数据通信网络基础

前言

- 在人类社会的起源和发展过程中，通信就一直伴随着我们。从20世纪七、八十年代开始，人类社会已进入到信息时代，对于生活在信息时代的我们，通信的必要性更是不言而喻的。
- 本节课所说的通信，是指借助数据通信网络进行连接的通信。本课程主要介绍通信及数据通信网络的概念，信息传递的过程，网络设备及其作用，网络类型及典型组网，最后还会简要介绍网络工程和网络工程师的相关概念。

目标

- 学完本课程后，您将能够：
  - 区分网络通信和数据通信网络的概念
  - 描述信息传递的过程
  - 区分不同的网络设备并了解其基本作用
  - 认识不同的网络类型及拓扑类型
  - 了解网络工程与网络工程师的相关概念

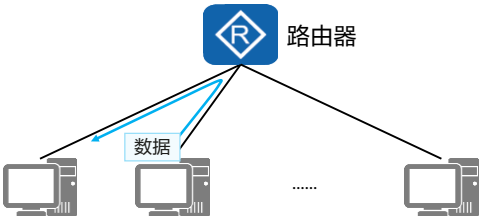
华为设备图标简介

目录

网络通信基本概念

- 通信，是指人与人、人与物、物与物之间通过某种媒介和行为进行的信息传递与交流。
- 网络通信，是指终端设备之间通过计算机网络进行的通信。
- 网络通信的例子：

A. 两台计算机（终端）之间通过网线传递文件



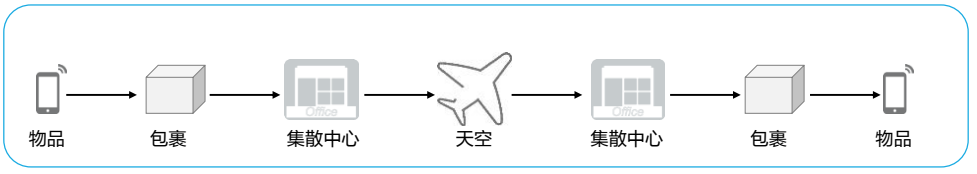
B. 多台计算机（终端）通过路由器传递文件

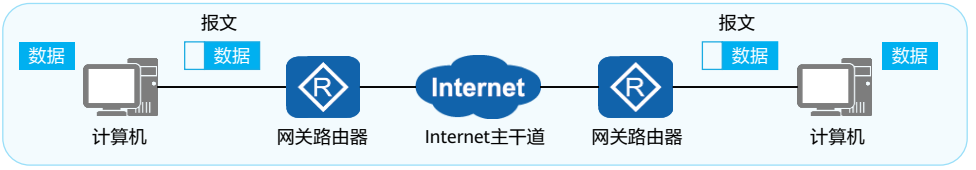
C. 计算机（终端）通过Internet下载文件

- 网络通信的例子：
  - A. 两台计算机通过一根网线相连，就组成了一个最简单的网络。
  - B. 由一台路由器（或交换机）和多台计算机设备组成的小型网络。在这样的网络中，通过路由器的中转，每两台计算机之间都可以自由地传递文件。
  - C. 当计算机想从某个网址获取文件时，必须先接入Internet，然后才能下载。
- Internet (译名：英特网、互联网、网际网等)，是目前世界上规模最大的计算机网络，其前身诞生于1969年的ARPAnet (Advanced Research Projects Agency Network)。Internet的广泛普及和应用是当今信息时代的标志性内容之一。

信息传递过程

- 虚拟的信息传递与真实的物品传递过程有许多相似之处。





  HUAWEI

- 快递过程与网络通信过程的对比：
- 需要快递的物品：
  - 应用程序生成需要传递的信息 (或数据)。
- 物品被包装起来形成包裹，并粘贴含有收货人姓名、地址的快递单：
  - 应用程序将数据打包成原始的“数据载荷”，并添加“头部”和“尾部”形成报文，报文中的重要信息是接收者的地址信息，即“目的地址”。
  - 在一个信息单元的基础上，增加一些新的信息段，使其形成一个新的信息单元，这个过程称为“封装”。
- 包裹被送到集散中心，集散中心对包裹上的目的地址进行分检，去往同一个城市的物品被放入同一架飞机，并飞向天空：
  - 报文通过网线到达“网关”，网关收到报文后，对其“解封装”，读取目的地址，再重新封装，并根据目的地址不同，送往不同的“路由器”，通过网关及路由器的传递，报文最终离开本地网络，进入Internet的干道进行传输。
  - 其中，网线所起的作用跟公路一样，它是信息传输的介质。
- 飞机抵达目的机场后，包裹被取出进行分检，去往同一地区的包裹，被送到了同一集散中心：
  - 报文经过Internet干道的传输，到达目的地址所在的本地网络，本地网络的网关或路由器对报文进行解封装和封装，并根据目的地址决定发往相应的下一台路由器，最终到达目的计算机所在网络的网关。
- 集散中心根据包裹上的目的地址进行分检，快递员送包裹上门，收件人拆开包裹，确认物品完好无损后收下。整个快递过程完成。
  - 报文到达目的计算机所在网络的网关，解封装和封装，然后根据目的地址发往相应的计算机。计算机收到报文后，对报文进行校验处理，校验无误后，接收下报文，并将其中的数据载荷交由相应的应用程序进行处理。一次完整的网络通信过程就结束了。

常见术语

术语	说明
数据载荷	最终想要传递的信息
报文	网络中交换与传输的数据单元
头部	在数据载荷的前面添加的信息段
尾部	在数据载荷的后面添加的信息段
封装	对数据载荷添加头部和尾部，形成新的报文的过程
解封装	去掉报文的头部和尾部，获取数据载荷的过程
网关	提供协议转换、路由选择、数据交换等功能的网络设备
路由器	为报文选择传递路径的网络设备
终端设备	数据通信系统的端设备，作为数据的发送者或接收者

- 数据载荷：可以理解为最终想要传递的信息，但实际上，在具有层次化结构的通信过程中，上一层协议传递给下一层协议的数据单元（报文）都可以称之为下一层协议的数据载荷。
- 报文：网络中交换与传输的数据单元，具有一定的内在格式，通常都具有头部+数据载荷+尾部的基本结构。传输过程中，报文的格式和内容可能发生改变。
- 头部：为了更好的传递信息，在组装报文时，在数据载荷的前面添加的信息段统称为报文的头部。
- 尾部：为了更好的传递信息，在组装报文时，在数据载荷的后面添加的信息段统称为报文的尾部。注意，很多报文是没有尾部的。
- 封装：分层协议所采用的一种技术，底层协议收到来自上层协议的消息时，将该消息附加到底层帧的数据部分。
- 解封装：是封装的逆过程，也就是去掉报文的头部和尾部，获取数据载荷的过程。
- 网关：是在采用不同体系结构或协议的网络之间进行互通时，用于提供协议转换、路由选择、数据交换等功能的网络设备。网关是一种根据其部署位置和功能而命名的术语，而不是一种特定的设备类型。
- 路由器：为报文选择传递路径的网络设备。
- 终端设备：数据通信系统的端设备，作为数据的发送者或接收者，提供用户接入协议操作所需必要功能，可以是计算机、服务器、VoIP、手机等。

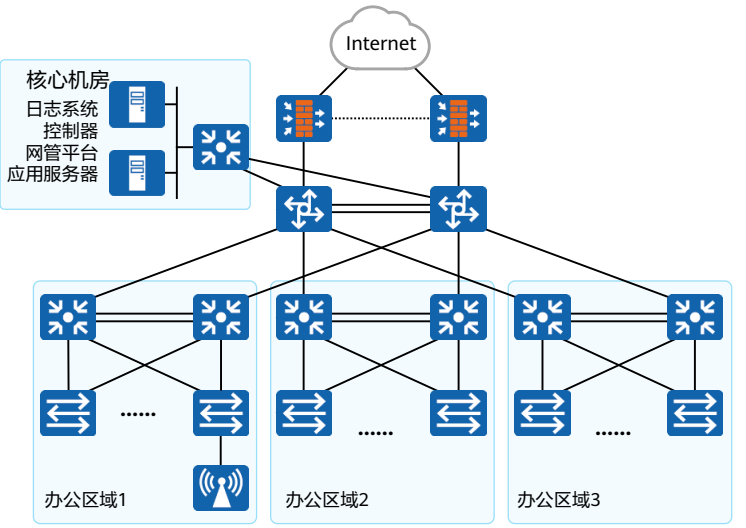
数据通信网络基本概念

- 数据通信网络：

  由路由器、交换机、防火墙、无线控制器、无线接入点，以及个人电脑、网络打印机、服务器等设备构成的通信网络。

- 功能：

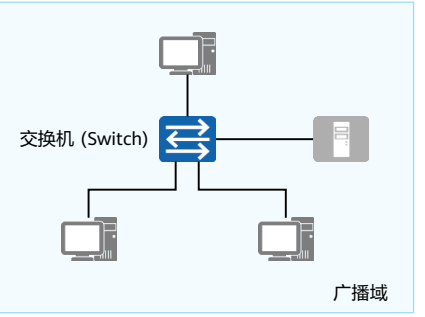
  数据通信网络最基本的功能是实现数据互通。



- 数据通信网络，Data Communication Network。

网络设备 - 交换机

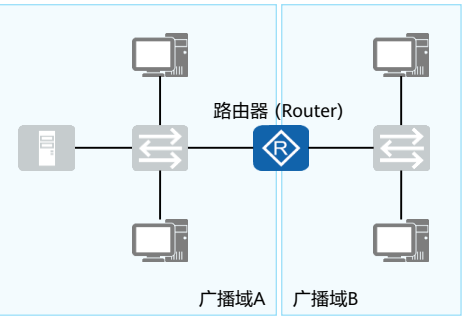
- 交换机：距离终端用户最近的设备，用于终端用户接入网络、对数据帧进行交换等。
  - 终端设备（PC、服务器等）的网络接入
  - 二层交换（Layer 2 Switching）



