颈带宽和往返时延）算法基于测量反馈来主动调整 CWND 和报文发送速率，从而提高带宽利用率，避免网络拥塞。

- BBR 的原理
  - 基于 ACK 反馈持续测量 RTT，并计算网络即时带宽 BW。采用将当前获得的 BW、RTT 和历史瓶颈带宽 BtlBw（初始值为 0）、历史 RTT 极小值 RTprop（初始值比较大，为十

六进制数全 f）比较，并根据 Kathleen Nichols 算法，得到当前瓶颈带宽 BtlBw 和当前 RTT 极小值 RTprop。

- BBR 使用得到的 BtlBw 和 RTprop，乘以 pacing_gain（发送速率增益系数）、cwnd_gain（发送窗口增益系数），计算出下次发送的 CWND 和发包速率。

- BBR 的实现过程

  如下表 1-8 所示，BBR 通过以下四个状态，来交替循环探测，得出 BtlBw 和 RTprop：StartUp（启动阶段）、Drain（排空阶段）、ProbeBW（BtlBw 探测阶段）、ProbeRTT（RTprop 探测阶段）。

为了获得 BtlBw，BBR 会增加发包数量，可能会造成报文被缓存在链路的队列中，导致 RTT 增加；而为了测量 RTprop（ RTprop 指链路中没有排队且无丢包时，一个包在链路中一个来回所需的时长），需要清空当前链路队列中的缓存。因此，不能同时探测得出 BtlBw 和 RTprop，需要交替循环探测得出 BtlBw 和 RTprop 。

表1-8 BBR 实现过程

BBR 状态	实现过程
StartUP	• 为了快速探测到最佳发包速率，BBR 需要探测当前即时带宽 BW、并与历史瓶颈带宽 BtlBw 进行比较，最大值为当前瓶颈带宽 BtlBw • 通过当前瓶颈带宽 BtlBw 和发包速率增益系数（此时的发包增益系数 pacing_gain=2.89），计算出当前发包速率（当前发包速率=pacing_gain*BtlBw = 2.89*BtlBw） • 如果 BBR 连续 3 次探测、计算得到的即时带宽 BW 均小于当前瓶颈带宽 BtlBw 的 1.25 倍，则说明已经获得了当前瓶颈带宽 BtlBw，且链路可能出现了发包缓存现象，BBR 进入 Drain 状态
Drain	• 因为在 StartUP 状态链路中可能出现了发包缓存现象，为了排空链路队列中缓存的报文，BBR 需要降低发包速度。此时的发包增益系数 pacing_gain=0.35，发包速率 =pacing_gain*BtlBw = 0.35*BtlBw • BBR 降低发包速率，直至 BBR 检测到链路中正在传输的报文数小于 “BtlBw*RTprop”。此时 BBR 认为链路队列中缓存的报文已经排空，进入 ProbeBW 状态
ProbeBW	BBR绝大部分时间处于ProbeBW状态。该阶段以8个RTT为周期循环执行，包括6个平稳周期、1个上升周期、1个减少周期： • 6 个平稳周期：为了充分利用网络带宽，BBR 在 6 个 RTT 内进行平稳发包。此时的发包增益系数 pacing_gain=1，发包速率=pacing_gain*BtlBw = 1*BtlBw • 1 个上升周期：为了探测更高的 BtlBw，BBR 在 1 个 RTT 内短暂地提高发包速率。此时的发包增益系数 pacing_gain=1.25，发包速率=pacing_gain*BtlBw = 1.25*BtlBw • 1 个减少周期：排空网络设备接收队列中可能缓存的报文, BBR 在 1 个 RTT 内短暂地降低发包速率。此时的发包增益系数 pacing_gain=0.75，发包速率=pacing_gain*BtlBw = 0.75*BtlBw

BBR 状态	实现过程
ProbeRTT	为了确保RTprop值的可靠性，BBR会检测RTprop值的刷新状态，若超过10秒RTprop还未更新，则无论当前BBR处在StartUP、Drain或ProbeBW中的哪个状态，均进入RTprop状态 • 进入 RTprop 状态后，BBR 将 CWND 减小到 4 并持续 200 毫秒，通过减少发包让链路中的报文排空来探测新的 RTprop • RTprop 探测完成后，BBR 根据最新数据来确定进入 StartUp 还是 ProbeBW 状态

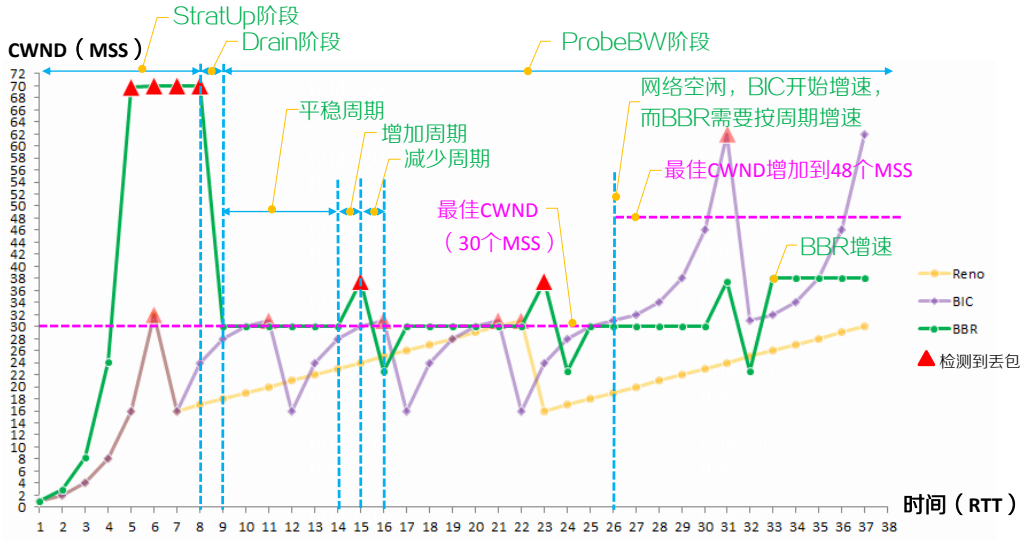
BBR 算法总结

如下图 1-54 所示，相比 Reno 和 BIC，BBR 的优势主要体现在以下方面：

- BBR 基于周期探测结果来调整 CWND，不以丢包为触发拥塞控制的条件，比 Reno 和 BIC 更合理，更高效。
- Reno 和 BIC 使用 CWND 控制发包，它的尽力发送势必导致链路中报文被队列缓存，报文在网络中的传输时延增加。队列缓存越大，时延越严重。BBR 使用发包速率和 CWND 双重参数控制发包，能将突发报文平滑发送，而且 BBR 会定期降速发包，清空链路中队列里缓存的报文，能够有效的降低网络时延。
- BBR 绝大部分时间平稳发包，而 Reno 和 BIC 需要循环探测、丢包、快速恢复、探测，BBR 带宽利用率更高，吞吐量更大。

当然，BBR 也有使用约束：由于 BBR 的周期性特点，以及 ProbeBW 状态增加周期增速限额（1.25* BtlBw）和减少周期对减速限额（0.75* BtlBw），导致当网络带宽增加或者减少时，BBR 的抢占能力不如 BIC。BIC 反应迅速，如果链路中大部分为持续时间短的 TCP 连接，则使用 BIC 的传输效率更高。

图1-54 BBR 算法总结



标准 BBR 的实现称为 BBRv1，为了弥补 BBRv1 在多种算法（BBR、BIC 算法）共存的网络中带宽抢占能力方面的使用约束，BBR 又衍生出了 BBRv2。BBRv2 是对 BBRv1 的优化，优化的方面包括但不仅限于：

- BBRv2 在 Startup 状态，增加了丢包退出判断条件。（为避免更严重地丢包，在 Startup 状态，只要检测到丢包，则退出 Startup 状态，而不用等到“连续 3 次探测计算得到的 BW 均小于当前 BtlBw 的 1.25 倍”）
- 为了快速响应网络变化，TCP 需要及时调节 CWND，通过加快 ProbeRTT 的更新可实现 CWND 值的快速调节。BBRv2 将进入 RTprop 状态的触发条件从 10 秒减小到 2.5 秒（若持续 2.5 秒 RTprop 一直未更新，则 BBRv2 立即进入 ProbeRTT 状态，来探测新的 RTprop）
- 为了减小进入 ProbeRTT 状态带来的带宽波动，BBRv2 进入 ProbeRTT 状态后，BBRv2 将 CWND 减少到 0.5*BtlBw*RTprop（BBRv1 将 CWND 减少到 4）

TCP 拥塞控制技术比较

Reno、BIC 通过 CWND 来控制发送端可发送的报文的数量，BBR 通过网络最大带宽 BtlBw 和 RTprop（极小 RTT）控制发送端发送报文的速度和数量。它们最大的差别在于链路最大带宽的探测机制不同。总体趋势来说：

- 收敛速度（从快到慢）：BBR>BIC>Reno
- 公平性（深缓存场景下的抢占能力，从强到弱）：BIC>BBR>Reno
- 抗丢包能力（从强到弱）：BBR>BIC>Reno
- 网络排队时延 （从低到高）：BBR<Reno<BIC

深缓存指链路中具有比较大的缓存，例如接收缓存、中间设备的缓存队列等。

如下表 1-9 所示，我们再来回顾一下各算法的关键技术，并根据其特点，给出推荐使用场景。

表1-9 TCP 拥塞控制算法比较

算法	关键技术	使用场景
Reno	• 将丢包作为拥塞控制的触发条件，不区分丢包原因是错包还是真正的网络拥塞，容易误判 • 检测到丢包时，CWND 减半 • 线性增长探测 CWND，每经过一个 RTT，CWND 加 1	推荐用于低丢包率、低带宽、对时延要求低的场景
BIC	• 将丢包作为拥塞控制的触发条件，不区分丢包原因是错包还是真正的网络拥塞，容易误判 • 检测到丢包时，CWND 减半 • 用二分搜索法和乘性增加法探测 CWND	推荐用于低丢包率、高带宽、对时延要求低的场景
BBRv1	不将丢包作为拥塞控制的触发条件，以周期探测的时延和带宽为依据，计算CWND和发送速率，平稳发包	推荐用于存在一定丢包率、高带宽、对时延要求高、BBR算法单独使用的场景
BBRv2	相比于BBRv1: • BBRv2 在 Startup 状态，只要检测到丢包，则退出 Startup 状态，来避免更严重地丢包 • BBRv2 将进入 RTprop 状态的触发条件从 10 秒减小到 2.5 秒（若持续 2.5 秒 RTprop 一直未更新，则 BBRv2 立即进入 ProbeRTT 状态，来探测新的 RTprop），来及时调节 CWND，以便快速响应网络变化 • BBRv2 进入 ProbeRTT 状态后，BBRv2 将 CWND 减少到 0.5*BtlBw*RTprop（BBRv1 将 CWND 减少到 4），来减小进入 ProbeRTT 状态带来的带宽波动	推荐用于存在一定丢包率、高带宽、对时延要求高、各种算法共存的场景

自适应音视频保障

简介

FEC（ Forward Error Correction，前向纠错）是一种广域网差错控制技术，用于解决实时音、视频的网络丢包问题，保障重要音、视频流穿越广域网后仍能流畅播放，增强 RTP 流在广域网传输的可靠性。FEC 的基本原理是：发送端将多个原始报文，通过 RS（Reed Solomon）算法编码计算并增加冗余包后，以编码块为单位发送；接收端根据收到的原始报文和冗余包恢复出丢失的报文，还原原始数据。

平均抗丢包率

FEC 技术中的重点和难点是生成适量的冗余包。生成的冗余包数量过多，会浪费网络带宽；生成的冗余包数量过少，则无法恢复出丢失的报文。冗余包的数量受编码块包数、编码算法和平均抗丢包率的影响。其中，编码块包数指编码块中原始报文的个数，由用户配置决定；编码算法出厂时

已固定，不可配置；平均抗丢包率用于衡量 FEC 抗丢包的能力，通常会设置成大于链路平均丢包率的值。当编码块包数和编码算法已确定，平均抗丢包率对传输网络的影响如下：

- 平均抗丢包率越大，FEC 抗丢包能力越强，但同时会产生更多的冗余包，占用更多的链路带宽，传输效率越低。
- 平均抗丢包率越小，FEC 抗丢包能力越弱，同时产生的冗余包就少，占用的链路带宽越少，传输效率越高。

D-FEC 和 A-FEC

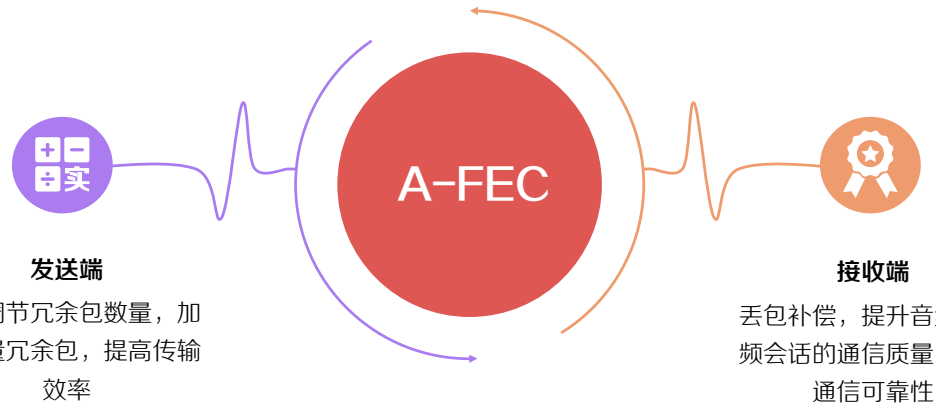
根据平均抗丢包率的配置方式不同，FEC 分为 D-FEC（Determined-FEC，固定的 FEC）和 A-FEC（Adaptive-FEC，自适应的 FEC）两种类型。

- D-FEC 使用固定的平均抗丢包率，适用于链路丢包率比较稳定的场景。
- A-FEC 根据实时链路丢包率动态调节平均抗丢包率，实现机制比 D-FEC 复杂。可在保证通信质量和可靠性的前提下，尽量减小冗余包数量，提高传输效率。在广域网环境下，A-FEC 应用更广泛。

技术价值

A-FEC 技术的价值如下图 1-55 所示。

图1-55 技术价值



处理流程

如下图 1-56 所示，A-FEC 处理流程：

(1) 发送端识别业务报文，进行 FEC 编码处理后发送。

a.报文识别：发送端根据管理员配置的策略匹配需要重点保障的音、视频报文。对于匹配成功的报文，则进入 FEC 处理环节；对于匹配失败的报文，则正常转发。

b.缓存、编码：积攒多个原始报文（例如报文 1、2、3、4），对这些报文进行编码，生成一个或多个冗余包 R。冗余包报文头中携带原始报文和冗余包的个数、原始报文和冗余包的起始序号等信息。

c.组合发送：按顺序发送原始报文 1、2、3、4 和冗余包 R。

(2) 在广域网中传输报文，编码块中的原始报文和冗余包都可能出现丢包。

(3) 接收端识别业务报文，进行 FEC 解码后发送给目的端。

a.报文识别：接收端根据和发送端相同的策略匹配音、视频报文。对于匹配成功的报文，则进入 FEC 解码环节；对于匹配失败的报文，则正常转发。

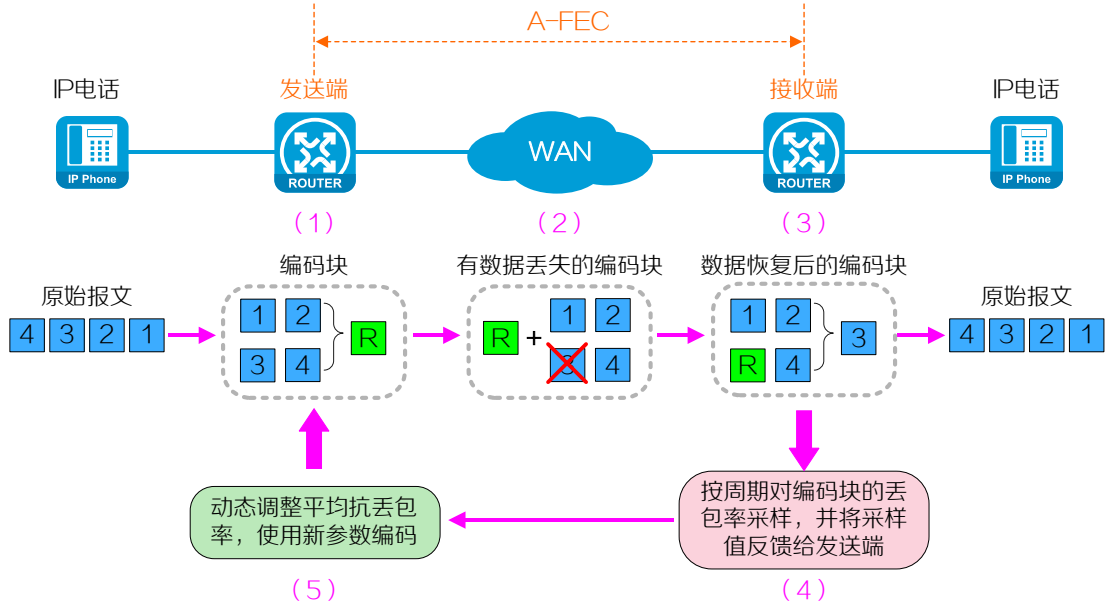
b.缓存、解码：接收端将原始报文缓存，收到冗余包时会根据冗余包中携带的信息尝试解码。如果缓存的同一编码块中的原始报文与冗余包个数之和大于等于冗余包中携带的原始报文的个数，则可恢复丢失的数据，解码成功。如下图所示，接收端根据原始报文 1、2、4 和冗余包 R 可恢复出丢失的原始报文 3。如果解码失败，则继续等待下一个冗余包来解码，或者等待解码超时，再进入发送环节。

c.发送：丢弃冗余包，发送原始报文。

(4) 接收端在执行步骤(3)的同时会计算每个编码块的丢包率，按周期对编码块的丢包率采样，并将采样的丢包率反馈给发送端。

(5) 发送端根据收到的丢包率计算平均值，动态调整平均抗丢包率，并从下一次编码开始，使用调整后的平均抗丢包率进行 FEC 编码。

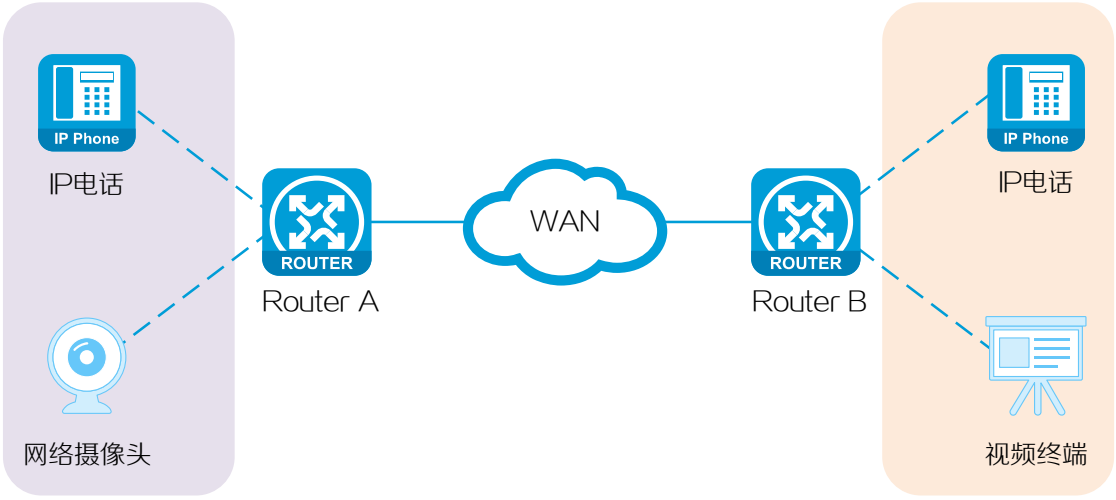
图1-56 A-FEC 处理流程



典型组网

如下图 1-57 所示，对于网络丢包有较高要求的语音、视频等数据流，可在流量接入广域网的设备 Router A 和 Router B 上开启 A-FEC 功能，来保障关键的音、视频通信在 20%丢包的情况下无卡顿、无花屏，能够为音、视频会议的正常举行提供有力保障。