- 交换机：
  - 在园区网络中，交换机一般来说是距离终端用户最近的设备，接入层的交换机一般为二层交换机，又称为以太网交换机，二层是指TCP/IP参考模型的数据链路层；
  - 以太网交换机可以实现：数据帧的交换、终端用户设备的接入、基本的接入安全功能、二层链路的冗余等；
  - 广播域：一个节点发送一个广播报文其余节点都能够收到的节点的集合。

网络设备 - 路由器

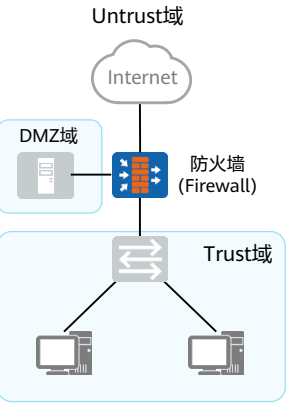
- 路由器：网络层设备，可以在因特网中进行数据报文转发。路由器根据所收到的报文的目的地址选择一条合适的路径，将报文传送到下一个路由器或目的地，路径中最后的路由器负责将报文送交目的主机。
  - 实现同类型网络或异种网络之间的通信
  - 隔离广播域
  - 维护路由表（Routing Table）、运行路由协议
  - 路径（路由信息）选择、IP报文转发
  - 广域网接入、网络地址转换
  - 连接通过交换机组建的二层网络



- 路由器：
  - 路由器工作在TCP/IP参考模型的网络层；
  - 路由器可以实现：维护路由表和路由信息、路由发现及路径选择、数据转发、隔离广播域、广域网接入和网络地址转换及特定的安全功能。

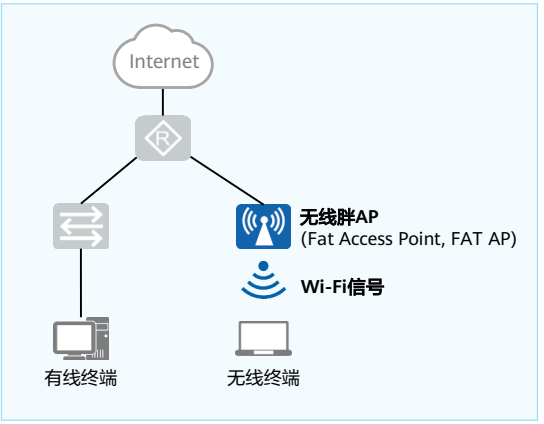
网络设备 - 防火墙

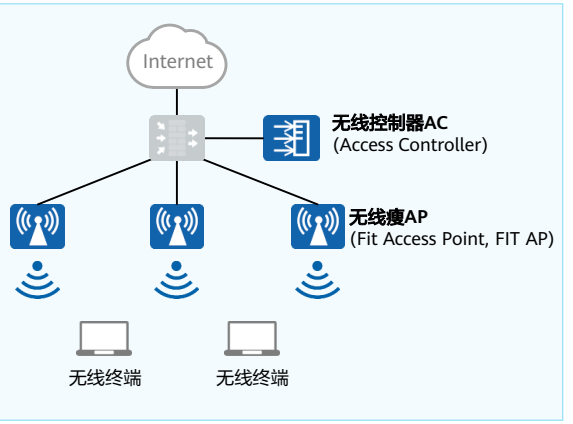
- 防火墙：网络安全设备，用于控制两个网络之间的安全通信。它通过监测、限制、更改跨越防火墙的数据流，尽可能地对外部屏蔽网络内部的信息、结构和运行状况，以此来实现对网络的安全保护。
  - 隔离不同安全级别的网络
  - 实现不同安全级别的网络之间的访问控制（安全策略）
  - 用户身份认证
  - 实现远程接入功能
  - 实现数据加密及虚拟专用网业务
  - 执行网络地址转换
  - 其他安全功能



- 防火墙：
  - 是位于两个信任程度不同的网络之间（如企业内部网络和Internet之间）的设备，它对两个网络之间的通信进行控制，通过强制实施统一的安全策略，防止对重要信息资源的非法存取和访问，以达到保护系统安全的目的。

网络设备 - 无线设备





- 无线局域网WLAN广义上是指以无线电波、激光、红外线等无线信号来代替有线局域网中的部分或全部传输介质所构成的网络。而常见的Wi-Fi是指IEEE 802.11标准上的无线局域网技术。
- 在WLAN中，常见的设备有胖AP、瘦AP和无线控制器AC。
  - 无线接入点 (AP, Access Point)
    - 一般支持FAT AP、FIT AP和云管理工作模式，根据网络规划的需求，可以灵活地在各种模式下切换。
    - FAT AP：适用于家庭，独立工作，需单独配置，功能较为单一，成本低。
    - FIT AP：适用于大中型企业，需要配合AC使用，由AC统一管理和配置，功能丰富。
    - 云管理：适用于中小型企业，需要配合云管理平台使用，由云管理平台统一管理和配置，功能丰富，即插即用。
  - 无线接入控制器 (AC, Access Controller)
    - 一般位于整个网络的汇聚层，提供高速、安全、可靠的WLAN业务。
    - 提供大容量、高性能、高可靠性、易安装、易维护的无线数据控制业务，具有组网灵活、绿色节能等优势。

目录

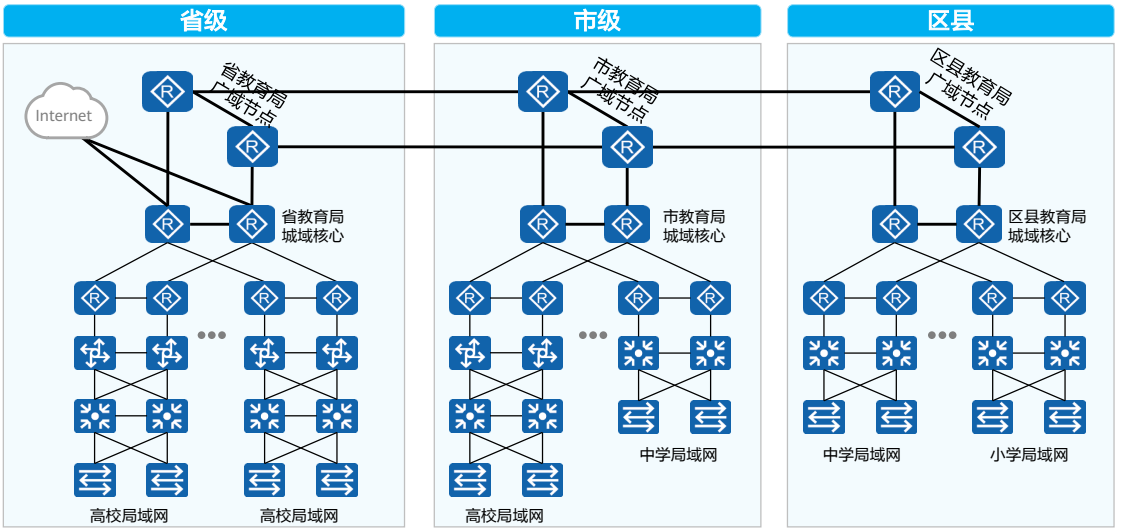
局域网、城域网、广域网

- 按照地理覆盖范围来划分，网络可以分为局域网（Local Area Network）、城域网（Metropolitan Area Network）和广域网（Wide Area Network）。
  - 局域网（LAN）：
    - 在某一地理区域内由计算机、服务器以及各种网络设备组成的网络。局域网的覆盖范围一般是方圆几千米以内。
    - 典型的局域网有：一家公司的办公网络，一个网吧的网络，一个家庭网络等。
  - 城域网（MAN）：
    - 在一个城市范围内所建立的计算机通信网络。
    - 典型的城域网有：宽带城域网、教育城域网、市级或省级电子政务专网等。
  - 广域网（WAN）：
    - 通常覆盖很大的地理范围，从几十公里到几千公里。它能连接多个城市甚至国家，并能提供远距离通信，形成国际性的大型网络。
    - 典型的广域网有：Internet（因特网）。

- 网络类型可以根据覆盖的地理范围，划分成局域网和广域网，以及介于局域网和广域网之间的城域网。
- 局域网：
  - 基本特点：
    - 覆盖范围一般在几公里之内；
    - 主要作用是把分布距离较近 (如: 有一个家庭内、一座或几座大楼内、一个校园内，等等) 的若干终端电脑连接起来。
  - 使用技术：以太网、Wi-Fi等。
- 城域网：
  - 基本特点：
    - 城域网是较大型的局域网，需要的成本较高，但可以提供更快的传输速率。它改进了局域网中的传输介质，扩大了局域网的访问范围，范围可以包含一个大学校园或城市；
    - 主要作用是将同一城市内不同地点的主机、数据库以及局域网等连接起来；
    - 与广域网作用相似，但实现方式和性能不同。
  - 使用技术：基于大型的局域网，与局域网技术相似，如：以太网 (10Gbps/100Gbps)、WiMAX (全球互通微波访问)。

- 广域网：
  - 基本特点：
    - 覆盖范围一般在几公里以上，可大至几十、几百或几千公里；
    - 主要作用是把分布较远 (如: 跨越城市、跨越国家， 等等) 的若干局域网或城域网连接起来；
    - 会用到电信运营商的通信线路。
  - 使用技术：HDLC、PPP等。

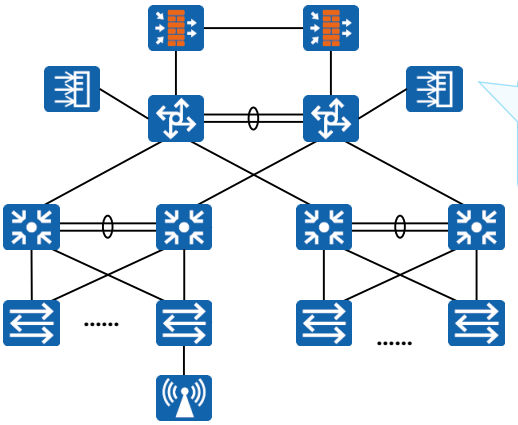
教育行业中的局域网、城域网及广域网



HUAWEI

网络拓扑

- 网络拓扑（Network Topology）是指用传输介质（例如双绞线、光纤等）互连各种设备（例如计算机终端、路由器、交换机等）所呈现的结构化布局。

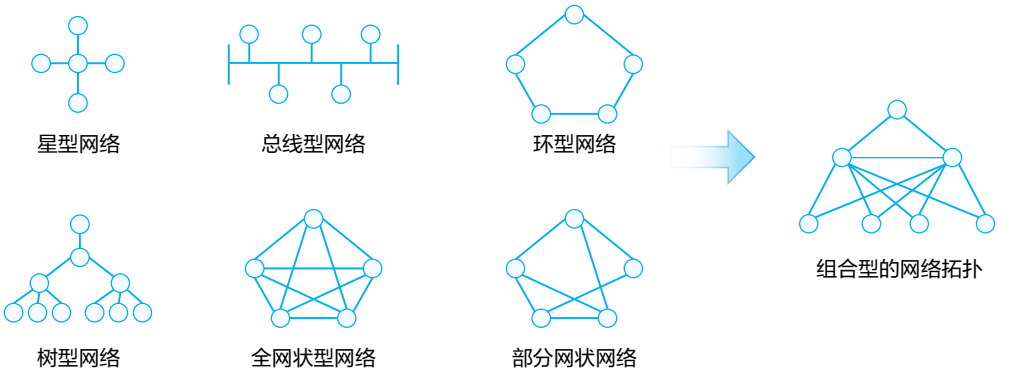


网络拓扑图是在网络工程领域用于描述网络的物理或逻辑结构，是一种非常重要的网络内容。

- 网络拓扑的绘制：
  - 掌握专业的网络拓扑图绘制技巧是非常重要的，这需要大量的练习。
  - Visio及Power Point是两种用于绘制网络拓扑图的常见工具。

网络拓扑形态

- 按照网络的拓扑形态来划分，网络可分为星型网络、总线型网络、环形网络、树形网络、全网状网络和部分网状网络。



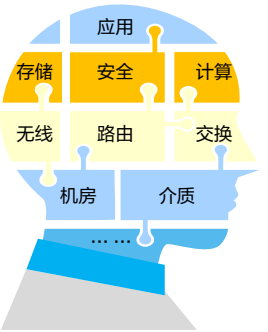
- 星型网络：
  - 所有节点通过一个中心节点连接在一起。
  - 优点：容易在网络中增加新的节点。通信数据必须经过中心节点中转，易于实现网络监控。
  - 缺点：中心节点的故障会影响到整个网络的通信。
- 总线型网络：
  - 所有节点通过一条总线（如同轴电缆）连接在一起。
  - 优点：安装简便，节省线缆。某一节点的故障一般不会影响到整个网络的通信。
  - 缺点：总线故障会影响到整个网络的通信。某一节点发出的信息可以被所有其他节点收到，安全性低。
- 环型网络：
  - 所有节点连成一个封闭的环形。
  - 优点：节省线缆。
  - 缺点：增加新的节点比较麻烦，必须先中断原来的环，才能插入新节点以形成新环。
- 树型网络：
  - 树型结构实际上是一种层次化的星型结构。
  - 优点：能够快速将多个星型网络连接在一起，易于扩充网络规模。
  - 缺点：层级越高的节点故障导致的网络问题越严重。

- 全网状网络：
  - 所有节点都通过线缆两两互联。
  - 优点：具有高可靠性和高通信效率。
  - 缺点：每个节点都需要大量的物理端口，同时还需要大量的互连线缆。成本高，不易扩展。
- 部分网状网络：
  - 只是重点节点之间才两两互连。
  - 优点：成本低于全网状网络。
  - 缺点：可靠性比全网状网络有所降低。
- 在实际组网中，通常都会根据成本、通信效率、可靠性等具体需求而采用多种拓扑形态相结合的方法。

目录

网络工程

- 网络工程：
  - 在信息系统工程方法和完善的组织机构指导下，根据网络应用的需求，按照计算机网络系统的标准、规范和技术，规划设计可行性方案，将计算机网络硬件设备、软件和技术系统地集成在一起，以成为满足用户需求、高性价比的网络系统的组建工作。
- 网络工程所涵盖的技术模块：



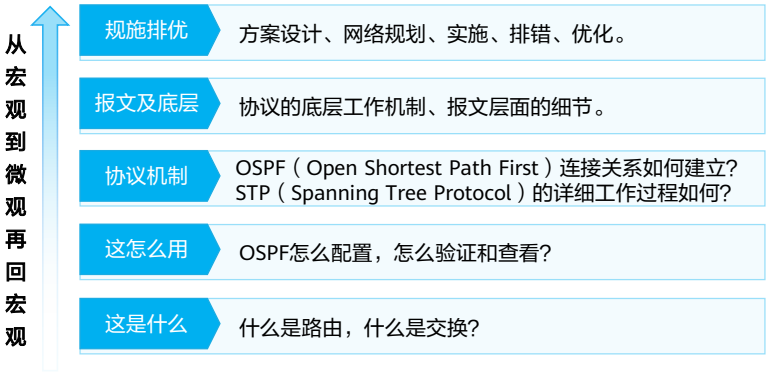
HUAWEI

- 网络工程，就是围绕着网络进行的一系列的活动，包括：网络规划、设计、实施、调试、排错等。
- 网络工程设计的知识领域很宽广，其中路由和交换是计算机网络的基本。

网络工程师

- 网络工程师：
  - 是在网络工程领域，掌握专业的网络技术，具备一定的职业技能及职业素养，具有一定项目实施经验，能够在项目现场与客户或者其他项目干系人充分沟通，根据客户的需求及环境因素制定实施方案及项目计划（得到项目干系人认可），并充分调动各方资源保证项目按时、保质保量落地，以及在项目实施后对干系人进行培训及工程文档交付的职业。
- 网络工程师综合能力模型：

网络工程师的技术成长之路



华为认证，为企业人才培养注入活力

认证考试

- 对接行业，培养既懂技术又懂业务的“行家”。
- 基于华为云，培养平台建设与服务应用专家。
- 聚焦ICT基础设施，培养全技术领域架构人才。

提供人才成长路径

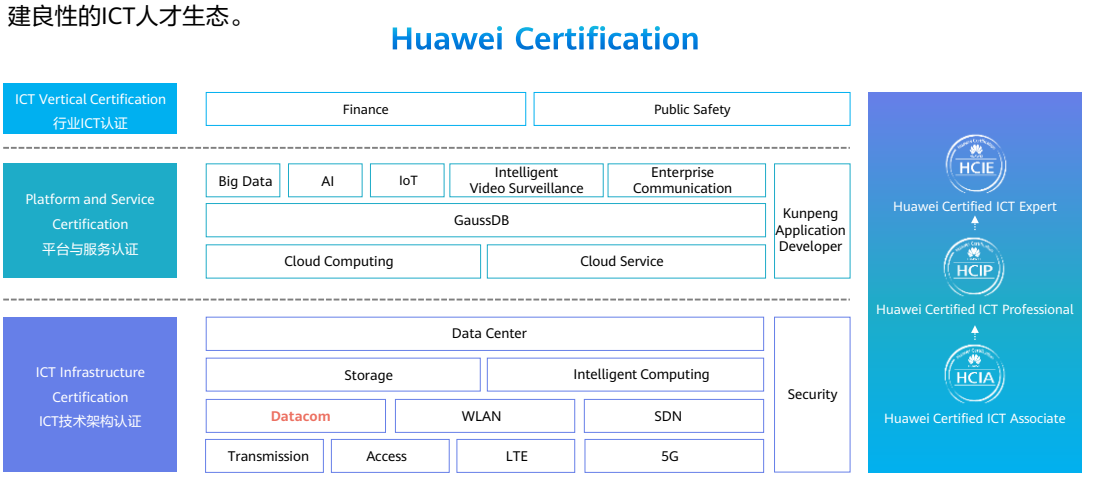
- 满足企业人才不断进阶的职业角色演变：工程师->高级工程师->专家。
- 层次化的认证进阶设计，适配岗位能力要求，可专业纵深，可融合扩展，提供可定制的人才成长路径，缩减企业人才培养成本。