图1-57 A-FEC 典型组网



多路径包复制

简介

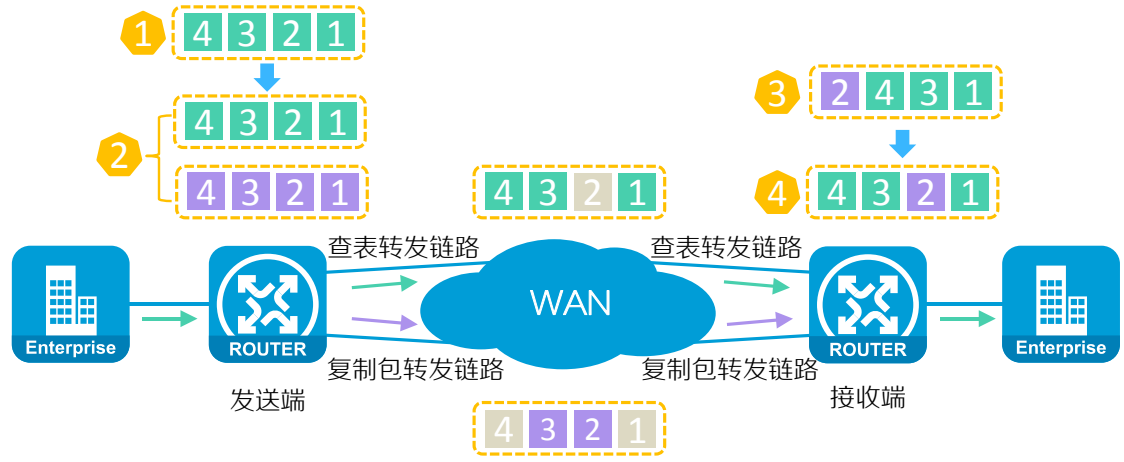
多路径包复制是一种冗余纠错技术。对于可靠性要求很高并且流量较小的业务（例如工业自动控制、付款、紧急呼叫、重要音视频通话等），多路径包复制技术利用大部分局域网普遍采用多条 WAN 链路接入广域网的多链路优势，在发送端将原始业务报文复制后从两条链路发送出去，在接收端将原始业务报文和复制报文互相补充，整合成原始报文流，从而有效降低甚至解决因单条链路丢包导致的业务中断等问题。

工作机制

如下图 1-58 所示，多路径包复制对报文的处理流程包括：

- 在发送端进行 ① 报文识别、 ② 报文复制。
- 在接收端进行 ③ 报文去冗余、④ 报文保序。

图1-58 多路径包复制对报文的处理流程

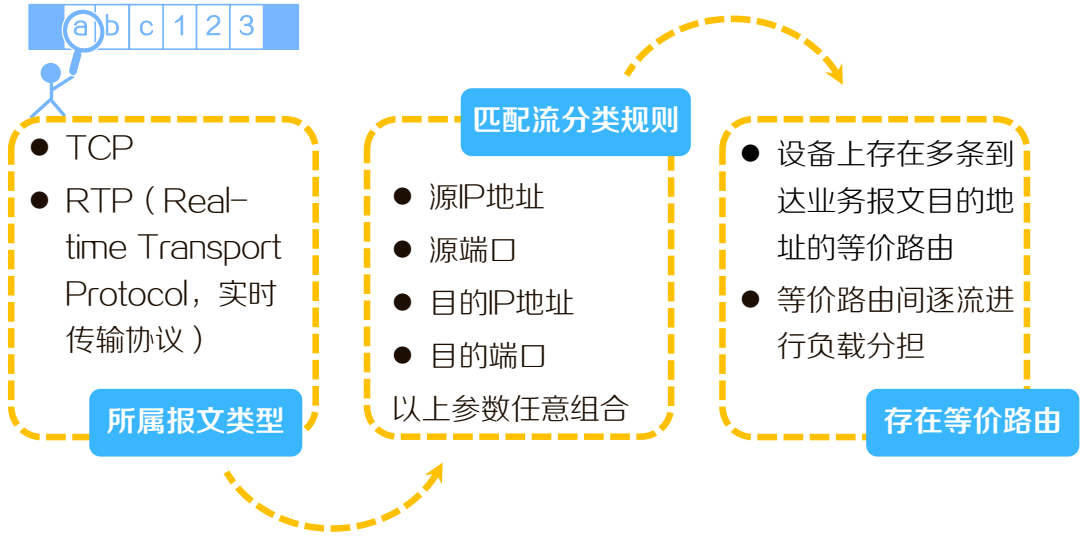


相同序号的复制报文和原始业务报文完全相同。本文为更好的展示原理，在复制报文图标的左上角添加了字母D，以示和原始业务报文的区别。

- 报文识别

发送端收到业务报文后，对报文进行识别。只有类型符合要求、匹配 TCP/UDP 流分类规则，且存在逐流负载分担等价路由的报文，设备才会对其进行包复制处理，如下图 1-59 所示。

图1-59 报文识别处理流程



- 报文复制

发送端将满足条件的多个原始业务报文积攒成一个数据块(本文假设 4 个报文作为一个数据块), 进行复制并转发。

- 原始报文通过查表转发路径（即路由出接口对应的链路）发送。
- 复制报文通过复制包转发路径（即路由表中第一个被查询到的、除查表转发路径外的其它等价链路）发送。

- 报文去冗余

如下图 1-60 所示，接收端会从两条链路接收到两份报文，需要对收到的重复报文进行去冗余处理，以便确保发送给目的端的报文中相同序号的原始业务报文和复制报文中仅保留一个。

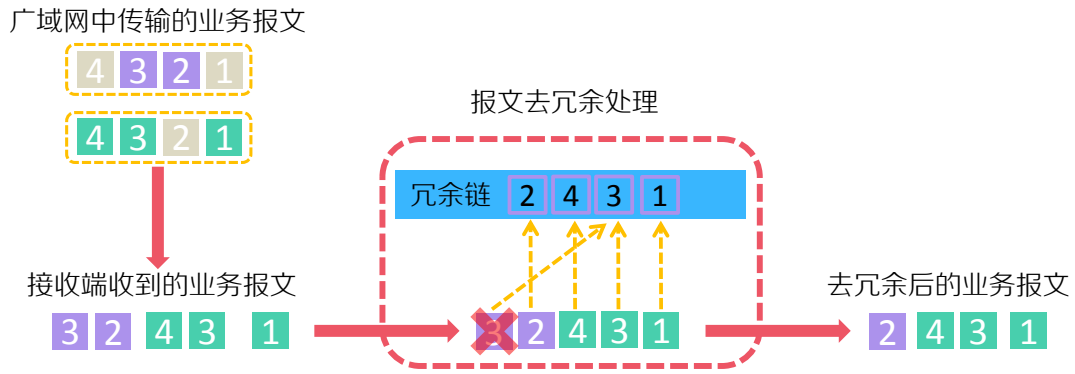
因为查表转发路径和复制包转发路径的速率不同、时延不同，且可能丢包，现假设接收端先后收到原始业务报文 1、3、4 和复制报文 2 和 3，报文去冗余流程大致如下：

(1) 报文去冗余处理模块第一次收到原始业务报文 1、3、4 和复制报文 2，将这些报文的信息记入冗余链，允许报文通过。

(2) 收到复制报文 3，因为冗余链中已经有报文 3 的信息记录，系统判断复制报文 3 是重复报文，于是，丢弃复制报文 3。

(3) 将去冗余后得到的报文 1、3、4、2 依次发送给“报文保序”环节处理。

图1-60 报文去冗余处理流程



冗余链是记录报文信息的缓冲区，用于判断报文是否为冗余报文。对于非冗余报文，允许通过；对于冗余报文，直接丢弃。

- 报文保序

因为两条链路的时延不同，报文到达接收端的先后顺序很可能和报文本身的序号顺序不一致，所以，接收端对报文进行去冗余处理之后，还需要进行保序处理，以尽量保证报文按照序号从小到大的顺序发送出去。

假设业务报文去冗余后顺序依次为 1、3、4、2，则报文保序流程大致如下图 1-61 所示：

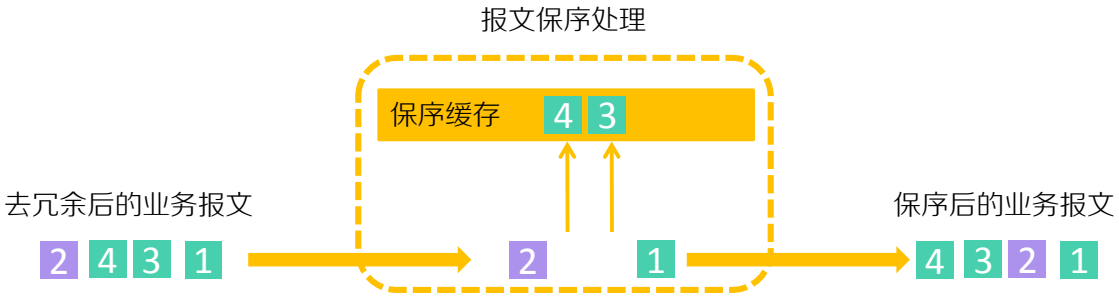
(1) 报文保序处理模块收到报文 1 后，直接发送，并将下次期望收到的报文序号（以下简称为期望序号）设置为当前收到的报文序号加 1，即期望序号为 2。

(2) 收到报文 3 和 4，报文序号均大于期望序号 2，则将报文 3 和 4 放入保序缓存，期望序号仍为 2。

(3) 收到复制报文 2，报文序号等于期望序号 2，则发送复制报文 2，更新期望序号为 3，从保序缓存中取出并发送序号为 3 的报文。同时更新期望序号为 4，从保序缓存中取出并发送序号为 4 的报文，更新期望序号为 5。以此类推。

(4) 将保序后得到的报文 1、2、3、4 依次发送给目的端。

图1-61 报文保序处理流程

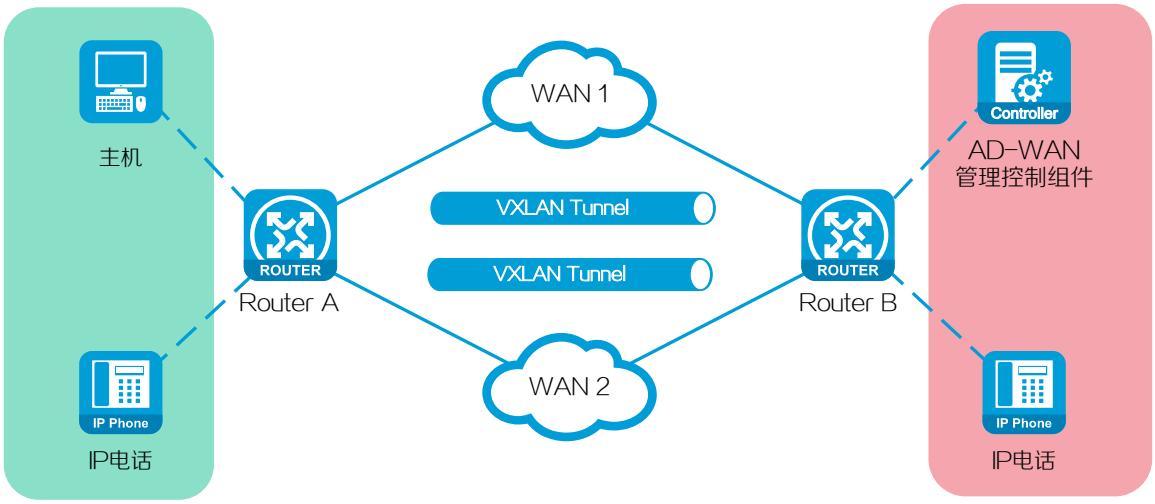


保序缓存是存放报文的缓冲区，用于尽量保证按照报文序号大小顺序将报文发送给下一跳。如果收到的报文的序号大于期望序号，即期望报文延迟到达，接口会将当前报文暂时存入保序缓存，以便等待期望的报文发送后，再发送这些报文。

典型应用

如下图 1-62 所示，某企业的分支机构通过双 WAN 链路跨 VXLAN 网络和总部连接，且存在一些可靠性要求较高的业务（例如通过网络向 AD-WAN 管理控制组件下发指令、重要音视频通话等）。可利用双 WAN 链路组网特点，在 Router A 和 Router B 上部署多路径包复制功能，使得这些业务报文在 WAN 链路传输时，能有效降低甚至解决单条链路中的丢包、乱序问题，保证业务的高可靠性。

图1-62 多路径复制应用



Web 应用缓存

简介

Web Cache（Web 缓存）是一种 Web 应用缓存功能，可将用户访问过的 Web 页面内容缓存在本地用于直接响应后续相同请求。

工作流程

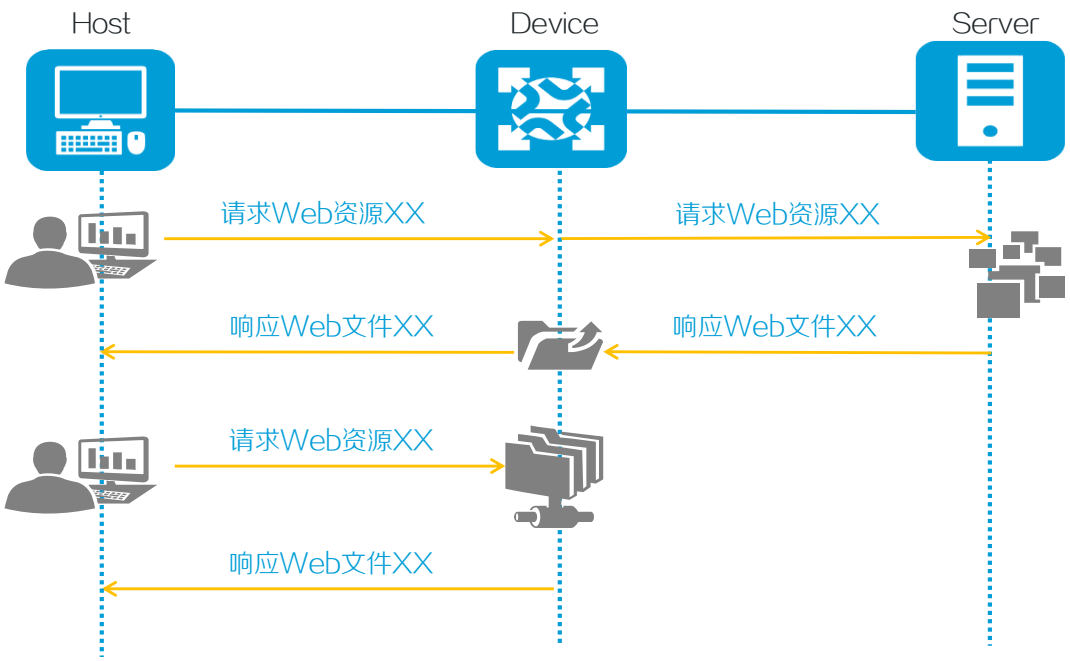
如下图 1-63 所示是 Web Cache 的工作流程：

(1) 用户通过 HTTP/HTTPS 协议请求指定服务器的 Web 页面内容时， Web Cache 将服务器响应的内容缓存在本地文件中。

(2) 后续用户访问该服务器时，Web Cache 获取报文的 URL，依据 URL 在本地缓存文件中进行查找。

(3) Web Cache 使用查找到的本地缓存文件中的内容直接响应用户。

图1-63 Web Cache 工作流程



技术价值

Web Cache 功能的技术价值如下图 1-64 所示。

图1-64 Web Cache 技术价值

提高响应速度

提高了用户访问网站的速度及响应时间，增强用户体验

减少网络带宽消耗

减少设备到服务器的访问流量，降低设备传输成本

缓解服务器压力

减少了目的服务器的访问流量，缓解目的端服务器压力

相关说明

本功能支持缓存基于 HTTP/HTTPS 协议的 Web 应答文件。

智能网络质量分析

简介

iNQA（Intelligent Network Quality Analyzer，智能网络质量分析）是一种适用于大规模 IP 网络、可快速测量网络丢包性能的机制，可测量单向和双向丢包信息（包括报文丢失数、报文丢失率、字节丢失数、字节丢失率）。网络管理员利用 iNQA 的测量结果可快速定位丢包时间、丢包位置、丢包严重程度。

技术优势

如下图 1-65 所示是 iNQA 技术的优势。

图1-65 技术优势

测量范围广

iNQA支持对二层和三层网络中的丢包信息进行测量

定位速度快

iNQA按周期自动测量目标流转发路径上各段的丢包信息。与发现丢包后再逐点检测、逐步排查故障点的测量方法相比，iNQA定位故障的速度更快

支持场景多样

支持点到点、点到多点、多点到多点等多种场景

测量精度高

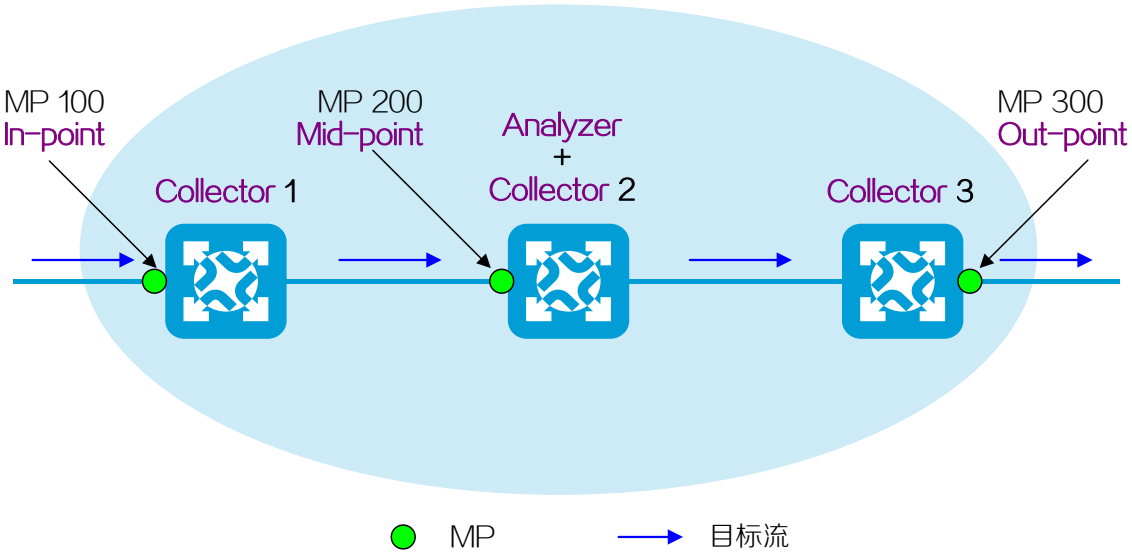
相比NQA等通过模拟业务报文进行测量的技术，iNQA通过标识业务报文，直接对业务报文进行丢包测量，测量数据可以真实反映网络质量，丢包计算更精准