助力企业创新与转型

- 经过权威认证的ICT人才，可保证企业项目交付质量，促进客户满意度提升。
- 提升企业整体绩效与生产力。
- 加快业务创新转型，实现运营效率整体提升。

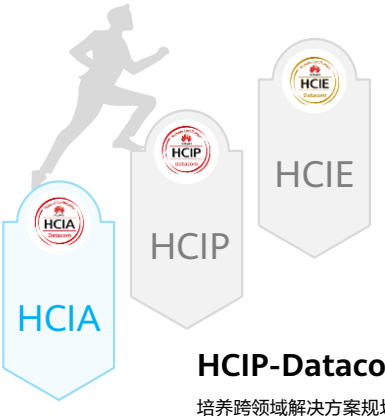
华为认证体系

- 华为认证覆盖ICT全领域，致力于提供领先的人才培养体系和认证标准，培养数字化时代的新型ICT人才，构建良性的ICT人才生态。



- 华为人才生态网站：https://e.huawei.com/cn/talent/#/home

华为数通认证进阶路径



HCIE-Datacom

培养具备跨领域解决方案坚实理论及部署能力的网络专家

HCIA-Datacom

培养具备数通基础理论技能的网络工程师

HCIP-Datacom

培养跨领域解决方案规划设计或单领域规划及部署的网络高级工程师

- HCIA-Datacom：一门课程（考试）
  - 数据通信基本概念、路由交换技术基础、安全、WLAN，SDN与NFV，编程自动化基础、网络部署案例
- HCIP-Datacom：一门必选课程（考试），六门任选子认证课程（考试）
  - 必选课程（考试）：
    - HCIP-Datacom-Core Technology
  - 任选课程（考试）：
    - HCIP-Datacom-Advanced Routing & Switching Technology
    - HCIP-Datacom-Campus Network Planning and Deployment
    - HCIP-Datacom-Enterprise Network Solution Design
    - HCIP-Datacom-WAN Planning and Deployment
    - HCIP-Datacom-SD-WAN Planning and Deployment
    - HCIP-Datacom-Network Automation Developer
- HCIE-Datacom：一门课程（考试），融合两大模块
  - 经典网络：
    - 基于命令行的经典数通技术理论
    - 基于命令行的经典数通技术部署
  - 华为SDN解决方案：
    - 企业SDN解决方案技术理论
    - 企业SDN解决方案规划部署

思考题

1. （单选）以下哪种类型的网络具有最高的可靠性？（　　）

   A. 星型网络

   B. 环型网络

   C. 全网状网络

   D. 树型网络

1. C

本章总结

- 在本章节中，介绍了网络通信和数据通信网络的概念，数据通信网络最基本的功能就是实现网络通信。
- 还介绍了各种网络设备，认识了局域网、城域网和广域网的区别，并且介绍了各种网络拓扑，在实际组网中，通常都会根据多方需求而采用多种拓扑形态相结合的方法。
- 最后，介绍了网络工程和网络工程师，并且介绍了华为数通认证进阶路径。


谢 谢
www.huawei.com

网络参考模型

前言

- 数字化时代，各种信息以数据的形式充斥着我们的生活。什么是数据？数据又是如何传递的?
- 本章我们将通过网络参考模型去简单了解数据的“一生”。

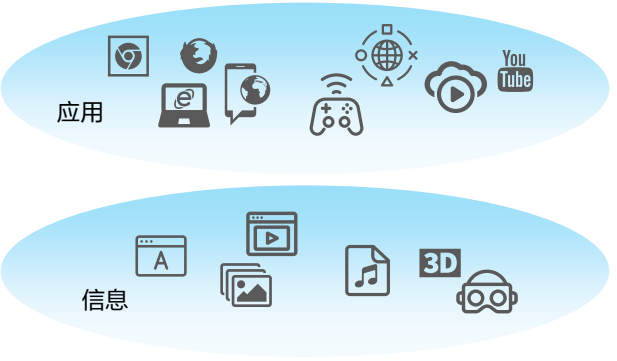
目标

- 学完本课程后，您将能够：
  - 理解数据的定义及传递过程
  - 理解网络参考模型概念及优势
  - 了解常见的标准协议
  - 掌握数据封装与解封装过程

目录

故事的起源 - 应用

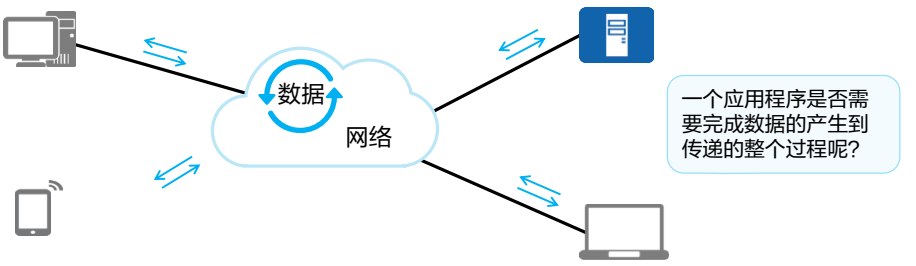
- 应用的存在，是为了满足人们的各种需求，比如访问网页，在线游戏，在线视频等。
- 伴随着应用会有信息的产生。比如文本，图片，视频等都是信息的不同呈现方式。



HUAWEI

应用的实现 - 数据

- 数据的产生
  - 在计算机领域，数据是各种信息的载体。
- 数据传输
  - 大部分应用程序所产生的数据需要在不同的设备之间传递。



HUAWEI

- 计算机只能识别0和1的组成的电子数据(digital data)。它不具备读取各种信息的能力，所以信息需要通过一定的规则翻译成数据。
- 而对人来说，我们不具备读取电子数据的能力，所以在读取信息的时候，需要将数据转成人能理解的信息。
- 对于一名网络工程师来说，需要更关注数据的端到端传递的过程。

目录

OSI参考模型

7. 应用层	对应用程序提供接口。
6. 表示层	进行数据格式的转换，以确保一个系统生成的应用层数据能够被另外一个系统的应用层所识别和理解。
5. 会话层	在通信双方之间建立、管理和终止会话。
4. 传输层	建立、维护和取消一次端到端的数据传输过程。控制传输节奏的快慢，调整数据的排序等等。
3. 网络层	定义逻辑地址；实现数据从源到目的地的转发。
2. 数据链路层	将分组数据封装成帧；在数据链路上实现数据的点到点、或点到多点方式的直接通信；差错检测。
1. 物理层	在媒介上传输比特流；提供机械的和电气的规约。

- OSI 模型(Open Systems Interconnection Model)，由国际化标准组织ISO (The International Organization for Standardization) 收录在ISO 7489标准中并于1984年发布。
- OSI参考模型又被称为七层模型，由下至上依次为：
  - 物理层：在设备之间传输比特流，规定了电平、速度和电缆针脚等物理特性。
  - 数据链路层：将比特组合成字节，再将字节组合成帧，使用链路层地址（以太网使用MAC地址）来访问介质，并进行差错检测。
  - 网络层：定义逻辑地址，供路由器确定路径，负责将数据从源网络传输到目的网络。
  - 传输层：提供面向连接或非面向连接的数据传递以及进行重传前的差错检测。
  - 会话层：负责建立、管理和终止表示层实体之间的通信会话。该层的通信由不同设备中的应用程序之间的服务请求和响应组成。
  - 表示层：提供各种用于应用层数据的编码和转换功能，确保一个系统的应用层发送的数据能被另一个系统的应用层识别。
  - 应用层：OSI参考模型中最靠近用户的一层，为应用程序提供网络服务。

TCP/IP参考模型

- 因为OSI协议栈比较复杂，且TCP和IP两大协议在业界被广泛使用，所以TCP/IP参考模型成为了互联网的主流参考模型。

HUAWEI

- TCP/IP模型在结构上与OSI模型类似，采用分层架构，同时层与层之间联系紧密。
- TCP/IP标准参考模型将OSI中的数据链路层和物理层合并为网络接入层，这种划分方式其实是有悖于现实协议制定情况的，故融合了TCP/IP标准模型和OSI模型的TCP/IP对等模型被提出，后面的讲解也都将基于这种模型。