网络模型

iNQA 使用多点（多个 Collector）收集、单点（单个 Analyzer）计算的网络模型，其模型中包含以下元素，如下图 1-66 所示：

- Collector（采集器）：负责管理和控制 MP，周期性收集 MP 产生的统计数据并上报给 Analyzer。
- Analyzer（AD-WAN 智能分析组件）：负责收集 Collector 上送的统计数据并完成数据的汇总和计算。Analyzer 可以独立部署在一台设备上，也可以和 Collector 部署在同一台设备上。
- 目标流：iNQA 统计的目标对象，是网络中符合指定匹配规则的业务报文流。可以通过源 IP 地址/网段、目的 IP 地址/网段、协议类型、源端口号、目的端口号参数的任意组合来定义一条目标流。
- MP（测量点）：负责执行测量动作和产生测量数据，是目标流的实际测量点。MP 和 Collector 上的接口绑定，完成对接口收发报文丢包情况的测量。根据职责不同，MP 分为 In-point（流量入口测量点）、Out-point（流量出口测量点）和 Mid-point（中间测量点）三种类型。

图1-66 网络模型



应用场景

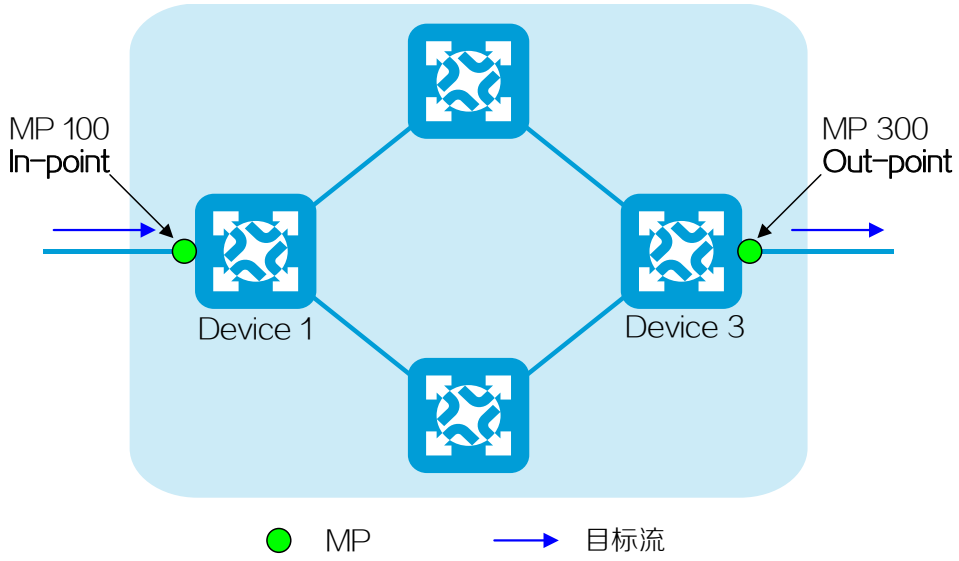
(1) 端到端丢包统计

端到端丢包统计用于测量流量从 In-point MP 进入网络、Out-point MP 离开网络过程中的丢包信息。

- 点到点网络丢包统计

如下图 1-67 所示，该场景下，In-point MP 和 Out-point MP 均只有一个，且均在同一个网络（如均在广域网）内。例如，目标流从 MP 100 进入网络，从 MP 300 离开网络，通过在 Device 1 和 Device 3 上部署 iNQA，可以测量报文穿越该网络时的丢包情况。

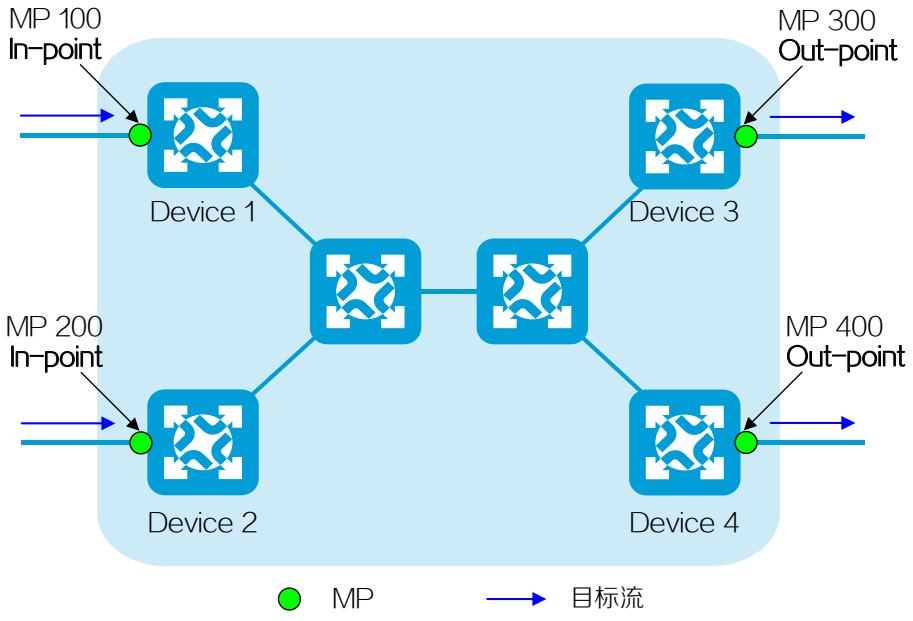
图1-67 点到点网络丢包



- 多点到多点网络丢包统计

  如下图 1-68 所示，该场景下，In-point MP 和 Out-point MP 均有多个，且均在同一个网络。例如，目标流从 MP 100 和 MP 200 进入网络，从 MP 300 和 MP 400 离开网络。通过在 Device 1、Device 2、Device 3 和 Device 4 上分别部署 iNQA，可以测量报文多路径穿越该网络时的丢包情况。

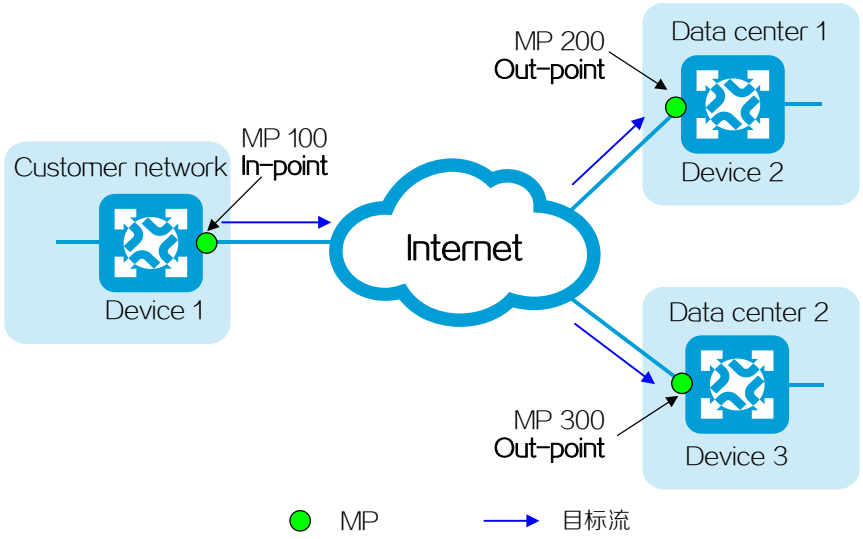
图1-68 多点到多点网络丢包



- 跨网络丢包统计

  如下图 1-69 所示，该场景下， In-point MP 和 Out-point MP 不在同一个网络内，中间跨越了其它网络。例如，用户通过 Internet 访问数据中心，业务流量在互为备份的数据中心 1 和数据中心 2 之间进行负载分担。在用户网络和数据中心网络边缘设备上部署 iNQA，可以统计流量穿越 Internet 时是否存在丢包。

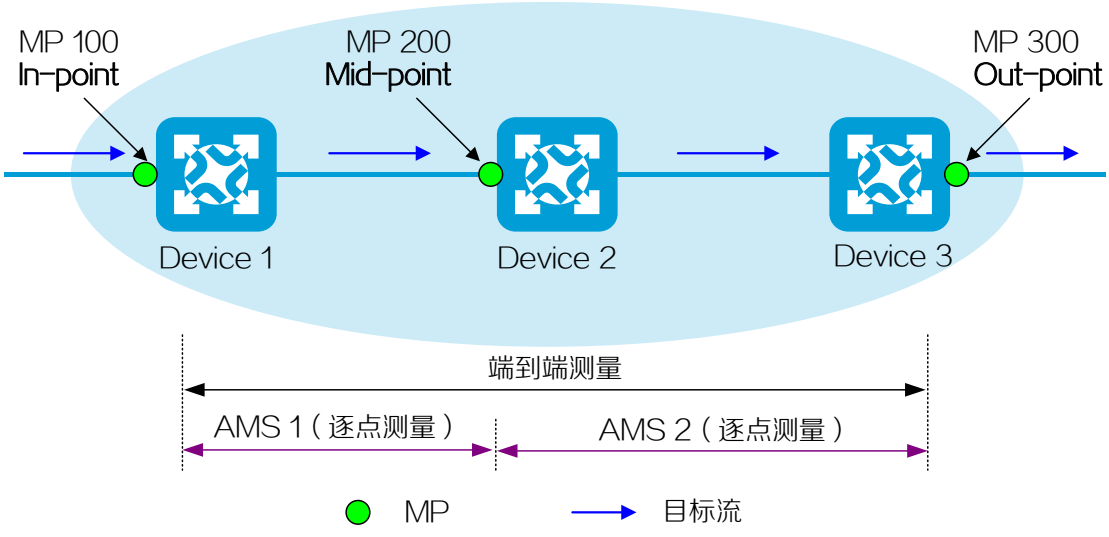
图1-69 跨网络丢包



(2) 逐点丢包统计

如果 In-point MP 和 Out-point MP 之间存在多台设备、多段线路，可以使用 Mid-point MP 将端到端测量网络划分为多个更小的测量单元 AMS（Atomic Measurement Span，原子测量段）。使用 AMS 可以测量目标流传输路径上任意两个物理接口间是否存在丢包，协助进一步定位丢包位置，这就是逐点丢包统计，如下图 1-70 所示。

图1-70 逐点丢包

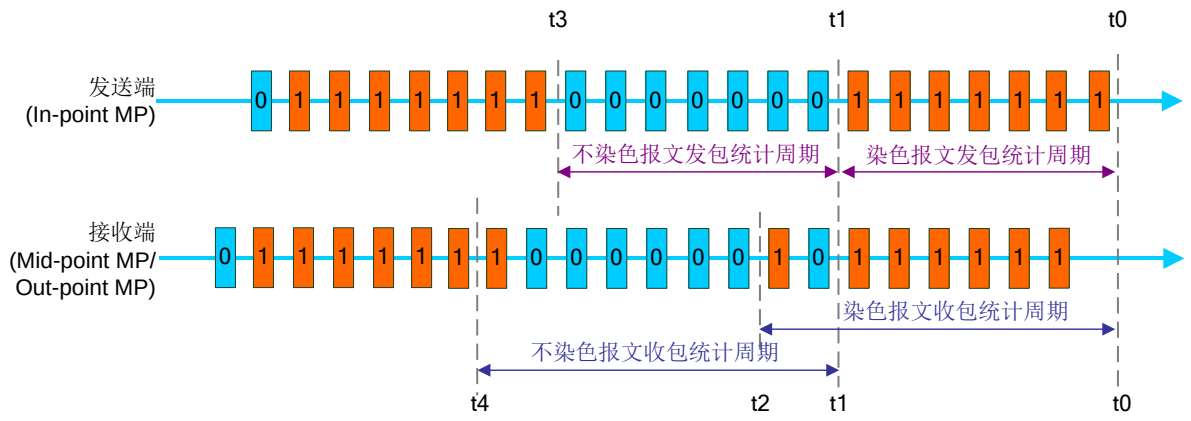


报文染色和计数机制

为方便用户及时了解网络丢包情况，iNQA 按周期测量丢包率，用户可查询每个周期内网络的丢包情况。iNQA 通过以下技术实现按周期测量并确保测量的准确性：

- 通过染色技术区别相邻统计周期内的报文。iNQA 使用 IP 报文头中 ToS 字段的 5～7 位中的任意比特作为染色位。 将染色位设置为 1 表示染色，设置为 0 表示不染色。 In-point MP 对目标流按周期交替地进行染色、不染色处理；Out-point MP 进行去染色处理，即将所有统计报文的染色位设置为 0。
- 在染色周期内，启用染色报文计数器统计染色报文的数量；在不染色周期内，启用不染色报文计数器统计不染色报文的数量。如下图 1-71 中的不染色报文 Y 不会统计在染色报文收包统计周期中。

图1-71 报文染色和数量统计



- iNQA 自动适当放宽收包统计周期，让收包统计周期比发包统计周期长一点，这样可以最大程度地避免网络延时与传输乱序对统计结果的不良影响。如图中染色报文 X 延时到达，接收端仍会将其统计到染色报文收包统计周期中。

工作机制

iNQA 工作过程分为三个阶段，如下图 1-72 所示：

(1) 所有参与测量的设备通过 NTP 或者 PTP 功能达到时间同步。在测量开始前，为确保各 Collector 能够基于相同的周期进行报文染色、上报、统计，所有 Collector 必须时间同步。如果时间不同步，会导致 iNQA 计算结果不准确。同时，为便于管理维护， 建议 Analyzer 和所有 Collector 之间时间同步。

(2) Collector 周期性收集 MP 产生的统计数据并上报给 Analyzer。

(3) Analyzer 对相同周期内相同目标流的报文进行丢包分析，计算报文丢失数（LostPkts）、报文丢失率（PktLoss%）、字节丢失数（LostBytes）、字节丢失率（ByteLoss%）。计算规则如下：

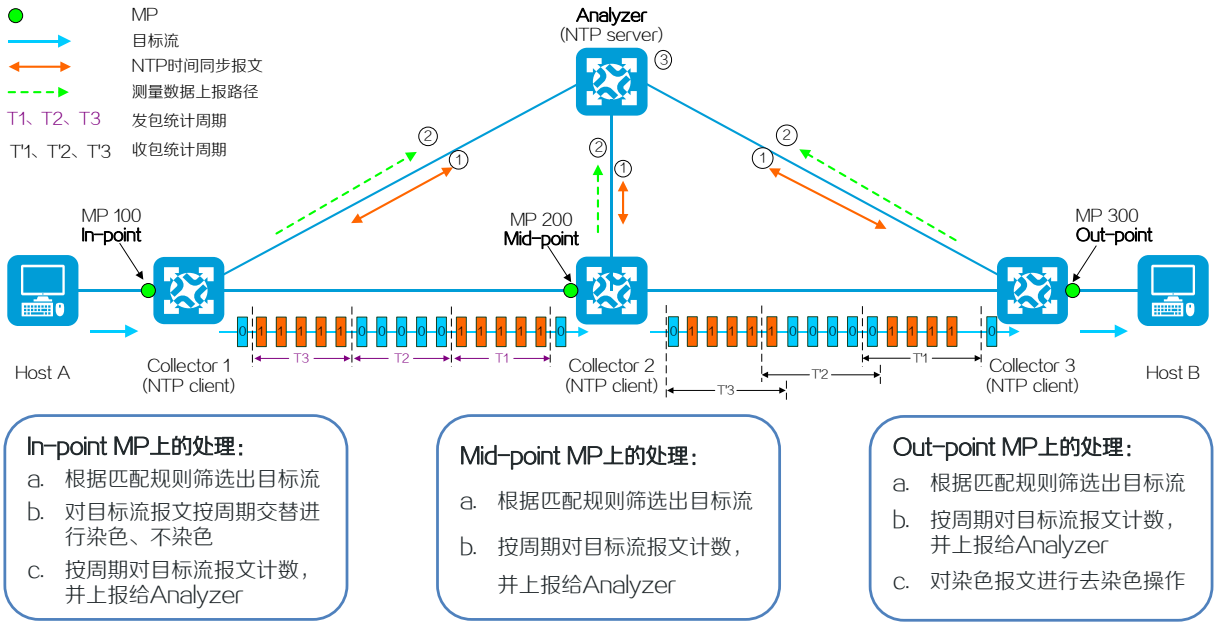
LostPkts = PktsIngress – PktsEgress　　PktLoss% = LostPkts/PktsIngress

LostBytes = BytesIngress – BytesEgress　　ByteLoss% = LostBytes/BytesIngress

表1-10 报文丢包分析

Period	LostPkts	PktLoss%	LostBytes	ByteLoss%
T1	1	20%	100	5%
T2	0	0	0	0
T3	0	0	0	0

图1-72 工作机制



音视频质量分析

音视频质量分析简介

多媒体音视频业务在日常生活中应用广泛，包括音视频会议、视频监控、视频播放等，用户对多媒体服务体验的要求也日益增高。通过部署音视频质量分析方案，设备可以实时监控基于 SIP 和 H.323 协议的音视频流量，并优先转发音视频流量。音视频质量分析方案还可以通过可视化方式展示音视频流量的质量分析结果，以便管理员快速发现和排除网络故障，改善并解决音视频质量问题，为用户提供良好的多媒体服务体验。

网络构成

- Client：使用音视频应用的终端。
- Device：音视频流量途经的设备，需要监控音视频流量、记录音视频会话信息并上报流量质量信息和会话信息。
- Voice server：语音服务器，用于管理语音用户的注册和呼叫。

- SeerAnalyzer（AD-WAN 智能分析组件）：SNA（SeerNetwork Architecture，先知网络架构）上的核心组件，可以实时采集网络业务流量数据，并通过大数据分析技术和人工智能算法可视化展示网络的运行情况。

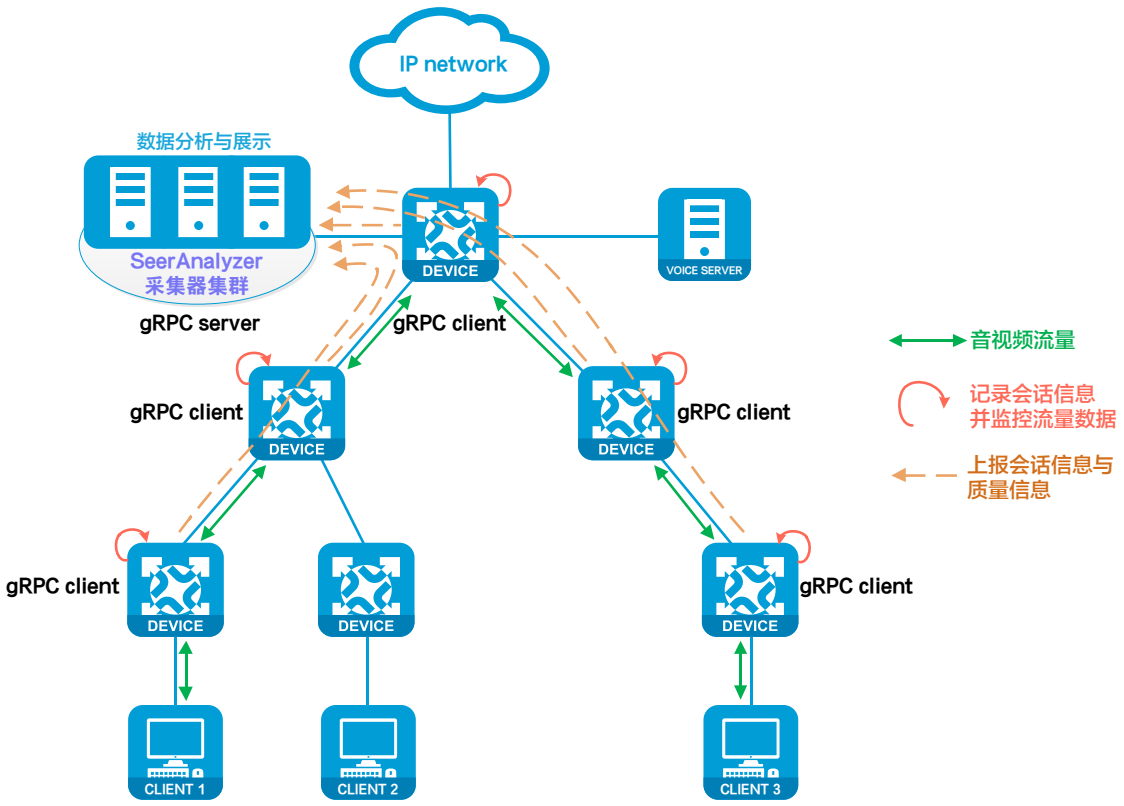
工作机制

如下图 1-73 所示，音视频质量分析方案中，音视频流途经的设备会通过 SQA 和 eMDI 功能监控音视频流量、记录音视频会话信息，并使用 gRPC 功能将会话和质量信息上报给 SeerAnalyzer，由 SeerAnalyzer 进行数据分析与可视化展示。

表1-11 术语解释

术语	中英文全称
SQA	Service quality analysis，服务质量分析
eMDI	Enhanced Media Delivery Index，增强型媒体传输质量指标
gRPC	Google Remote Procedure Call，Google远程过程调用

图1-73 音视频质量分析工作机制



(1) 音视频流量途经的设备通过 SQA 功能识别基于 SIP 或 H.323 协议的音视频流量，获取报文的五元组信息（源/目的 IP 地址、源/目的端口号和协议号），并提高音视频流量的转发优先级，以便优先转发音视频流量。

(2) 设备上的 SQA 模块将获取到的五元组信息通知给 eMDI 模块，eMDI 根据五元组信息对音视频流量的丢包、乱序和抖动数据进行监控。

(3) 设备作为 gRPC 客户端，通过 gRPC 协议报文将 eMDI 监控到的丢包、乱序、抖动等质量信息以及 SQA 记录的 SIP、H.323 会话信息上送给 SeerAnalyzer。

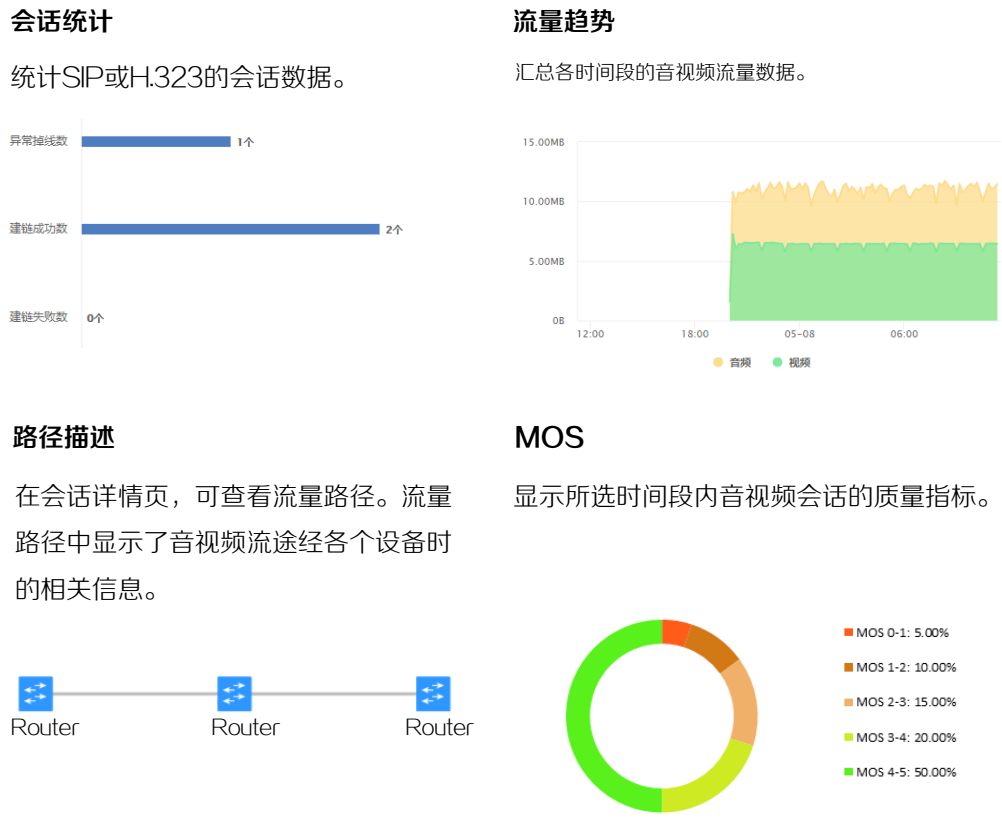
(4) SeerAnalyzer 根据上报的质量信息，对所有 SIP 和 H.323 流量的质量进行分析，并以可视化的方式展示音视频质量分析结果。网络管理员通过查看 SeerAnalyzer 展示的音视频质量分析数据和会话信息，了解网络状况，定位并解决网络中的问题，从而改善音视频业务的质量。

除了在 SeerAnalyzer 上查看音视频质量信息和会话信息，还可以在每台设备上通过 eMDI 的显示命令查看音视频流量的丢包率、时延、抖动、乱序率等信息，以及通过 SQA 的显示命令查看通话的详细信息，其中包含 MOS（Mean Opinion Score，服务质量评估值）信息。MOS 是音视频会话的质量指标，值越低表示质量越差。

可视化展示

如下图 1-74 所示，SeerAnalyzer 的音视频质量分析页面上，可以显示音视频流量的会话和质量信息，包括 SIP 和 H.323 的会话统计、流量趋势、路径描述、MOS 等信息。

图1-74 可视化展示



1.5 典型组网

总部+分支（Hub-Spoke 组网）

- 适用场景

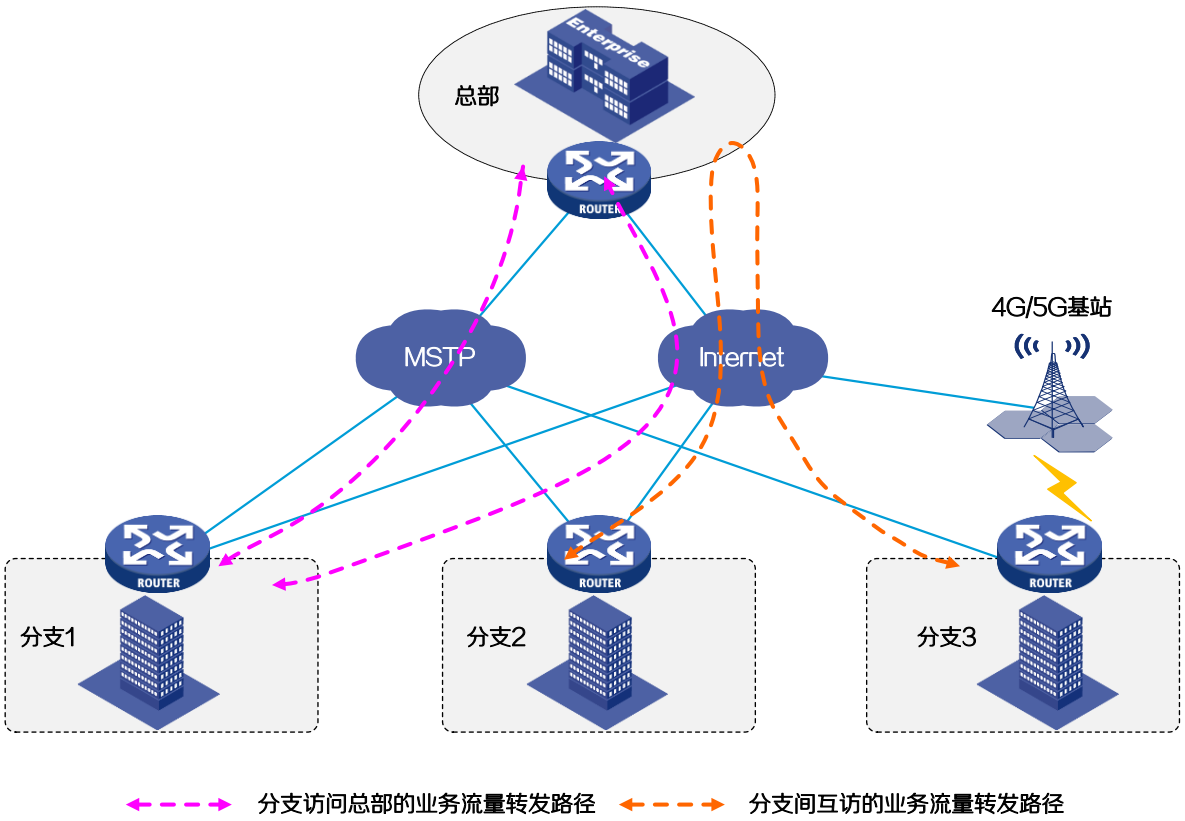
  适用于业务主要部署在总部，并且分支之间没有或者仅有少量的业务互访需求的企业。

- 特点
  - 分支通过 MSTP 专线或 Internet（例如 5G 网络）访问部署在总部的业务。
  - 分支设备可选用支持 5G 插卡的路由器实现 5G 网络接入功能。分支 5G 路由器具有大带宽，兼容 4G 网络、双 5G 卡高可靠性等特点。
  - 分支之间进行业务交互时，需要通过总部中转。

- 组网拓扑

  如图 1-75 所示，分支到总部采用单层 Hub-Spoke 组网，组网扁平化。其中，总部作为 Hub 站点、各分支作为 Spoke 站点。

图1-75 总部+分支（Hub-Spoke 组网）



分支入云（Hub-Spoke 组网）

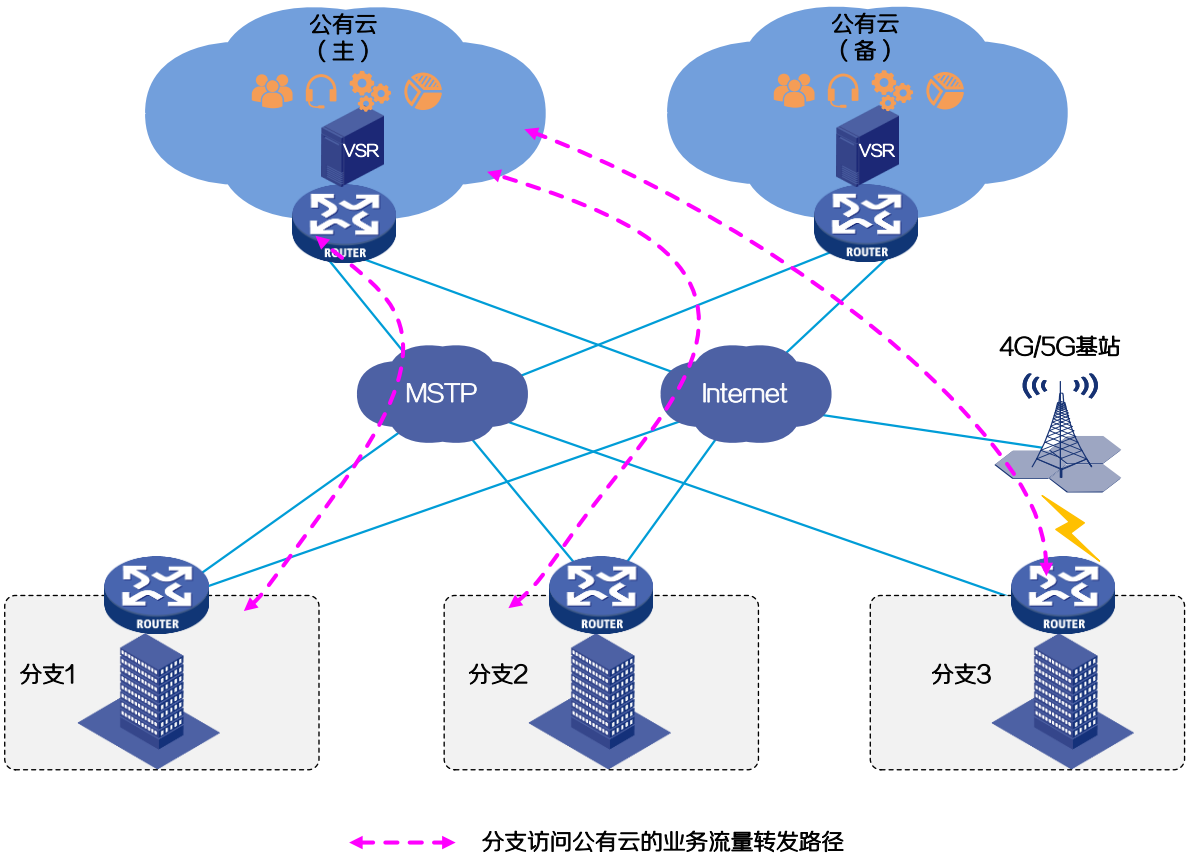
- 适用场景

  适用于业务部署在公有云，分支通过公有云访问这部分业务的企业。

- 特点
  - VSR 可以直接在公有云部署。
  - 支持部署主备公有云，当主公有云故障时，分支业务可切到备公有云上，提高可靠性。
  - 支持多种云部署，例如：阿里云、紫光云、天翼云等。
- 组网拓扑

  如图 1-76 所示，分支到公有云采用单层 Hub-Spoke 组网，组网扁平化。其中，公有云作为 Hub 站点、各分支作为 Spoke 站点。

图1-76 分支入云（Hub-Spoke 组网）



总部+分支（多总部组网）

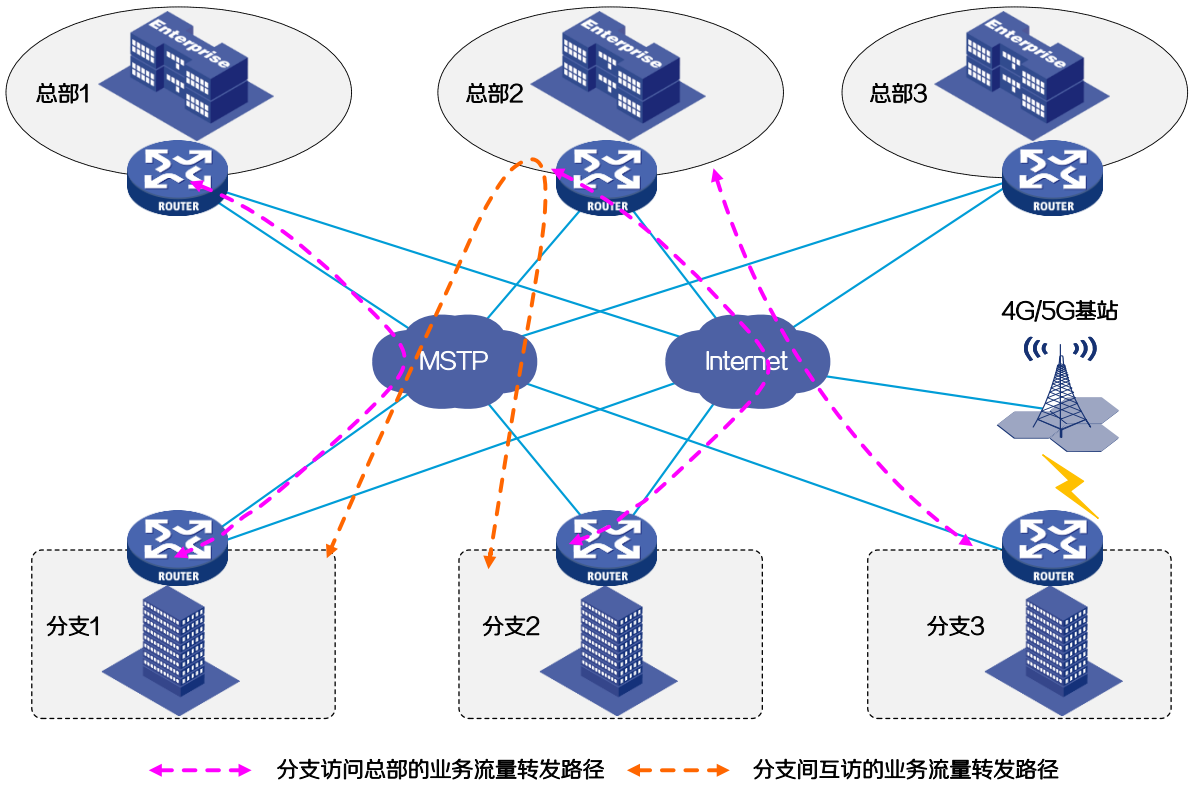
- 适用场景

  适用于业务部署在多个总部，并且每个总部均需要为分支提供业务支持的企业。

- 特点
  - 支持多总部多活，即该组网中可以同时部署多个总部，多主多备，其中主总部负责处理业务，备总部提供备份，例如三总部，则可部署 2 主 1 备。采用多主多备的方式，既可以在多活总部之间进行负载分担和业务备份，又可以通过备总部进一步提供备份，增强可靠性保障。
  - 分支之间进行业务交互时，需要通过总部中转。
- 组网拓扑

  如图 1-77 所示，多总部的本质是采用多个 Hub 站点的 Hub-Spoke 组网。

图1-77 总部+分支（多总部组网）



总部+分支（Full-mesh 组网）

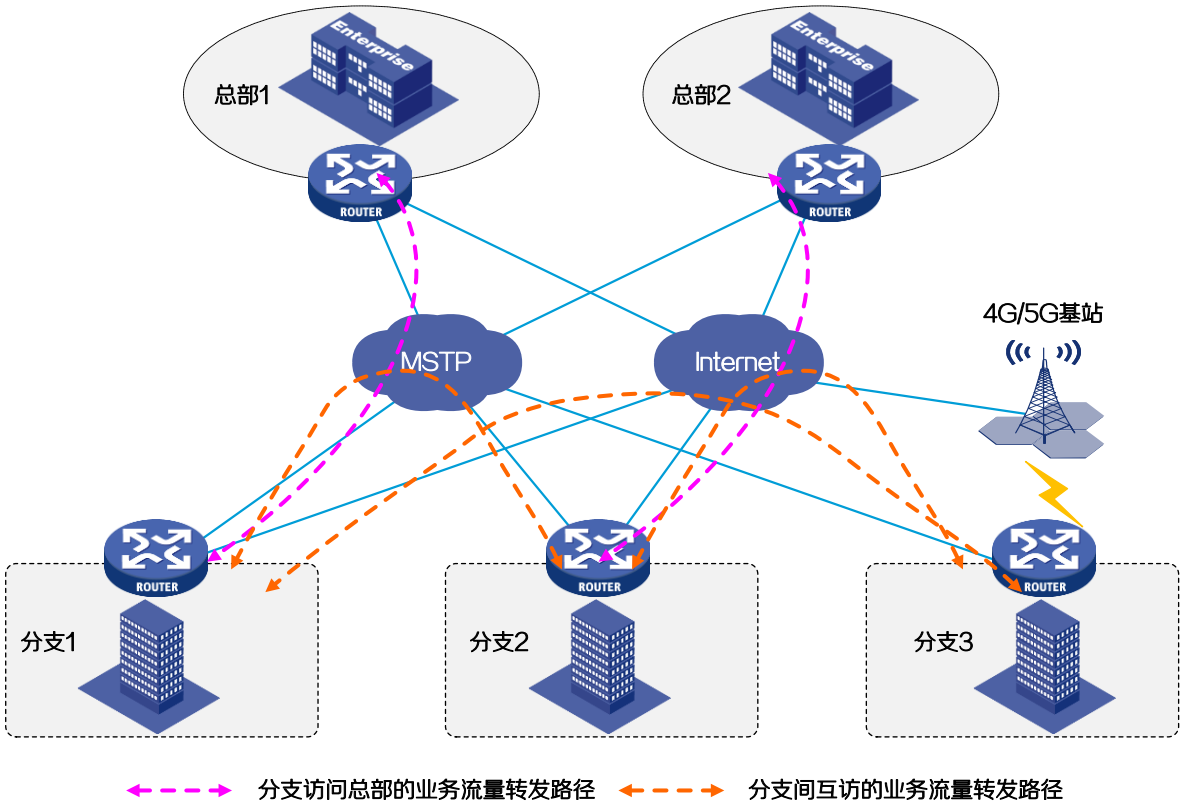
- 适用场景

  适用于分支不多，或者分支之间有大量业务互访需求的企业。
- 特点
  - 灵活拓扑、弹性扩容。
  - 分支之间支持 Full-mesh 组网和 Partial-mesh 组网。
  - 分支之间进行业务交互时，不需要通过总部中转，可以直接互通。