TCP/IP常见协议

- TCP/IP协议栈定义了一系列的标准协议。

<table>
<tr><td rowspan="2">应用层</td><td>Telnet</td><td>FTP</td><td>TFTP</td><td>SNMP</td></tr>
<tr><td>HTTP</td><td>SMTP</td><td>DNS</td><td>DHCP</td></tr>
<tr><td>传输层</td><td colspan="2">TCP</td><td colspan="2">UDP</td></tr>
<tr><td rowspan="2">网络层</td><td colspan="2">ICMP</td><td colspan="2">IGMP</td></tr>
<tr><td colspan="4">IP</td></tr>
<tr><td rowspan="2">数据链路层</td><td colspan="4">PPPoE</td></tr>
<tr><td colspan="2">Ethernet</td><td colspan="2">PPP</td></tr>
<tr><td>物理层</td><td colspan="4">``````</td></tr>
</table>

- 应用层
  - HTTP（Hypertext Transfer Protocol，超文本传输协议）：用来访问在网页服务器上的各种页面。
  - FTP（File Transfer Protocol，文件传输协议）：为文件传输提供了途径，它允许数据从一台主机传送到另一台主机上。
  - DNS（Domain Name Service，域名称解析服务）：用于实现从主机域名到IP地址之间的转换。
- 传输层
  - TCP（Transmission Control Protocol，传输控制协议）：为应用程序提供可靠的面向连接的通信服务。目前，许多流行的应用程序都使用TCP。
  - UDP（User Datagram Protocol，用户数据报协议）：提供了无连接通信，且不对传送数据包进行可靠性的保证。
- 网络层
  - IP（Internet Protocol，互联网协议）：将传输层的数据封装成数据包并完成源站点到目的站点的转发，提供无连接的、不可靠的服务。
  - IGMP（Internet Group Management Protocol，因特网组管理协议）：负责IP组播成员管理的协议。它用来在IP主机和与其直接相邻的组播路由器之间建立、维护组播组成员关系。
  - ICMP（Internet Control Message Protocol，网际报文控制协议）：基于IP协议在网络中发送控制消息，提供可能发生在通信环境中的各种问题反馈。通过这些信息，使管理者可以对所发生的问题作出诊断，然后采取适当的措施解决。

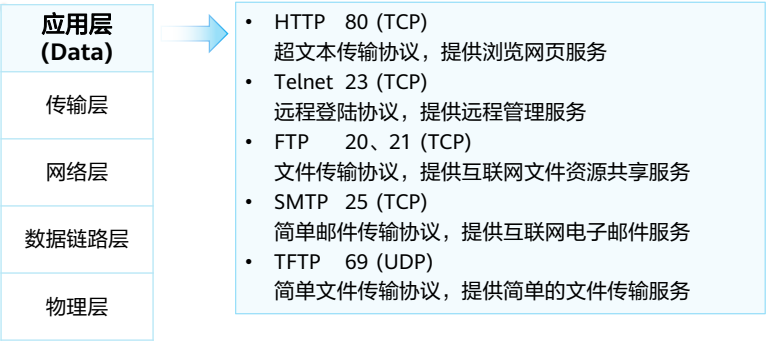
- 数据链路层
  - PPP（ Point-to-Point Protocol，点对点协议）：一种点对点模式的数据链路层协议，多用于广域网。
  - Ethernet(以太网协议)：一种多路访问广播型数据链路层协议，是当前应用最为广泛的局域网技术。
  - PPPoE（ Point-to-Point Protocol over Ethernet，以太网承载PPP协议）：PPPoE提供通过简单桥接访问设备（接入设备）把一个网络的多个主机连接到远程访问集中器的功能。常见的应用有家庭宽带拨号上网。

常见协议标准化组织

- IETF(Internet Engineering Task Force)
  - 负责开发和推广互联网协议（特别是构成TCP/IP协议族的协议）的志愿组织，通过RFC发布新的或者取代老的协议标准。
- IEEE(Institute of Electrical and Electronics Engineers)
  - IEEE制定了全世界电子、电气和计算机科学领域30%左右的标准，比较知名的有IEEE802.3(Ethernet)、IEEE802.11(Wi-Fi)等。
- ISO(International Organization for Standardization)
  - 在制定计算机网络标准方面，ISO是起着重大作用的国际组织，如OSI模型，定义于ISO/IEC 7498-1。

应用层

- 应用层为应用软件提供接口，使应用程序能够使用网络服务。应用层协议会指定使用相应的传输层协议，以及传输层所使用的端口等。
- 应用层的PDU被称为Data（数据）。



- TCP/IP每一层都让数据得以通过网络进行传输，这些层之间使用PDU（Packet Data Unit，协议数据单元）彼此交换信息，确保网络设备之间能够通信。
- 不同层的PDU中包含有不同的信息，因此PDU在不同层被赋予了不同的名称。

常见应用层协议 - FTP

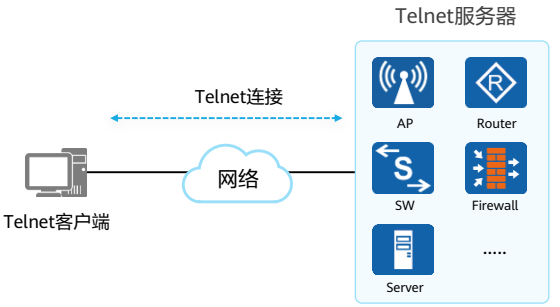
- FTP（File Transfer Protocol）是一个用于从一台主机传送文件到另一台主机的协议，用于文件的“下载”和“上传”，它采用C/S（Client/Server）结构。

FTP客户端：提供本地设备对远程服务器的文件进行操作的命令。用户在PC上通过应用程序作为FTP Client，并与FTP服务器建立连接后，可以对FTP Server上的文件进行操作。

FTP服务器：运行FTP服务的设备。提供远程客户端访问和操作的功能，用户可以通过FTP客户端程序登录到服务器上，访问设备上的文件。

常见应用层协议 - Telnet

- Telnet是数据网络中提供远程登录服务的标准协议。 Telnet为用户提供了在本地计算机上完成远程设备工作的能力。



用户通过Telnet客户端程序连接到Telnet服务器。用户在Telnet客户端中输入命令，这些命令会在服务器端运行，就像直接在服务端的控制台上输入一样。

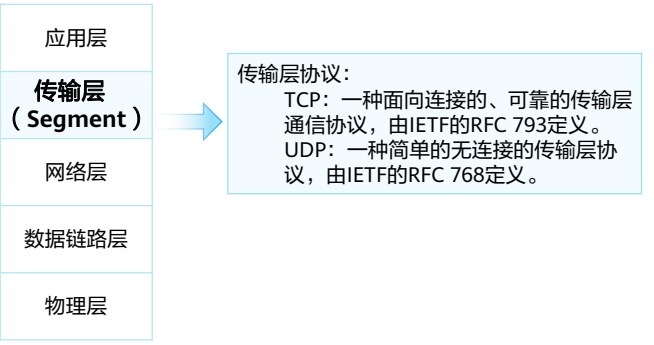
常见应用层协议 - HTTP

- HTTP（HyperText Transfer Protocol）是互联网上应用最为广泛的一种网络协议。设计HTTP最初的目的是为了提供一种发布和接收HTML页面的方法。

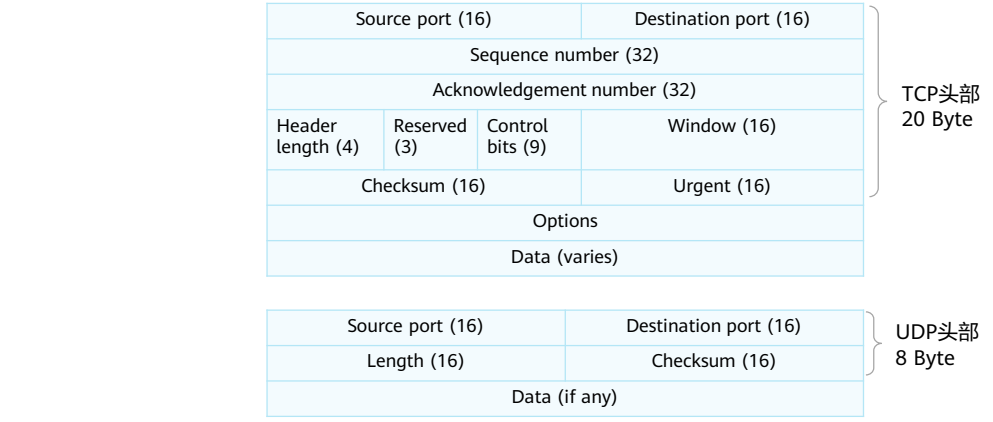
HUAWEI

传输层

- 传输层协议接收来自应用层协议的数据，封装上相应的传输层头部，帮助其建立“端到端”（Port to Port）的连接。
- 传输层的PDU被称为Segment（段）。