- 组网拓扑

  如图 1-78 所示，在 Full-mesh 组网中，分支和总部之间，以及各个分支之间建立全连接。

图1-78 总部+分支（Full-mesh 组网）



总部+汇聚+分支（三级分层组网）

- 适用场景

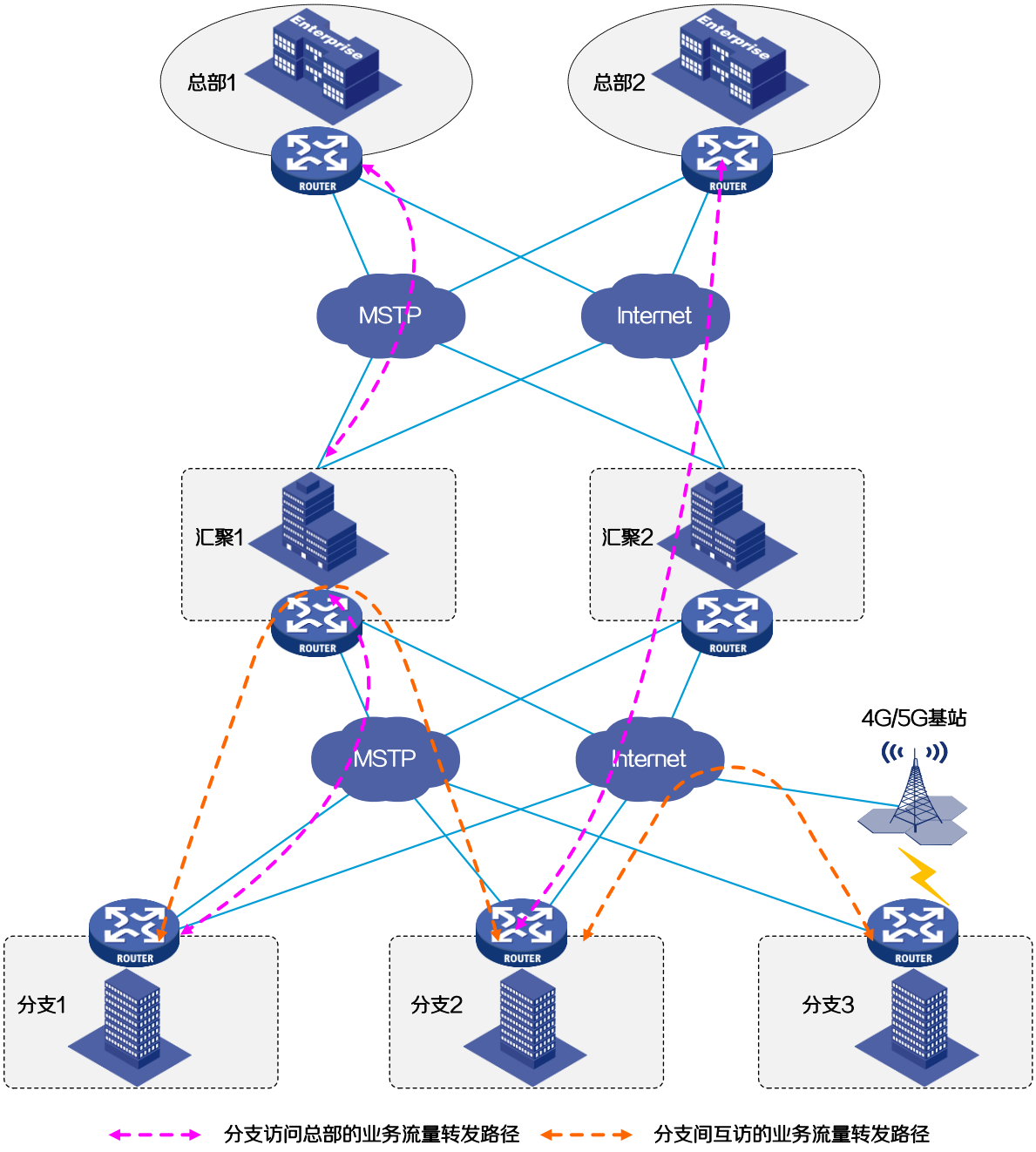
  适用于省-市-地县，或者全国-各省-各市等三级网络架构明确的企业。

- 特点
  - 网络结构清晰，网络规模易扩展。
  - 分支到总部可以逐级分段接入，也可以直通总部。
  - 分支之间进行业务交互时，可以通过汇聚节点中转，也可以直接互通。

- 组网拓扑

  如图 1-79 所示，三级分层组网可以看作是两段单层组网模型的叠加。在该组网中，通过将整个网络划分成多个区域，之后再将多个区域之间通过集中的汇聚节点进行互联。

图1-79 总部+汇聚+分支（三级分层组网）



多租户 POP 组网（MSP 运营组网）

- 适用场景

  适用于希望由运营商/MSP（Managed Service Provider，管理服务提供商）代建广域网分支网络的企业。

- 特点

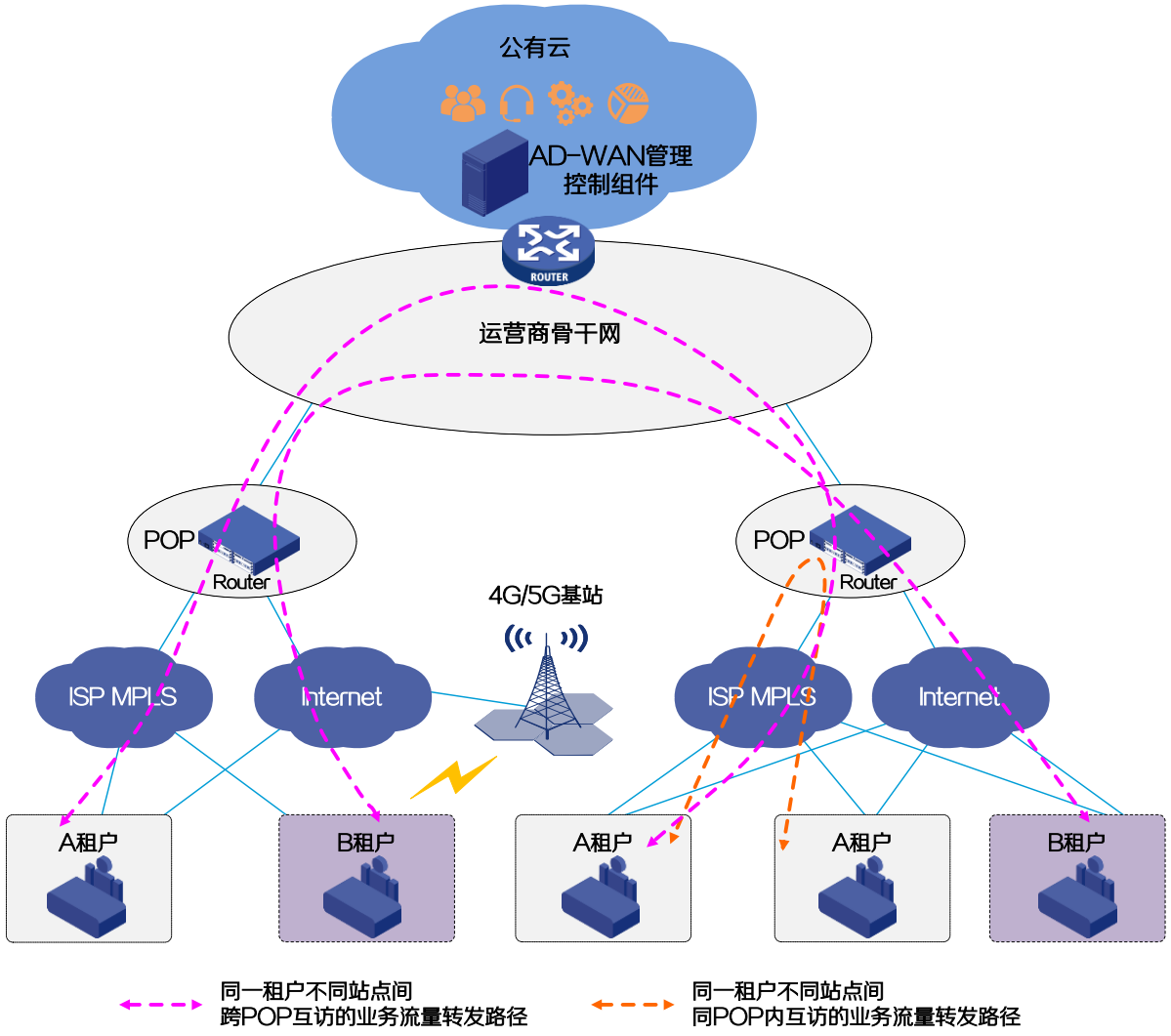
  - 企业可直接租用运营商的广域网分支网络服务，降低企业网络建设成本。

- POP（Point Of Presence，入网点）站点内相同租户通过 POP 站点互通，跨 POP 站点租户通过 POP 点接入运营商骨干网，以满足全球运营需求。
- 各租户路由、业务隔离，使用 HQoS 技术为不同的租户提供不同的服务。

- 组网拓扑

  如图 1-80 所示，多租户 POP 组网需要部署 POP 站点，租户站点和 POP 站点按照 Hub-Spoke 方式组网。

图1-80 多租户 POP 组网（MSP 运营组网）



1.6 成功实践

大型集团

业务挑战

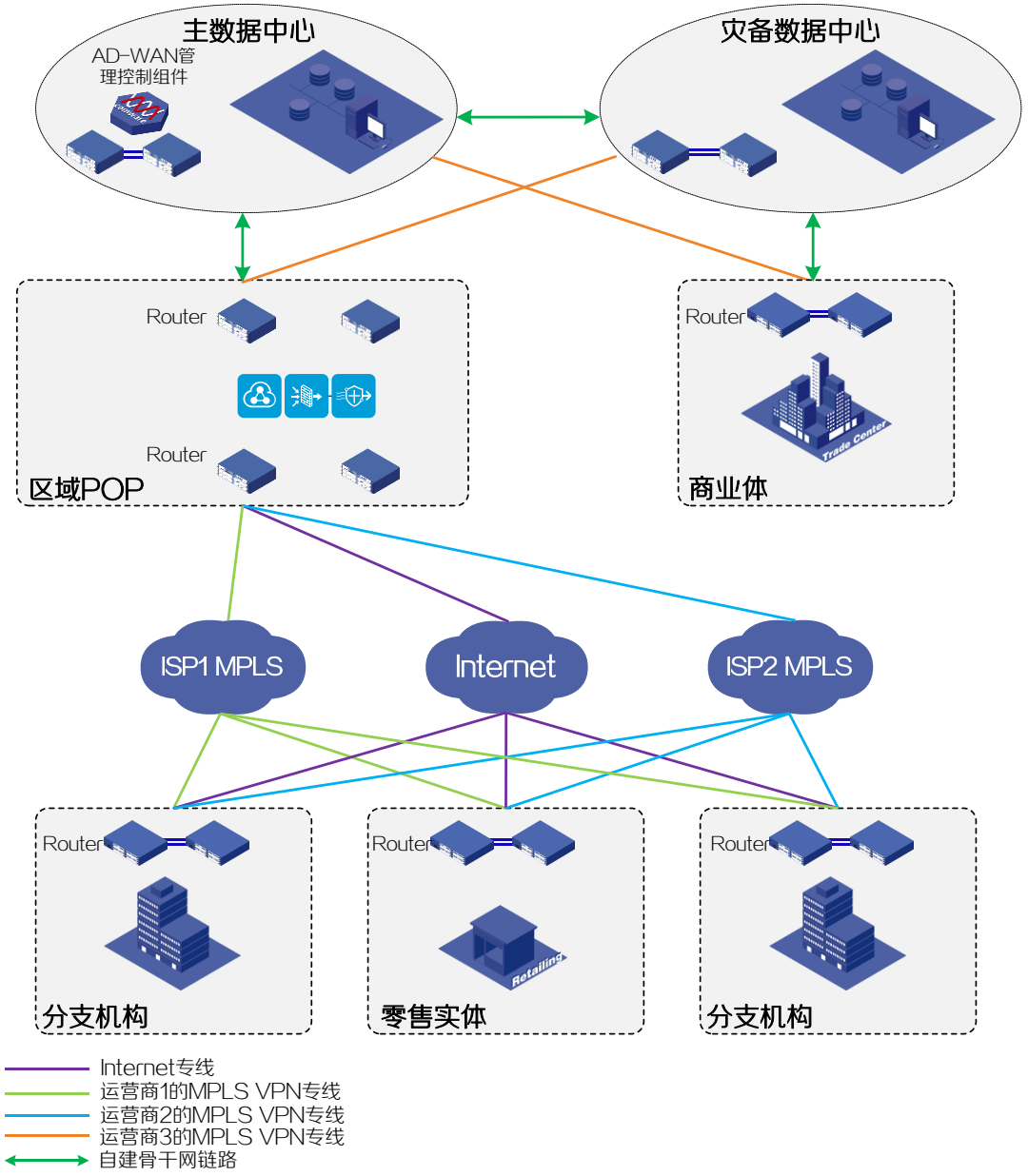
某集团公司是一家跨地产，制造，医药等行业，并且拥有众多实体公司的大型集团，现网存在如下问题：

- 分公司多且分散在全国各地，业务上线和站点开局的周期长。
- 专线带宽不足，扩容专线成本高。
- 现网的专线利用率不均衡，带宽未被充分合理地利用。
- 专业维护人员少，各分公司和总公司都缺少具有分级分权管理能力的统一管理平台，为实现业务流量负载分担和关键业务保障，需要复杂的手工配置，网络维护复杂且效率低下。

组网方案

大型集团 AD-WAN 分支解决方案架构如图 1-81 所示。

图1-81 大型集团 AD-WAN 分支解决方案架构



- 采用二级三级混合组网的架构部署集团组网方案，分支机构先接入区域 POP 点，再上联至数据中心，分支之间的互通避免绕行数据中心，部分支机构无法接入区域 POP 点则通过运营商链路直接接入数据中心。
- 分支机构采用 MPLS VPN 专线及互联网专线（IPsec VPN）接入数据中心。
- 双数据中心实现主备可靠性保护，AD-WAN 管理控制组件采用集群部署的方式部署在主数据中心。
- 所有分支机构均采用双设备及双链路的方式上行接入多 ISP 出口。

客户价值

缩短开通周期

部署 AD-WAN 分支解决方案后，站点开通时间从之前的 3 到 5 天缩短为 10 分钟，开通效率提升 95%以上。

降低专线成本

现网部署 AD-WAN 分支解决方案，将 MPLS 专线改为价格低廉的互联网线路，并采用灵活拓扑混合组网等接入方式，线路成本降低 70%。

提升运维效率

AD-WAN 分支解决方案实现全网自动化下发 VPN 业务、LAN 业务和 QoS 业务，可以基于应用、链路质量和时间等维度来调度流量，业务变更更加灵活快速，业务变更效率提升 400%。通过全网可视化和智能运维，缩短故障排错时间，网络运维成本节约超过 45%。

保险行业

业务挑战

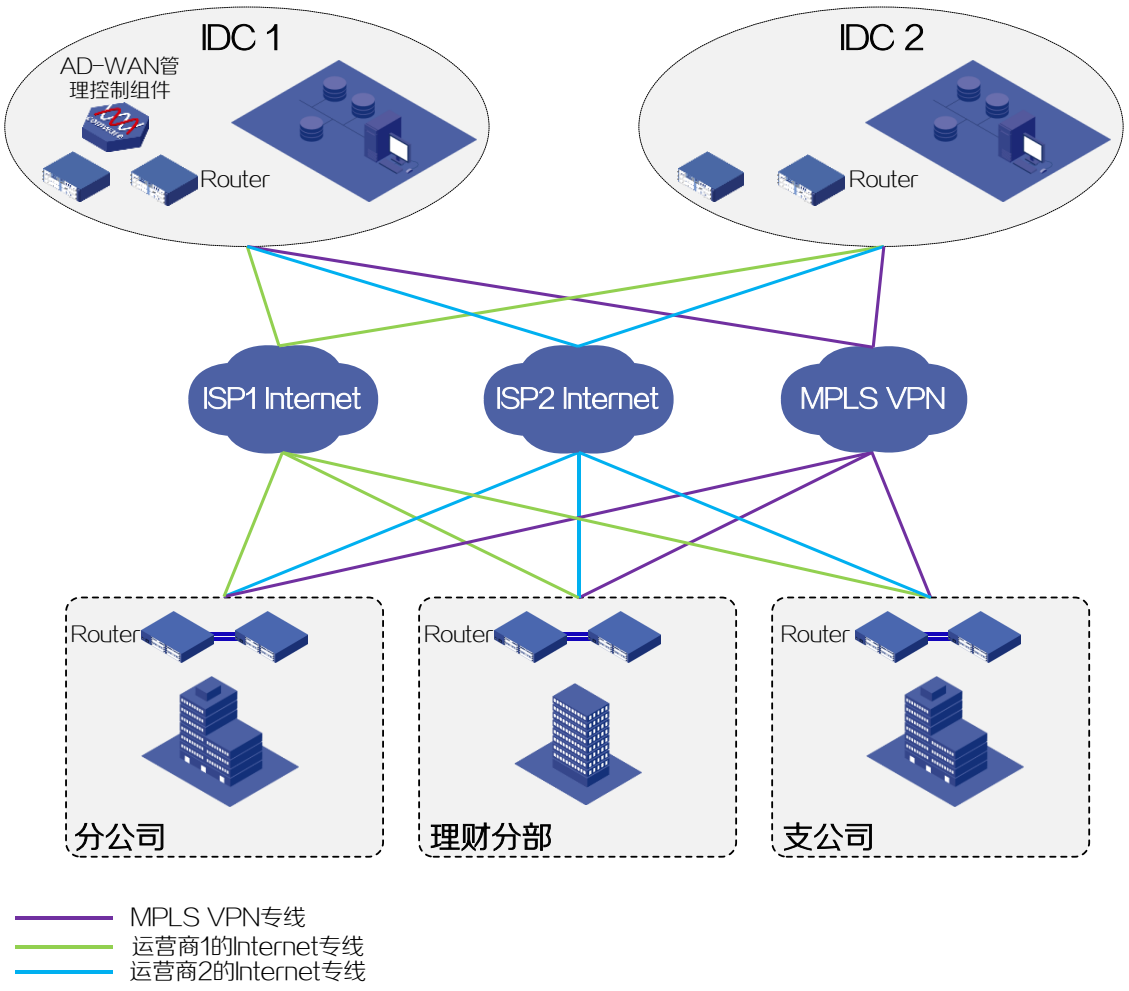
某保险公司现网面临以下问题：

- 现网视频、语音、图像等新业务快速增长导致线路带宽不满足业务需要，同时传统网络无法灵活地根据业务选择不同类型的线路，专线带宽扩容成本高，而且无法保障关键业务质量。
- 传统网络基于业务调度流量转发路径的配置复杂，维护难度大，不具备可视化维护界面。
- 分支机构的出口设备访问互联网时，需要具备 IPS、上网行为管理等安全管控能力，现网设备的网络与安全功能分离，设备多且复杂不易管理，维护成本高。

组网方案

保险行业 AD-WAN 分支解决方案架构如图 1-82 所示。

图1-82 保险保险行业 AD-WAN 分支解决方案



- 每个站点采用 1 条 MPLS VPN 专线及 2 条不同运营商的互联网专线接入数据中心。
- 部署双数据中心，每个数据中心部署双网关，数据中心保持双活状态。
- 本地访问互联业务通过设备内置的状态防火墙、IPS、用户行为管理等安全功能保证数据访问安全。

客户价值

灵活的业务调度

AD-WAN 分支解决方案提供丰富的智能流量调度策略。将大带宽的视频流业务调度至互联网专线承载大大降低 MPLS VPN 专线带宽压力，降低带宽扩容成本。将关键业务流量调度到 MPLS VPN 专线承载，同时使用互联网专线进行业务备份保障。基于应用对业务进行分级差异化 QoS 保障。

可视化运维

AD-WAN 分支解决方案提供可视化运维界面，通过简单的页面操作来完成互联网访问模式的切换，简化配置过程，提高运维效率。

银行系统

业务挑战

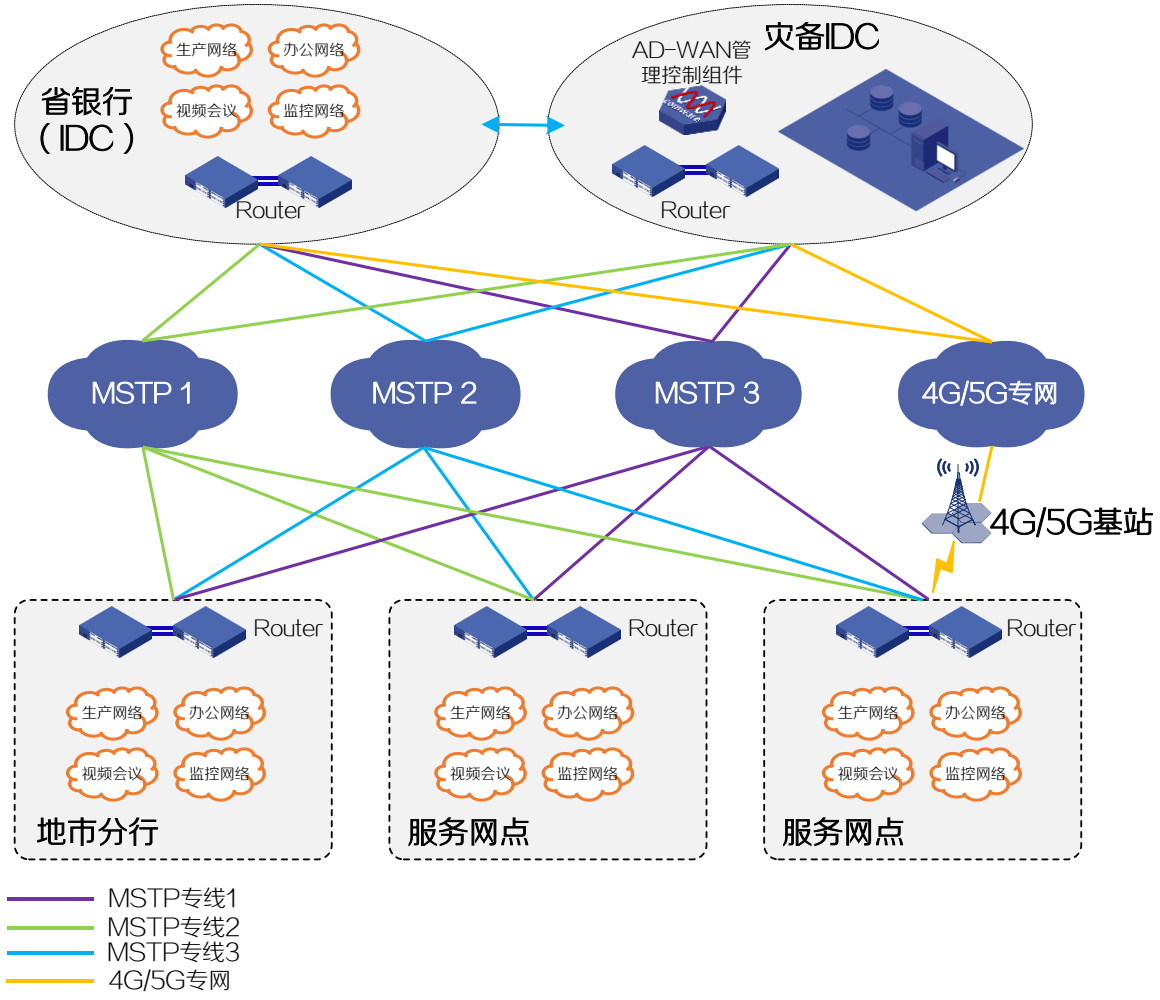
某银行现网面临的问题如下：

- 省行、地市支行网络分散管理，网络变更需要大量配置，维护成本高。
- 现网中生产、办公、监控、视频会议四条专线负载不均，造成带宽浪费，传统流量调度方式配置复杂。
- 银行网络中新业务越来越多样，例如人脸识别、用户自助等业务对时延及可靠性要求较高。而视频、图像等业务对带宽需求极大，现网无法满足不同业务多样化 QoS 需求。

组网方案

银行系统 AD-WAN 分支解决方案架构如图 1-83 所示。

图1-83 银行的 AD-WAN 分支解决方案架构



- 服务网点和地市分行通过多条专线上行至省银行及灾备数据中心，每个服务网点或地市分行均采用双设备接入，保证接入点的可靠性。
- 生产网络，办公网络，视频会议网络和监控网络通过不同专线承载，避免监控业务抢占关键的办公生产业务带宽。

客户价值

可视化运维

AD-WAN 分支解决方案提供可视化运维界面，通过简单的页面操作来完成配置和运维，简化配置过程，提高运维效率。

灵活的流量调度策略

AD-WAN 分支解决方案提供基于应用和业务类型的流量调度策略，可以采用多条专线的链路进行分担和备份保护，提高业务可靠性并均衡各条专线链路的带宽利用率。基于应用对业务进行分级差异化 QoS 保障。

零售行业

业务挑战

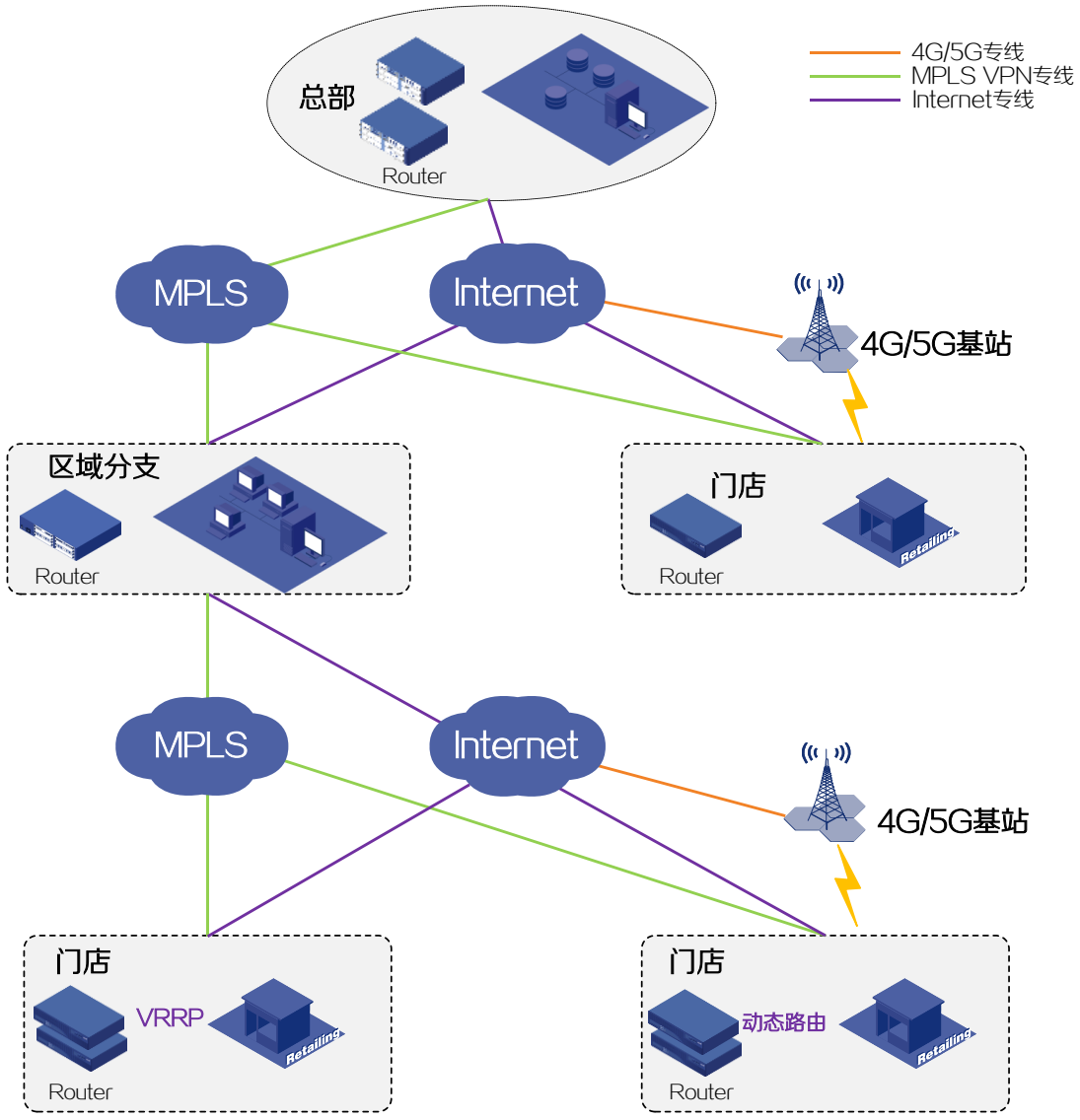
某大型连锁超市拥有分支和门店数量超过 700 个，分支和门店遍及省内及省外多个地市。目前现网面临如下运维挑战：

- 门店位置分散，门店网络的开通需专业人员到现场处理，开通周期长，人工成本高；
- 现网中 IPSec VPN、QoS 等业务均需要维护人员手工部署，工作量巨大，对于维护人员的技能要求较高。
- 分支和门店的网络故障定位难，网络运维不可视。

组网方案

AD-WAN 分支解决方案架构如图 1-84 所示。

图1-84 零售行业的 AD-WAN 分支解决方案架构



- 总部数据中心、分支区域数据中心、门店采用二级和三级混合组网的架构。在三级组网架构中，门店先连接到分支区域数据中心，分支区域数据中心连接到总部数据中心。在二级组网架构中，门店先直接连接到总部数据中心。
- 业务通过 MPLS、Internet 或 4G/5G 无线链路接入各数据中心，不同接入方式的链路故障时，业务可以进行快速切换，保证业务高可靠。
- 接入设备侧使用 VRRP 或动态路由实现保护。

客户价值

极简开局

部署 AD-WAN 分支解决方案后，可以使用零配置（ZTP，Zero Touch Provisioning）开局，无需专业人员往返现场，减少现场工作成本。每年节省现场运维成本 50 万，站点开通时间从之前用时 3～6 天缩短为 10 分钟。

自动化业务部署

部署 AD-WAN 分支解决方案后，IPSec VPN、QoS 等业务可以统一编排，自动化部署和配置，减少业务部署工作量。整体网络建设成本节约了超过 40%。

智能运维

AD-WAN 分支解决方案提供可视化界面集中监控门店设备状态、链路质量，辅助快速定位故障。运维效率提升 60%。

MSP (Managed Service Provider)

业务挑战

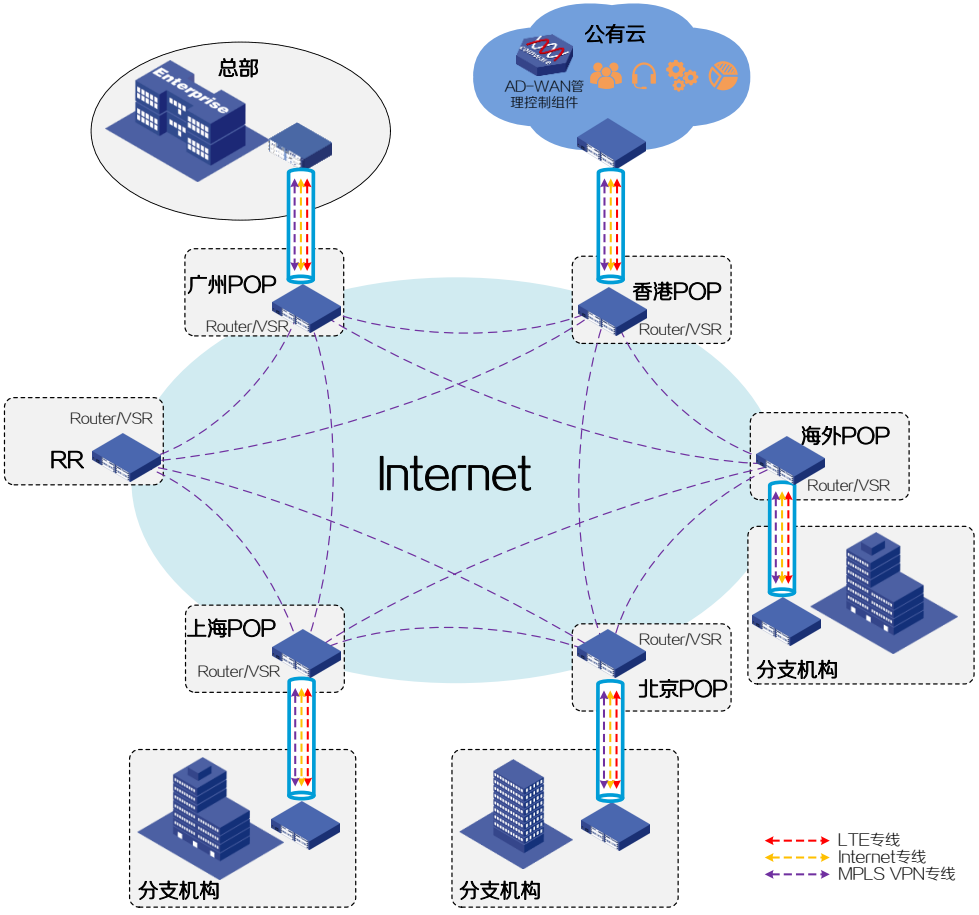
某机构将其网络集成服务和管理业务委托给国内某运营商/MSP，该机构现网面临如下问题：

- 缺少专业 ICT 运维人员，网络运维能力薄弱，希望能够有简单智能的运维手段。
- 下属的分支机构众多，遍布全球，分支机构的开通部署周期长，且需要专业人员现场提供技术支持。

组网方案

运营商/MSP 采用 AD-WAN 分支解决方案为该机构提供管理和运维服务，AD-WAN 分支解决方案架构如图 1-85 所示。

图1-85 运营商/MSP 采用的 AD-WAN 分支解决方案架构



- 每个分支机构都可以通过多线路（Internet 专线、LTE 专线和 MPLS VPN 专线）上行接入 POP 点。POP 点间通过 Internet 互联。
- 区域内 POP 点网关和 RR 合一部署，区域间部署独立 RR 设备。
- 分支站点可以同时接入两个 POP 点以保证接入可靠性。
- 多 VPN 部署可以满足不同客户业务隔离的需求。

客户价值

降低人力成本

由第三方为机构提供网络运维服务，机构无需自建专门 ICT 运维团队，节省了人力成本。

极简开局

AD-WAN 分支解决方案提供零配置开局方案，设备快速上线。分布在不同地区的分支机构需要开局时，无需专业人员往返现场。

智能运维

AD-WAN 分支解决方案提供可视化管理平台和自动化运维服务，简化传统的运维方式。使运营商代维代建成为可能，AD-WAN 管理控制组件掌握全网，不论最终客户在哪里，都可以就近接入 POP 点。运营商可以远程管理最终用户的接入设备。

第 2 章
云简广域网解决方案

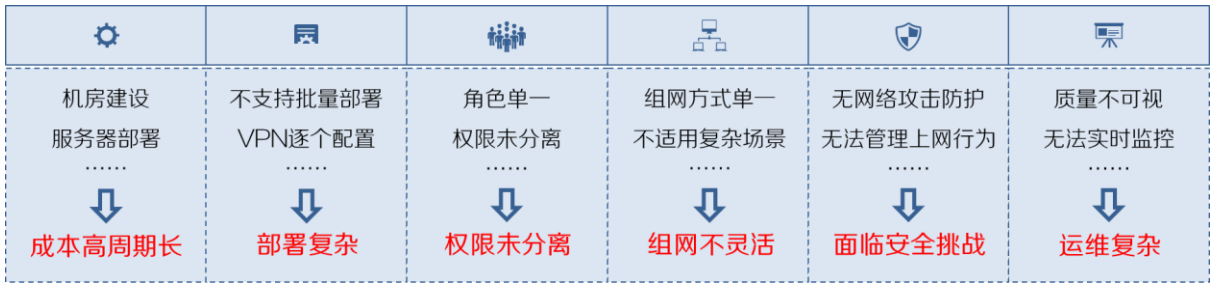
摘要

本章主要介绍传统 VPN 方案存在的问题、云简广域网方案概述、客户价值、典型组网和云简广域网解决方案在百行百业的成功实践。

2.1 传统VPN方案存在的问题

目前传统 VPN 集中管理方案虽然实现了远程部署，在一定程度上简化了用户工作，但仍存在一些不足，主要的问题如图 2-1 所示。

图2-1 传统 VPN 方案存在问题

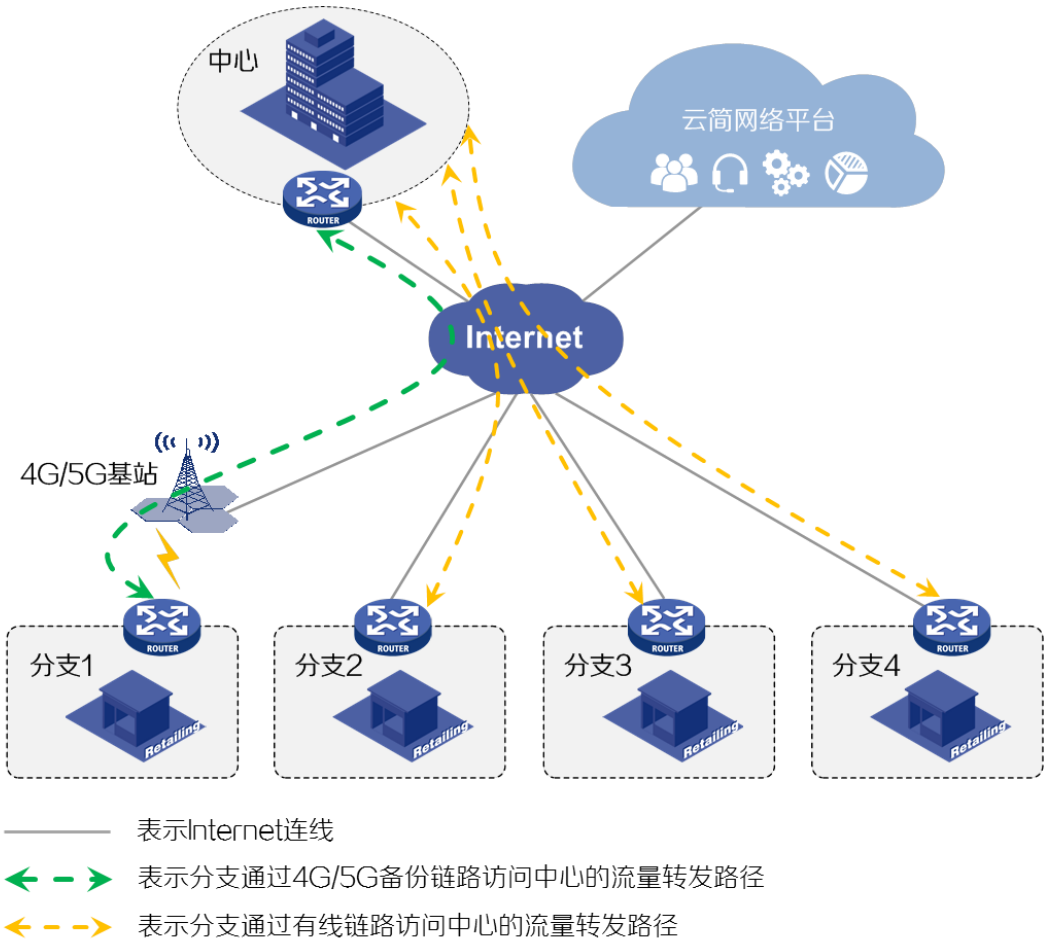


- 成本高周期长：机房占地空间大，建设成本高。需要部署服务器、安装软件，网络建设周期长。
- 部署复杂：多分支场景下进行 VPN 部署和管理时，需要逐个配置 VPN 地址、VPN 加密流量等参数，耗时长、效率低，且容易出错。
- 权限未分离：角色单一，未根据角色划分不同的管理权限，无法实现网络分权分级管理。
- 组网不灵活：组网方式单一，不能便捷地部署双中心、多出口等多分支组网架构，不能灵活适用复杂场景。
- 面临安全挑战：无法对用户的上网行为进行管理，容易遭受到网络攻击，且缺乏抵御网络攻击的能力。
- 运维复杂：无图形化管理界面，不具备质量可视化的功能。无法对设备状态、网络健康状态、终端接入情况等进行实时监控。