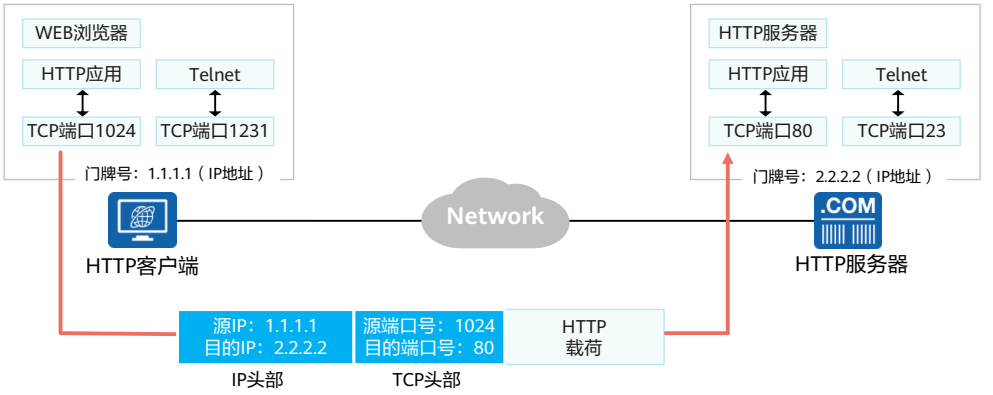
TCP和UDP – 报文格式



- TCP报文头部：
  - Source Port：源端口，标识哪个应用程序发送。长度为16比特。
  - Destination Port：目的端口，标识哪个应用程序接收。长度为16比特。
  - Sequence Number：序号字段。TCP链接中传输的数据流每个字节都编上一个序号。序号字段的值指的是本报文段所发送数据的第一个字节的序号。长度为32比特。
  - Acknowledgment Number：确认序列号，是期望收到对方下一个报文段数据的第1个字节的序号，即上次已成功接收到的数据段的最后一个字节数据的序号加1。只有Ack标识为1，此字段有效。长度为32比特。
  - Header Length：头部长度，指出TCP报文头部长度，以32比特（4字节）为计算单位。若无选项内容，则该字段为5，即头部为20字节。
  - Reserved：保留，必须填0。长度为3比特。
  - Control bits：控制位，包含FIN、ACK、SYN等标志位，代表不同状态下的TCP数据段。
  - Window：窗口TCP的流量控制，这个值表明当前接收端可接受的最大的数据总数（以字节为单位）。窗口最大为65535字节。长度为16比特。
  - Checksum：校验字段，是一个强制性的字段，由发端计算和存储，并由收端进行验证。在计算检验和时，要包括TCP头部和TCP数据，同时在TCP报文段的前面加上12字节的伪头部。长度为16比特。

    - Urgent:紧急指针，只有当URG标志置1时紧急指针才有效。TCP的紧急方式是发送端向另一端发送紧急数据的一种方式。紧急指针指出在本报文段中紧急数据共有多少个字节（紧急数据放在本报文段数据的最前面）。长度为16比特。
    - Options：选项字段（可选），长度为0-40字节。
- UDP报文头部：
    - Source Port:源端口，标识哪个应用程序发送。长度为16比特。
    - Destination Port:目的端口，标识哪个应用程序接收。长度为16比特。
    - Length:该字段指定UDP报头和数据总共占用的长度。可能的最小长度是8字节，因为UDP报头已经占用了8字节。由于这个字段的存在，UDP报文总长不可能超过65535字节（包括8字节的报头，和65527字节的数据）。
    - Checksum:覆盖UDP头部和UDP数据的校验和，长度为16比特。

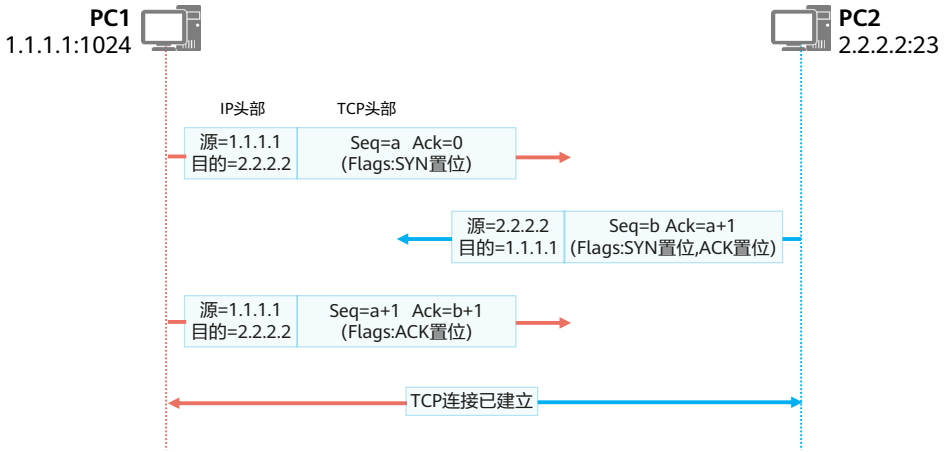
TCP和UDP – 端口号



- 客户端使用的源端口一般随机分配，目标端口则由服务器的应用指定；
- 源端口号一般为系统中未使用的，且大于1023；
- 目的端口号为服务端开启的应用（服务）所侦听的端口，如HTTP缺省使用80。

TCP的建立 - 三次握手

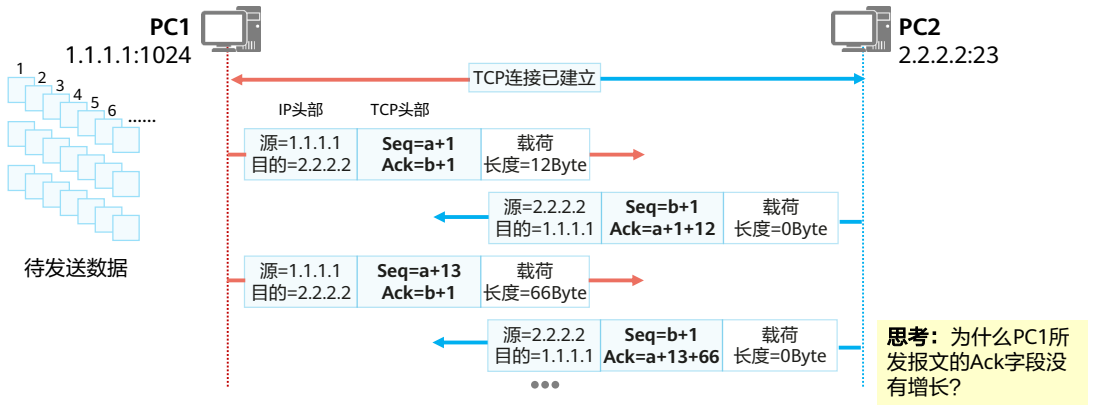
- 任何基于TCP的应用，在发送数据之前，都需要由TCP进行“三次握手”建立连接。



- TCP连接建立的详细过程如下：
  - 由TCP连接发起方（图中PC1），发送第一个SYN位置1的TCP报文。初始序列号a为一个随机生成的数字，因为没收到过来自PC2的任何报文，所以确认序列号为0；
  - 接收方（图中PC2）接收到合法的SYN报文之后，回复一个SYN和ACK置1的TCP报文。初始序列号b为一个随机生成的数字，同时因为此报文是回复给PC1的报文，所以确认序列号为a+1；
  - PC1接收到PC2发送的SYN和ACK置位的TCP报文后，回复一个ACK置位的报文，此时序列号为a+1,确认序列号为b+1。PC2收到之后，TCP双向连接建立。

TCP的序列号与确认序列号

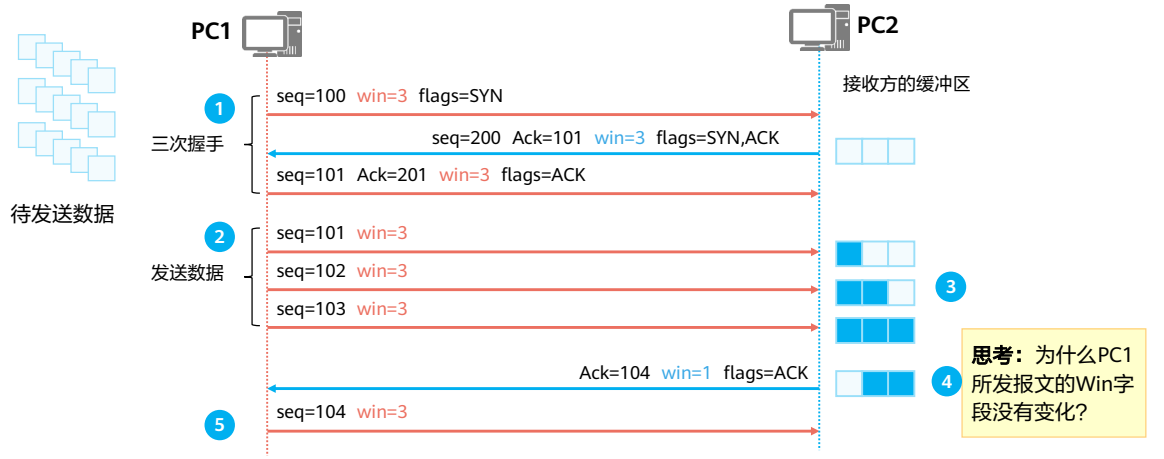
- TCP使用序列号和确认序列号字段实现数据的可靠和有序传输。



- 假设PC1要给PC2发送一段数据，传输过程如下：
  - 1. PC1将全部待TCP发送的数据按照字节为单位编上号。假设第一个字节的编号为“a+1”，第二个字节的序号为“a+2”，依次类推。
  - 2. PC1会把每一段数据的第一个字节的编号作为序列号（Sequence number），然后将TCP报文发送出去。
  - 3. PC2在收到PC1发送来的TCP报文后，需要给予确认同时请求下一段数据，如何确定下一段数据呢？序列号(a+1)+载荷长度=下一段数据的第一个字节的序号（a+1+12）
  - 4. PC1在收到PC2发送的TCP报文之后，发现确认序列号为“a+1+12”，说明“a+1”到“a+12”这一段的数据已经被接受，需要从“a+1+12”开始发送。
- 为了提升发送效率，也可以一次性发送多段数据，由接收方统一确认。