2.2 云简广域网解决方案介绍

为了解决传统 VPN 方案的六大不足，H3C 提出了云简广域网解决方案，组网架构如图 2-2 所示。

图2-2 云简广域网解决方案架构图



云简广域网解决方案通过如下表 2-1 所示的技术，解决了传统 VPN 方案面临的成本高周期长、部署复杂、权限未分离、组网不灵活、面临安全挑战和运维复杂的问题。

表2-1 云简广域网解决方案技术

传统 VPN 方案面临的问题	云简广域网解决方案技术
成本高周期长	依托云简网络平台，采用公有云服务，无需自建机房、配置服务器，注册即可享用云简网络管理服务
部署复杂	支持批量部署多分支组网，“三步”即可完成VPN业务的下发，极大地减少了网络管理时间和成本
权限未分离	支持创建子账号，并为不同的角色划分不同的管理权限，实现结构化、多层级、批量化的网络管理
组网不灵活	支持Internet、4G/5G等多种线路组网，同时提供专业化的4G/5G管理能力，支持单Hub双出口、双Hub双出口等多种组网模型

传统 VPN 方案面临的问题	云简广域网解决方案技术
面临安全挑战	支持多种深度安全特性，比如IPS（入侵防御系统）、APR（应用识别）、URL过滤等 云简广域网解决方案支持配置“防火墙功能”，支持对不同应用进行应用限速和带宽保障的配置
运维复杂	采用图形化界面，支持对所有分支网络一站管控，支持对各类型设备、网络健康状态、终端接入情况、VPN状态、VPN隧道质量等进行实时监控 提供专业化的4G/5G路由器管理能力 支持通过告警订阅功能提供多种告警机制，实现透明化的网络管理，提高运维效率

2.3 客户价值

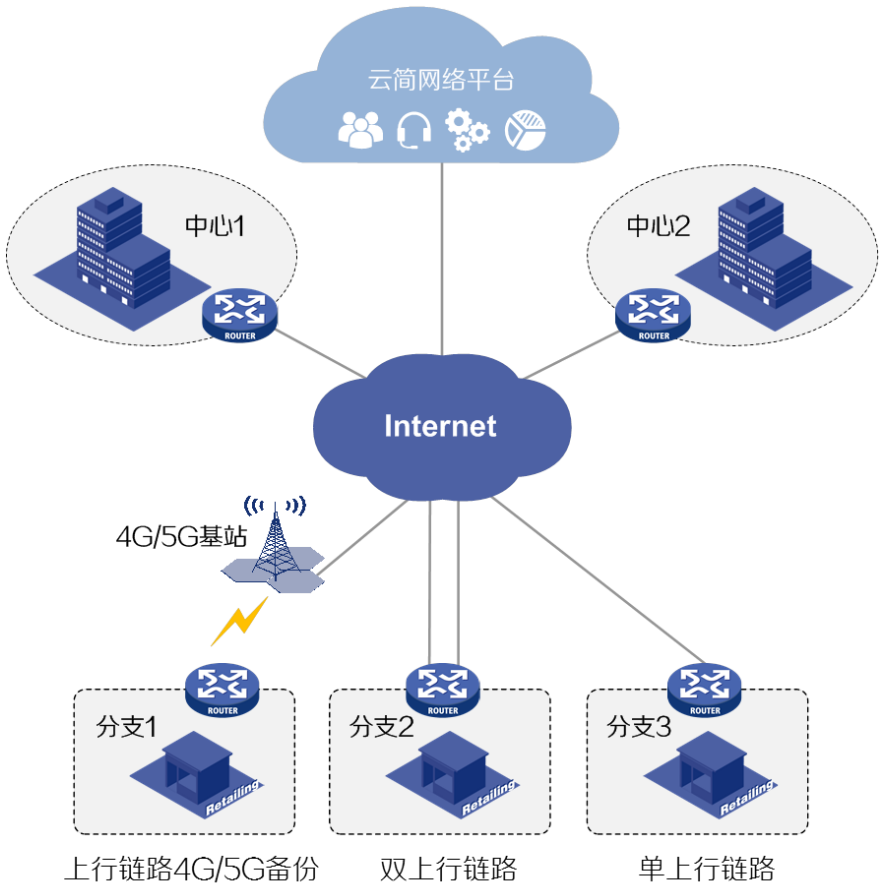
多——线路多样，特性丰富

线路多样，灵活可靠

- 线路多样

  云简广域网解决方案支持 Internet、4G/5G 等多种线路组网，如图 2-3 所示，分支设备上的两条链路可以使用双有线组合、单有线链路单 4G/5G 备份链路等组合方式。

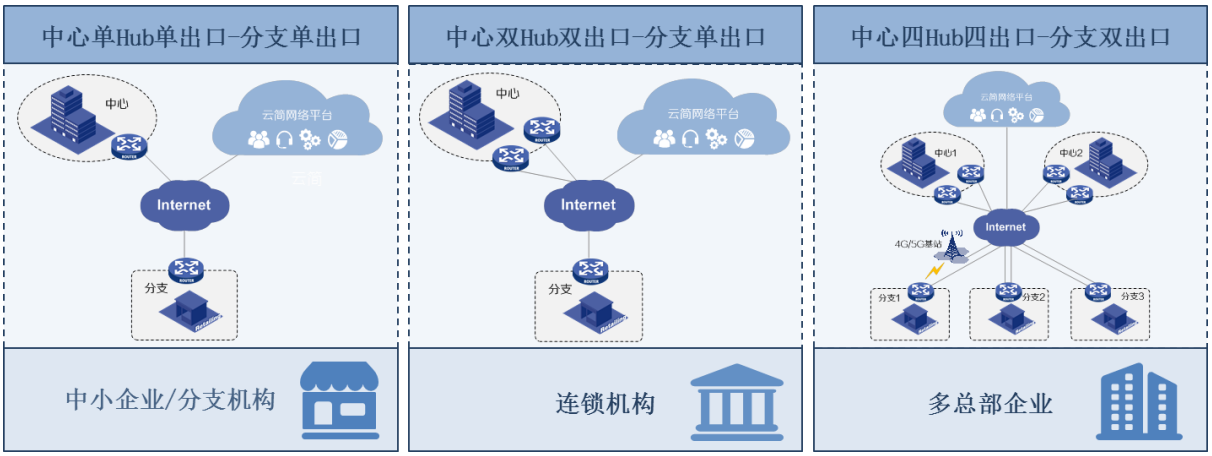
图2-3 多种线路组网图



- 组网灵活

  云简广域网解决方案支持分支设备穿越 NAT 场景，如图 2-4 所示，中心支持单 Hub 双出口、双 Hub 双出口等多种组网模型，分支设备可同时与中心建立多个 VPN 隧道，形成主备备份或者负载分担，提供数据传输的高可靠性。

图2-4 组网灵活



特性丰富，深度安全

云简广域网解决方案通过结合 H3C 路由器的深度安全特性，比如 IPS（入侵防御系统）、APR（应用识别）、URL 过滤等，能实现防御网络攻击、管理用户上网行为的目的

- IPS 通过分析流经设备的网络流量应用层信息实时检测和阻断各种恶意行为，实现保护企业信息系统和网络免遭攻击的目的。并且可将 IPS 功能的配置命令保存在模板文件中，通过云简网络平台，实现批量配置。
- APR 能够识别用户业务流量所属应用，并可针对应用实现限速、放行、丢弃等控制动作；
- URL 过滤能够对用户访问的 URL 进行控制，即允许或禁止用户访问 Web 资源。

IPS、APR、URL 过滤所使用的特征库不断对最新的网络攻击、应用、服务类型进行更新涵盖，特征库版本自动更新，实时保持业界最佳安全防护能力。

并且，云简网络平台支持配置“防火墙功能”，可配置基于五元组、应用组、URL 及 URL 黑白名单的防火墙规则，同时支持配置对不同的应用进行限速和带宽保障的配置。

快——集中部署，快速上线

集中部署，快速上线

- 快速配置

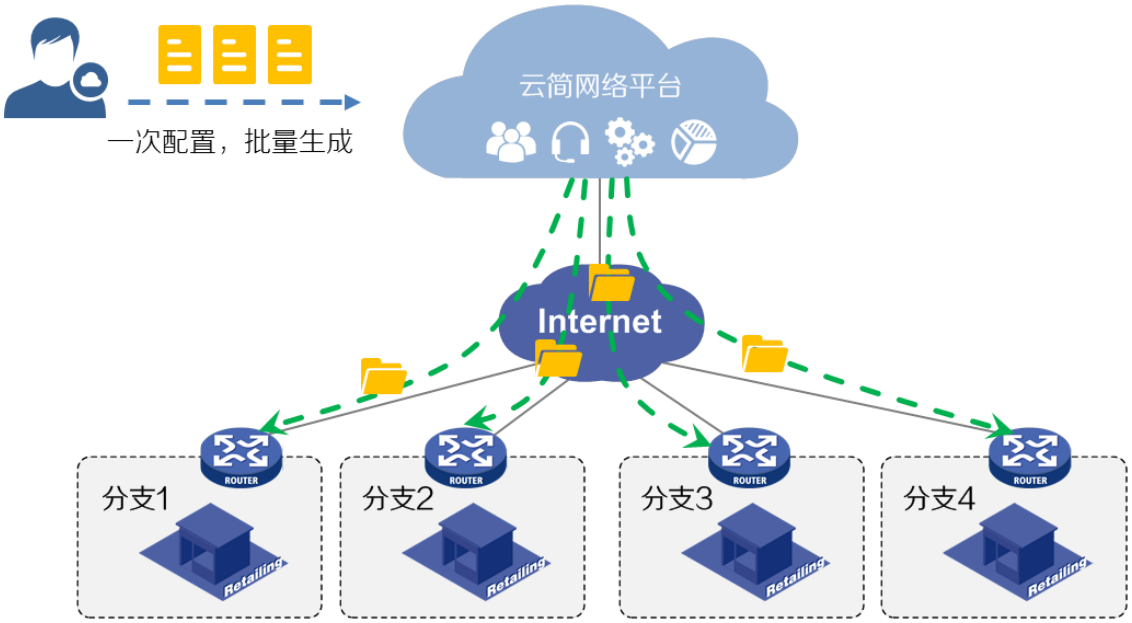
  云简广域网解决方案通过如图 2-5 所示的“创建中心 VPN-创建分支 VPN-配置安全策略”三步即可完成 VPN 业务的下发，极大地简化了 VPN 的配置复杂度。

图2-5 三步快速配置 VPN 业务

- 批量部署

  云简广域网解决方案支持批量部署多分支组网，如图 2-6 所示，极大地减少了网络管理时间和成本。

图2-6 批量部署多分支

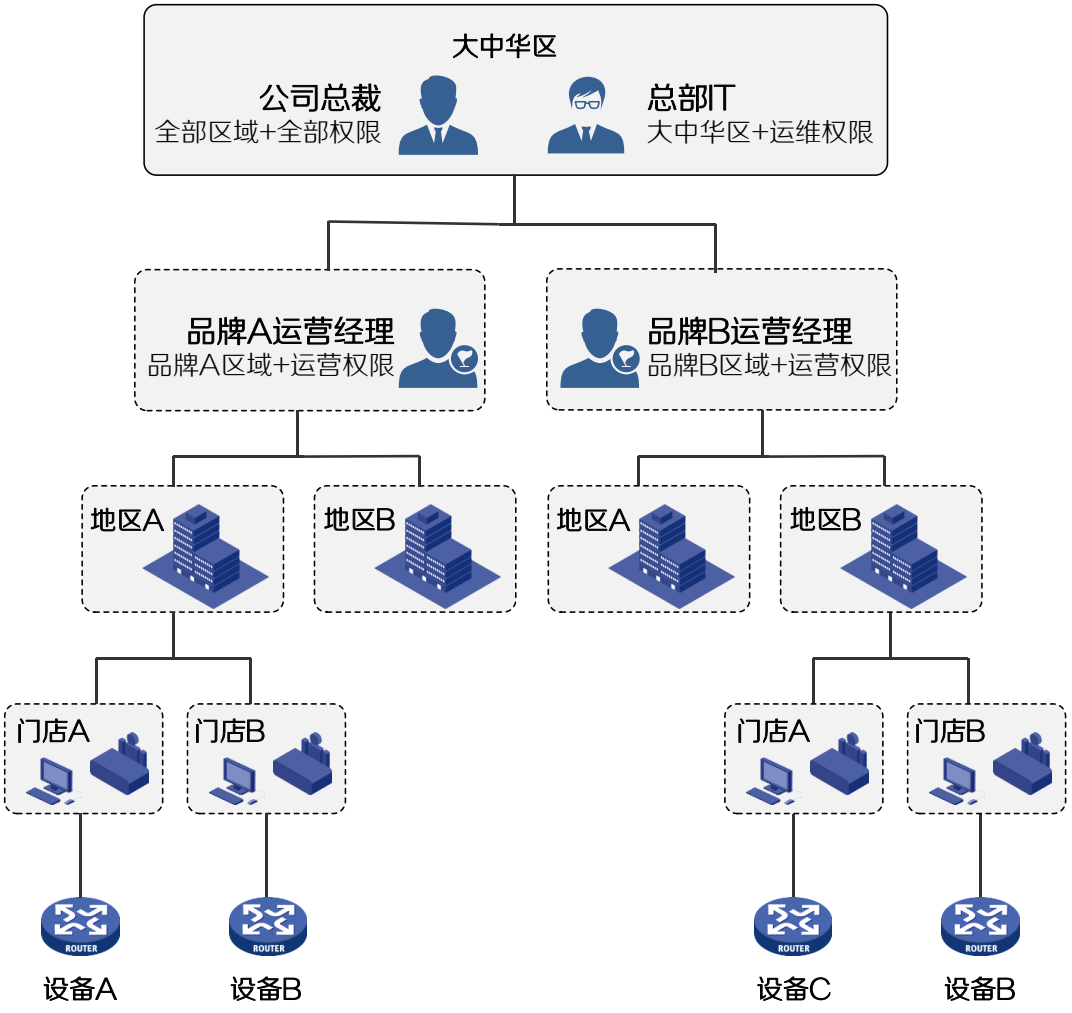


好——分权分域，整网可视

分权分域，管理清晰

如图 2-7 所示，云简广域网解决方案可贴合企业的组织架构建立网络组织架构，支持在云端根据品牌、地域等灵活划分分支网络，实现结构化、多层级、批量化的网络管理。同时，在云简网络平台上支持创建子账号，并为不同的角色划分不同的管理权限，实现分权分级管理。

图2-7 云简广域网解决方案分权分级管理图

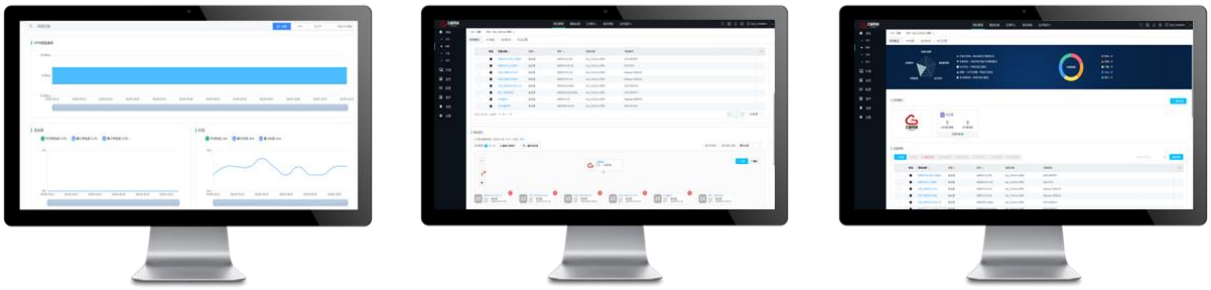


整网可视，透明管理

如图 2-8 所示，所有分支网络一站管控，各类型设备、网络健康状态、终端接入情况等数据可随时查看，同时支持对 VPN 状态、VPN 质量等进行实时监控，支持一键查看 VPN 相关信息。同时，云简广域网解决方案可以通过告警订阅功能提供多种告警机制，将告警信息实时推送至短信、微信、邮件等多个平台，提供自定义大屏，实现透明化的网络管理，提高运维效率。

图2-8 整网可视

VPN隧道质量可视　　组网拓扑可视　　告警实时监控



除此之外，云简广域网解决方案还支持提供专业化的 4G/5G 路由器管理能力，能够监控 4G/5G SIM 卡的基本信息、账号下所有 SIM 卡在线率统计以及基于每台设备 SIM 卡的信号质量趋势和流量趋势生成统计图表，如图 2-9 所示。

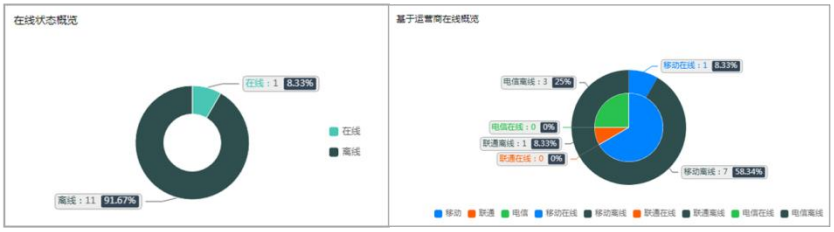
图2-9 4G/5G 路由器管理

4G/5G基本信息

IMSI码、所属运营商、在线时长、
信号强度、上下行流速、累计流量统计等

SIM卡在线统计

在线状态概览、在线数和离线数、
基于运营商在线概览、在线率趋势统计图



SIM卡信号质量

设备名称、IMSI、
信号质量、信号质量趋势统计图

SIM卡流量详情

设备名称、IMSI、总流量数据、
上下行流量数据、上下行流速、
流量趋势统计图、流速趋势统计图

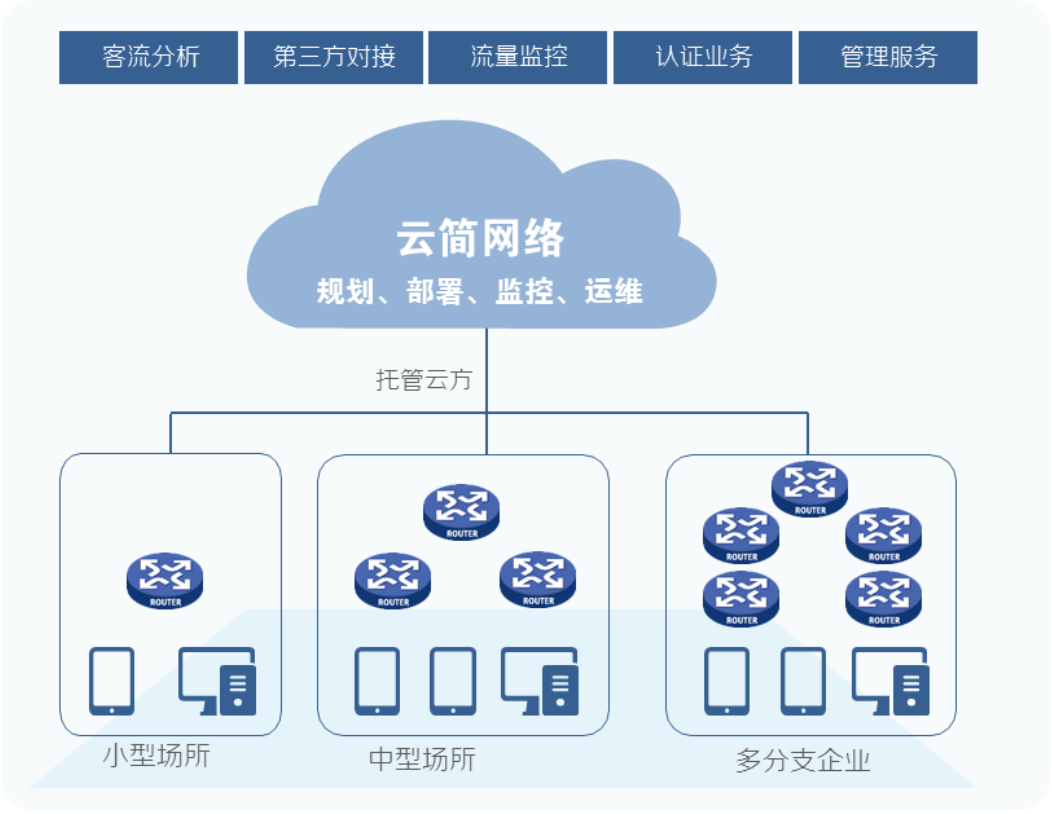
省——云管服务，智能运维

公有云服务，即时开通

如图 2-10 所示，云简广域网解决方案依托的云简网络是公有云方案，具有强大的 WiFi&IoT 融合能力和数据运营分析能力，提供客流分析、灵活认证和管理运维等服务，助力业务运营。

云简网络简化了用户网络运维的难度，用户无需自建机房、配置服务器，注册即可享用云简网络管理服务，极大地节省了部署时间和运维成本。

图2-10 云简网络示意图



智能运维，自动优化

如图 2-11 所示，云简网络平台支持全网设备与终端全实时问题扫描，支持云、网、端数据智能分析，依托于海量历史数据与进化算法，分析出问题根因并针对终端定制优化方案。无需人工干预，7×24 小时自动执行优化，能极大提高运维效率。

云简网络平台支持结合大数据和 AI 进行渐进优化，可根据不断变化的业务需求自动调整网络模型，如信道、功率、频宽等，让网络适配业务。闲时自动执行、温和渐进优化网络，网优不断网，能大幅提升用户体验感。

图2-11 智能运维

2.4 典型组网

单数据中心分支单线路组网

- 适用场景

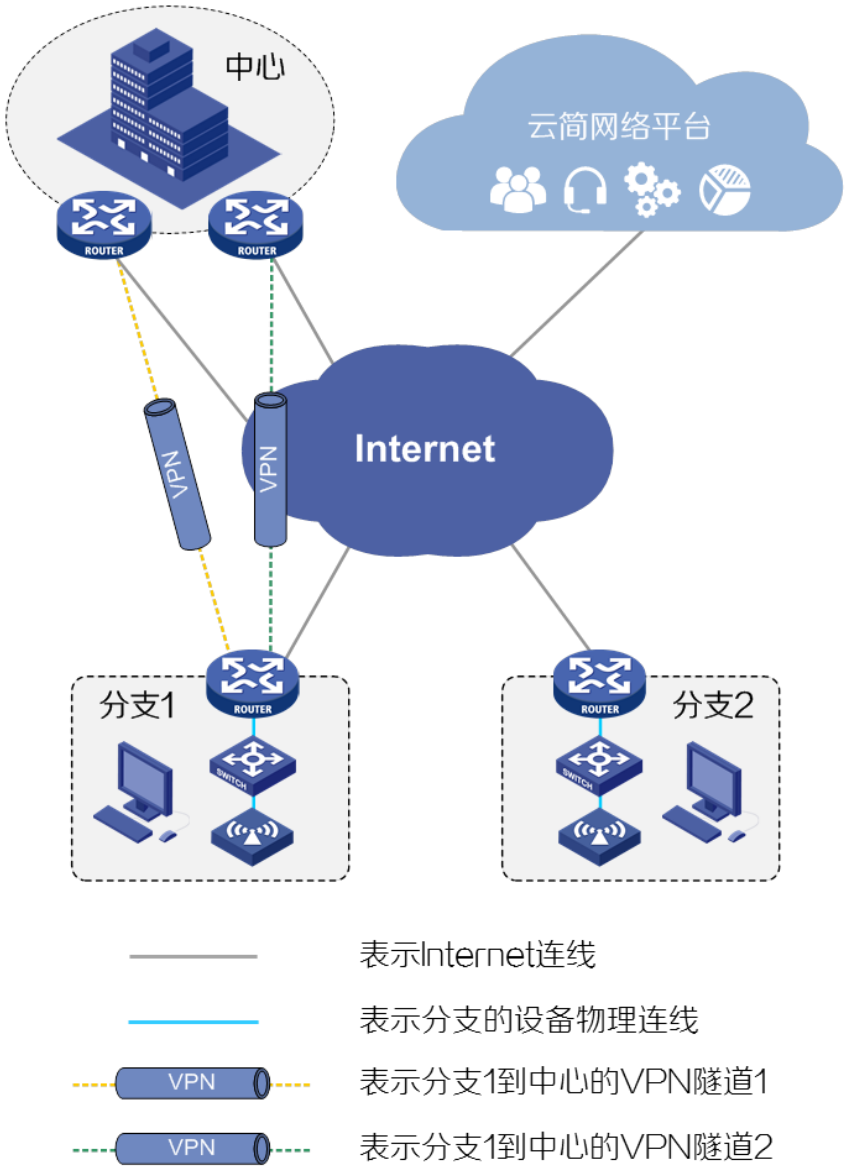
  适用于业务主要部署在总部，总部对数据有冗余备份要求的中小型网络场景，比如中小型企业、分支机构等。

- 特点
  - 每个分支设备和中心设备各自建立一条 VPN 隧道。
  - 中心两台 Hub 设备互为主备关系。

- 组网拓扑

  分支到中心采用 Hub-Spoke 组网，其中中心作为 Hub 站点，各分支作为 Spoke 站点。各分支使用一条有线链路接入中心，如图 2-12 所示。

图2-12 单数据中心分支单线路组网图



双数据中心分支双线路组网

- 适用场景

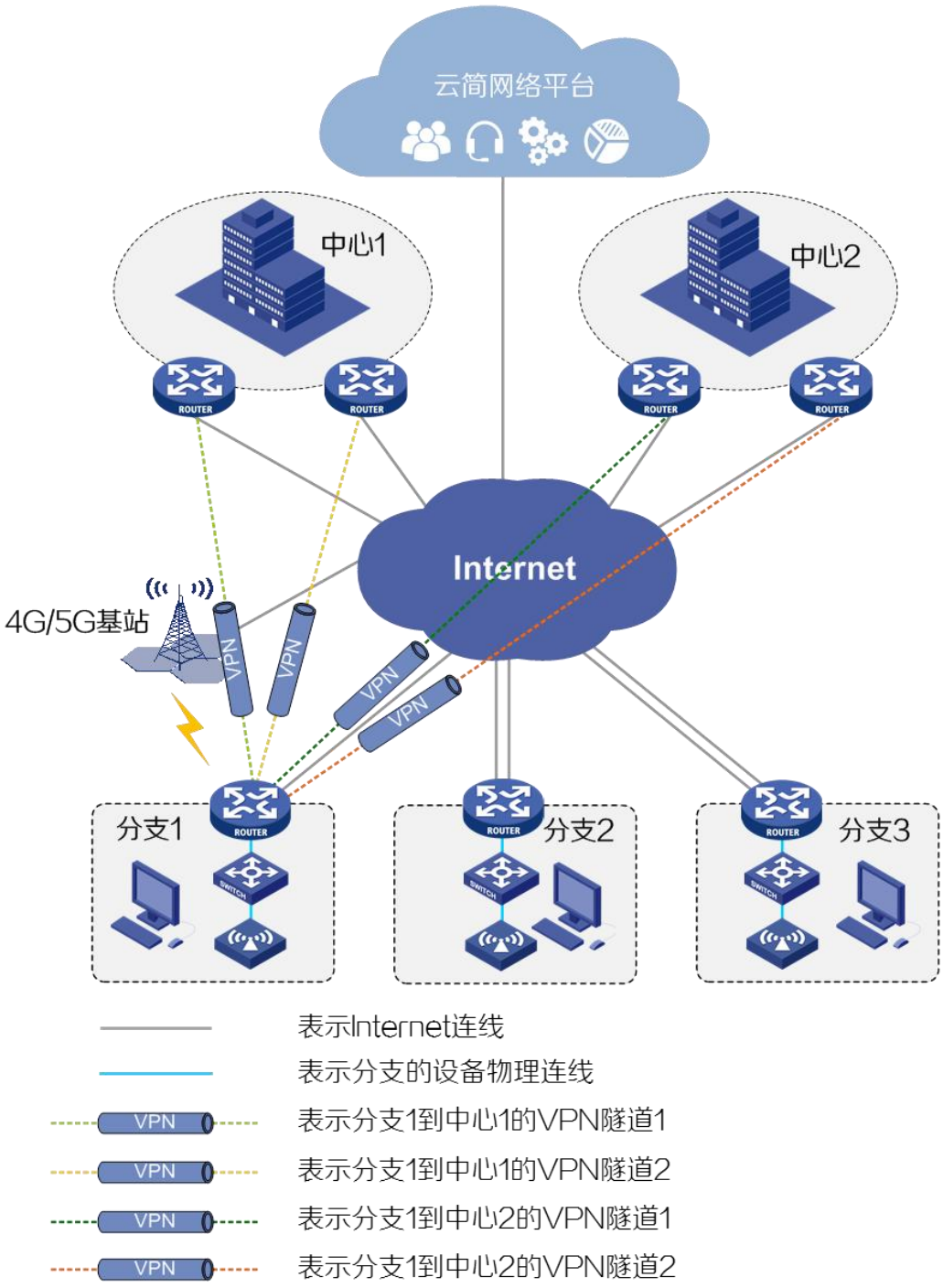
  适用于业务部署在多个总部，用户对数据可靠性要求较高的场景。

- 特点
  - 分支 1 分别使用一条有线链路和一条 4G/5G 备份链路接入中心，分支 2 和分支 3 均使用两条有线链路接入中心。
  - 当分支 1 的有线链路故障时，可以切换到 4G/5G 备份链路，不影响业务。
  - 每个中心两台 Hub 设备互为主备关系，分支 2 和分支 3 与每个中心设备建立的两条隧道也互为主备关系，当主隧道断开时，可以切换到备隧道，不影响业务。

- 组网拓扑

  分支到中心采用 Hub-Spoke 组网，其中中心作为 Hub 站点，各分支作为 Spoke 站点。分支 1 使用一条有线链路和一条 4G/5G 备份链路接入中心，分支 2 和分支 3 使用两条有线链路接入中心，如图 2-13 所示。

图2-13 双数据中心分支双线路组网图



2.5 成功实践

大型饮品公司

业务挑战

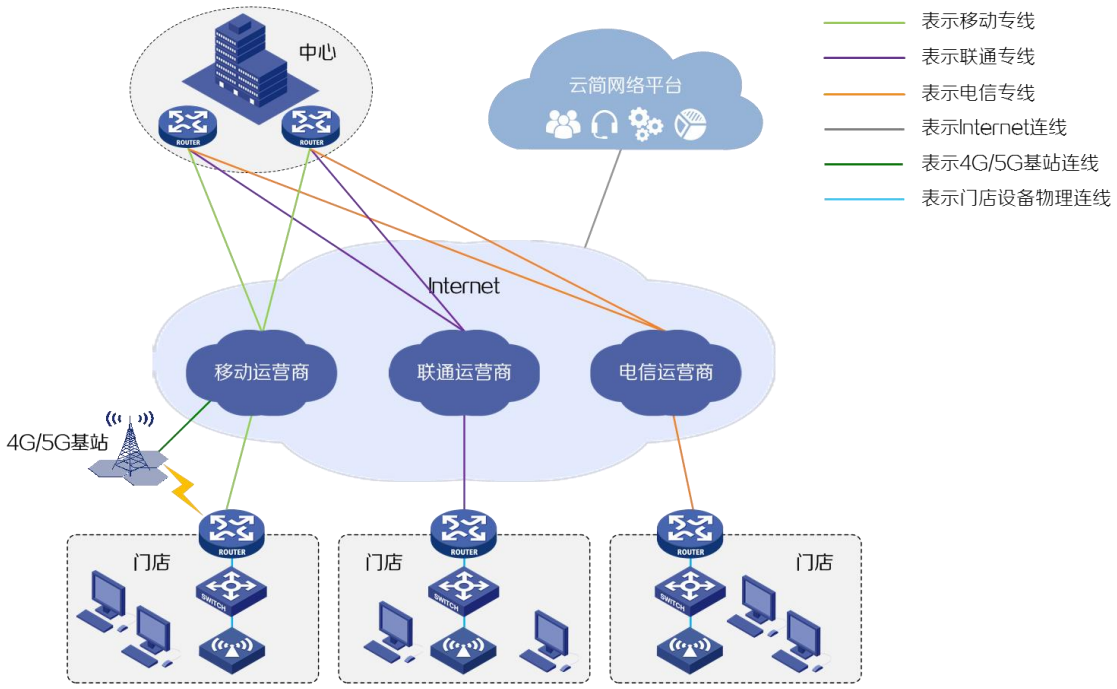
某大型饮品公司是全球连锁公司，在国内拥有几千家门店，业务覆盖地域广泛，网络是该公司生产业务运行的关键保障，因此对网络质量、可靠性、易管理性和需求响应及时度要求都非常高，当前该公司网络在运行和维护等方面面临如下问题：

- 设备亟需更新换代：现网设备硬件及软件版本老旧，设备已长期在线使用，亟需更新换代。
- 门店业务开通复杂：门店遍布全国各地，新业务开通部署难度大，设备开局繁琐，对现场实施人员技术要求高。
- 设备缺少统一运维管理：公司整体网络涉及设备类型众多，包括总部路由器、门店路由器、门店交换机、门店无线 AP 等，组网复杂，需要统一的设备状态监控和远程维护管理平台。
- VPN 部署复杂，质量不可视：总部与门店之间数据传输需要较高安全性，而安全隧道技术配置复杂，使用普通配置方式易出错，部署效率低。并且部署完成后无法监控链路质量和运行状态，影响业务稳定性。

组网方案

基于客户当前问题考虑，以简化开局部署难度、提高运维效率为目标，H3C 推出云简广域网解决方案，支持对整网路由器、交换机、AP、AC 等多种型号设备统一管理，同时通过图形化部署和快速建立 VPN 隧道，可以安全高效地完成业务传输。组网架构如图 2-14 所示。

图2-14 大型饮品公司云简广域网解决方案组网



- 公司总部部署两台高性能的 SR6600 系列路由器作为 VPN 的汇聚设备，并形成负载分担和主备。每台 SR6600 路由器各连接三根电信、联通、移动的互联网线路，保证全国各地门店均可同运营商接入总部；
- 各门店部署 MSR810 路由器、交换机和 AP 产品，满足门店内无线和有线终端接入需求，MSR810 路由器同时作为 AC 对门店的 AP 产品进行管理；
- 总部和门店的各型号网络设备统一接入 H3C 云简网络平台管理，通过云简网络平台进行业务部署和运维管理；
- 每个门店的 Spoke 设备采用单线路，分别与总部两台 Hub 设备建立 VPN 隧道，实现双隧道主备备份。同时可根据实际需求，增加备份有线线路或者 4G 线路，在每个门店灵活部署四隧道，提高备份可靠性。
- 为贴合公司门店的组织架构，在云简平台注册创建总部-大区-省市多级管理账号，实现分区域管理，省市管理员负责本地门店管理，快速灵活处理相关问题；总部/大区管理员负责整体网络监控。

客户价值

极简开局上线

H3C 云简网络平台支持 APP 扫码添加路由器，通过部署路由器，交换机、无线 AP 可以实现自动纳管上线。设备上线操作简单，降低了现场实施人员的操作难度，提高了开局效率。

VPN 易部署

简化 VPN 隧道配置，只需在中心 VPN、分支 VPN 页面进行简单配置即可建立总部与门店的 VPN 隧道。同时支持 VPN 隧道状态监控、隧道质量可视化、隧道通断实时告警等功能，用户可随时确认 VPN 业务状态。

组网高可靠性

双 Hub 组网，VPN 隧道实现主备备份。当承载业务隧道故障，能迅速切换到备份线路，保障用户业务稳定运行。

高效管理网络

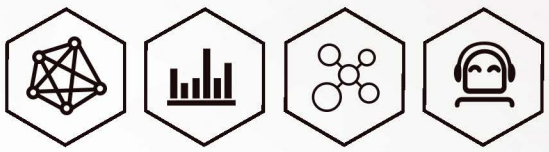
H3C 云简网络平台可统一纳管全网的有线和无线设备，支持设备运行状态监控、设备体检、终端数据报表、命令助手、个性化告警订阅等多种特色功能，满足运维管理需求。

轻松移动运维

使用与方案配套的 Cloudnet APP 可随时掌上监控管理网络、查看告警消息、使用 WiFi 检测等多种运维工具。

云端整网可视

各类型设备、网络健康状态、终端接入情况等数据可随时查看，还原每个关键节点的信息，运行故障快速告警，支持多种推送方式，实现透明化的网络管理，提高运维效率。



未 / 来 / 已 / 来

FURTURE
COME

中低端路由器产品 × 资料开发部联合出品