TCP的窗口滑动机制

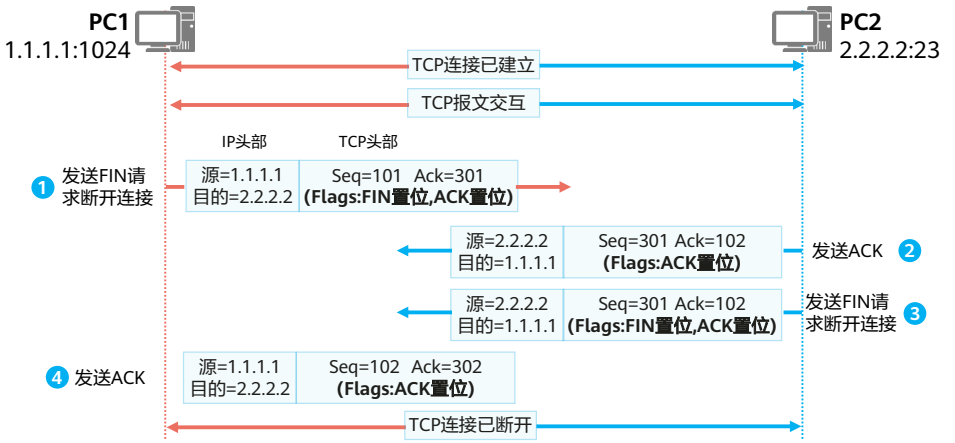
- TCP通过滑动窗口机制来控制数据的传输速率。



- 1. 在TCP三次握手建立连接时，双方都会通过Window字段告诉对方本端最大能够接受的字节数（也就是缓冲区大小）。
- 2. 连接建立成功之后，发送方会根据接受方宣告的Window大小发送相应字节数的数据。
- 3. 接受方接受到数据之后会放在缓冲区内，等待上层应用来取走缓冲的数据。若数据被上层取走，则相应的缓冲空间将被释放。
- 4. 接收方根据自身的缓存空间大小通告当前的可以接受的数据大小(Window)。
- 5. 发送方根据接收方当前的Window大小发送相应数量的数据。

TCP的关闭 - 四次挥手

- 当数据传输完成，TCP需要通过“四次挥手”机制断开TCP连接，释放系统资源。

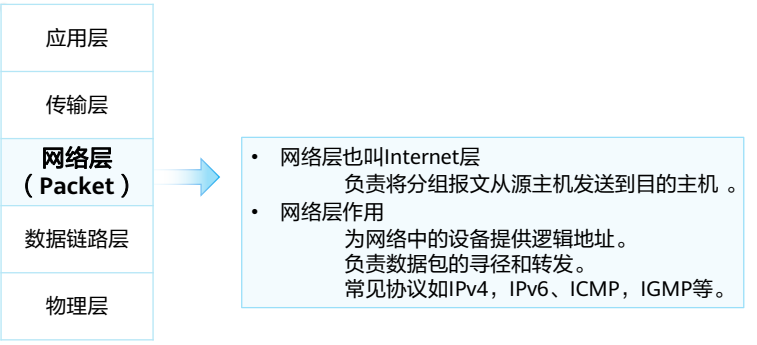


  HUAWEI

- TCP支持全双工模式传输数据，这意味着同一时刻两个方向都可以进行数据的传输。在传输数据之前，TCP通过三次握手建立的实际上是两个方向的连接，因此在传输完毕后，两个方向的连接必须都关闭。如图所示：
  - 1. 由PC1发出一个FIN字段置”1 ”的不带数据的TCP段；
  - 2. PC2收到PC1发来的FIN置位的TCP报文后，会回复一个ACK置位的TCP报文。
  - 3. 若PC2也没有需要发送的数据，则直接发送FIN置位的TCP报文。假设此时PC2还有数据要发送，那么当PC2发送完这些数据之后会发送一个FIN置位的TCP报文去关闭连接。
  - 4. PC1收到FIN置位的TCP报文，回复ACK报文，TCP双向连接断开。

网络层

- 传输层负责建立主机之间进程与进程之间的连接，而网络层则负责数据从一台主机到另外一台主机之间的传递。
- 网络层的PDU被称为Packet（包）。



- IPv4(Internet Protocol Version 4)，简称IP，是目前应用最广泛的网络层协议。

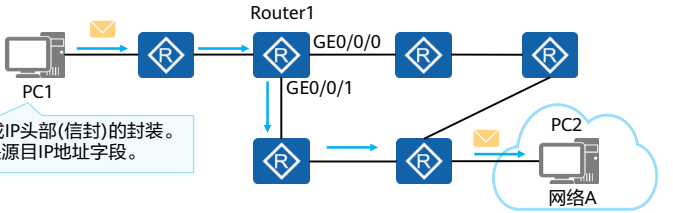
网络层协议工作过程

数据包的封装过程

信件：上层（例如传输层）提交的数据

信封：IP报文头部
发件人：源IP地址
收件人：目的IP地址

基于网络层地址的报文转发过程



PC完成IP头部(信封)的封装。
主要是源目IP地址字段。

Router1的路由表

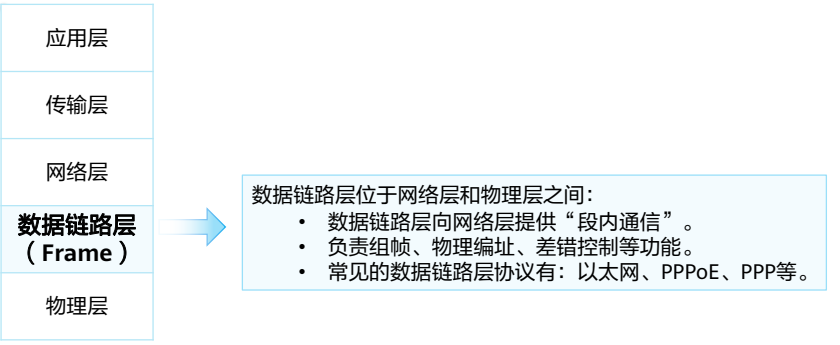
网络	出接口
网络A	GE0/0/1
...	...
...	...

- 源设备发出的报文会在其网络层头部携带该报文的源及目的设备的网络层地址；
- 具备路由功能的网络设备（例如路由器等）会维护**路由表**（相当于它的地图）；
- 当这些网络设备收到报文时，会读取其网络层携带的**目的地址**，并在其路由表中查询该地址，找到匹配项后，按照该表项的指示转发数据。

- 当采用IP作为网络层协议时，通信的双方都会被分配到一个“独一无二”的IP地址来标识自己。IP地址可被写成32位的二进制整数值形式，但为了方便人们阅读和分析，它通常被写成点分十进制的形式，即四个字节被分开用十进制表示，中间用点分隔，比如192.168.1.1。
- IP数据包的封装与转发：
  - 网络层收到上层（如传输层）协议传来的数据时候，会封装一个IP报文头部，并且把源和目的IP地址都添加到该头部中。
  - 中间经过的网络设备（如路由器），会维护一张指导IP报文转发的“地图”——路由表，通过读取IP数据包的目的地址，查找本地路由表后转发IP数据包。
  - IP数据包最终到达目的主机，目的主机通过读取目的IP地址确定是否接受并做下一步处理。
- IP协议工作时，需要如OSPF、IS-IS、BGP等各种路由协议帮助路由器建立路由表，ICMP帮忙进行网络的控制和状态诊断。

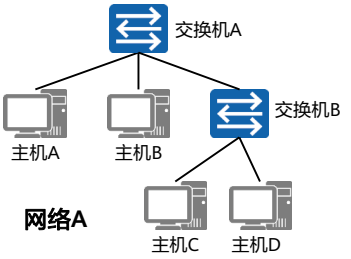
数据链路层

- 数据链路层位于网络层和物理层之间，可以向网络层的IP、IPv6等协议提供服务。数据链路层的PDU被称为Frame（帧）。
- 以太网（Ethernet）是最常见的数据链路层协议。



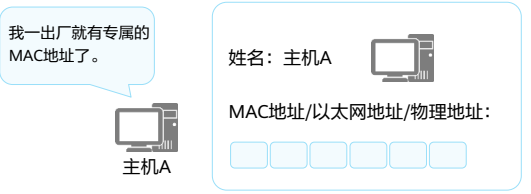
以太网与MAC地址

以太网定义



- 以太网是一种广播式数据链路层协议，支持多点接入。
- 个人电脑的网络接口遵循的就是以太网标准。
- 一般情况下，一个广播域对应着一个IP网段。

以太网MAC地址



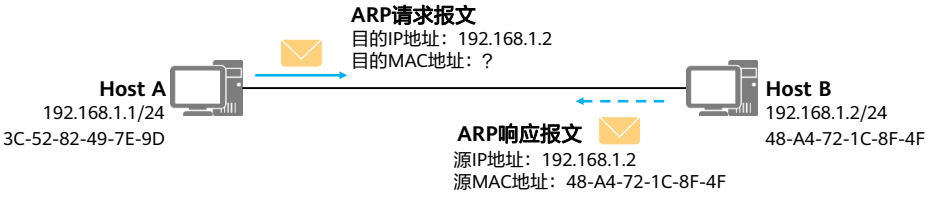
- MAC (Media Access Control)地址在网络中唯一标识一个网卡，每个网卡都需要且会有唯一的一个MAC地址。
- MAC用于在一个IP网段内，寻址找到具体的物理设备。
- 工作在数据链路层的设备。例如以太网交换机，会维护一张MAC地址表，用于指导数据帧转发。

HUAWEI

- MAC地址由48比特（6个字节）长，12位的16进制数字组成。例如：48-A4-72-1C-8F-4F

地址解析协议（ARP）

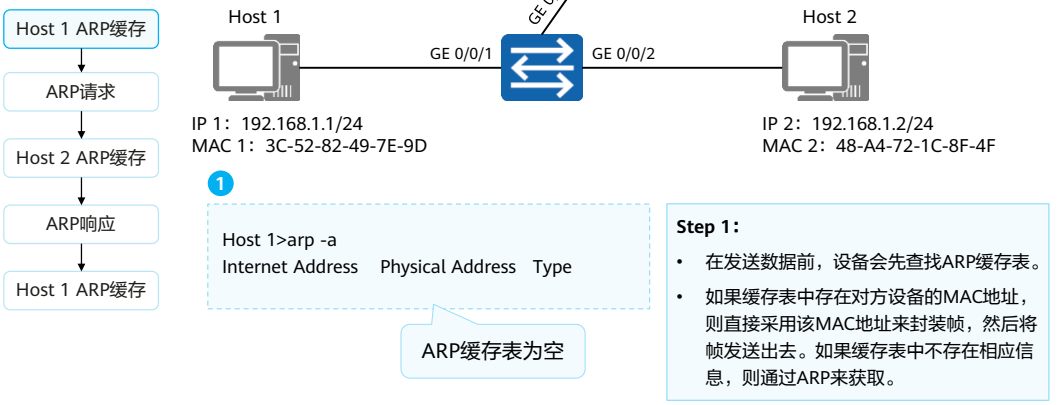
- ARP（Address Resolution Protocol）地址解析协议：
  - 根据已知的IP地址解析获得其对应的MAC地址。



HUAWEI

- ARP（Address Resolution Protocol，地址解析协议）是根据IP地址获取数据链路层地址的一个TCP/IP协议。
- ARP是IPv4中必不可少的一种协议，它的主要功能是：
  - 将IP地址解析为MAC地址；
  - 维护IP地址与MAC地址的映射关系的缓存，即ARP表项；
  - 实现网段内重复IP地址的检测。

ARP的工作原理（1）



- 网络设备一般都有一个ARP缓存（ARP Cache）。ARP缓存用来存放IP地址和MAC地址的关联信息。
- 在发送数据前，设备会先查找ARP缓存表。如果缓存表中存在对方设备的ARP表项，则直接采用该表项中的MAC地址来封装帧，然后将帧发送出去。如果缓存表中不存在相应信息，则通过发送ARP Request报文来获得它。
- 学习到的IP地址和MAC地址的映射关系会被放入ARP缓存表中存放一段时间。在有效期内（缺省：180s），设备可以直接从这个表中查找目的MAC地址来进行数据封装，而无需进行ARP查询。过了这段有效期，ARP表项会被自动删除。
- 如果目标设备位于其他网络，则源设备会在ARP缓存表中查找网关的MAC地址。然后将数据发送给网关。最后网关再把数据转发给目的设备。

ARP的工作原理 (2)

Step 2：

- 主机1通过发送ARP Request报文来获取主机2的MAC地址。
- 由于不知道目的MAC地址，因此ARP Request报文内的目的端MAC地址为0。

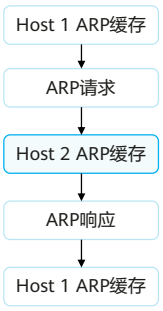
Step 3：

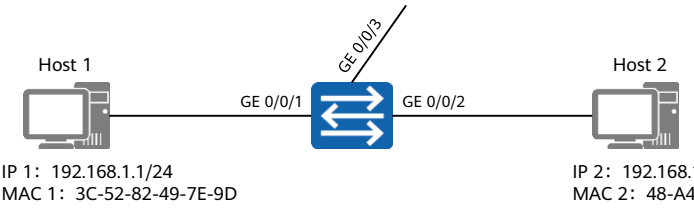
- ARP Request是广播数据帧，因此交换机收到后，会对该帧执行泛洪操作。

HUAWEI

- 主机1的ARP缓存表中不存在主机2的MAC地址，所以主机1会发送ARP Request来获取目的MAC地址。
- ARP Request报文封装在以太帧里。帧头中的源MAC地址为发送端主机1的MAC地址。此时，由于主机1不知道主机2的MAC地址，所以目的MAC地址为广播地址FF-FF-FF-FF-FF-FF。
- ARP Request报文中包含发送端MAC地址、发送端IP地址、目的端MAC地址、目的端IP地址，其中目的端MAC地址的值为0。ARP Request报文会在整个网络上传播，该网络中所有主机包括网关都会接收到此ARP Request报文。

ARP的工作原理 (3)





Step 4:

- 所有的主机接收到该ARP Request报文后，都会检查它的目的端IP地址字段与自身的IP地址是否匹配。
- 主机2发现IP地址匹配，则会将ARP报文中的发送端MAC地址和发送端IP地址信息记录到自己的ARP缓存表中。

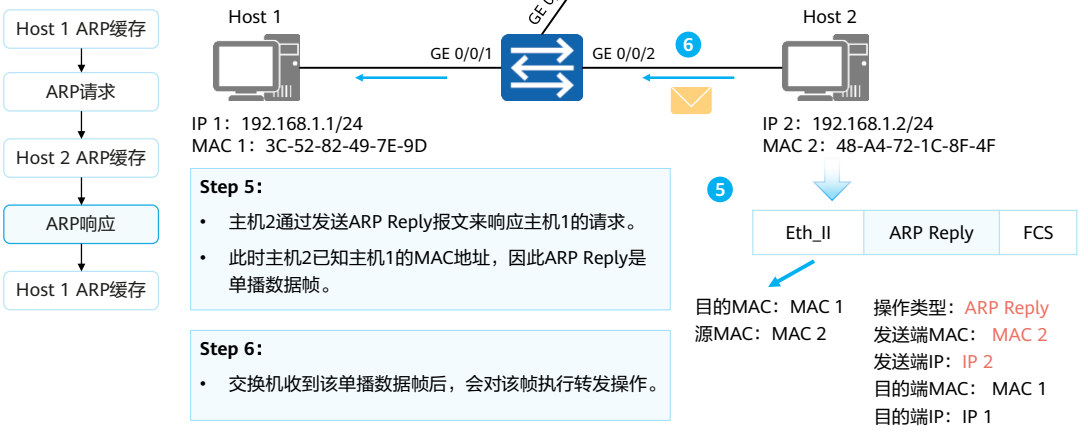
4

```
Host 2>arp -a
Internet Address      Physical Address      Type
192.168.1.1           3C-52-82-49-7E-9D     Dynamic
```


HUAWEI

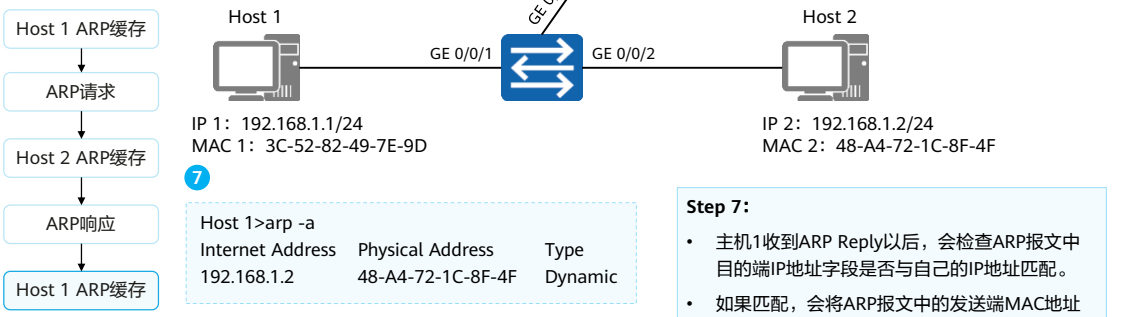
- 所有的主机接收到该ARP Request报文后，都会检查它的目的端IP地址字段与自身的IP地址是否匹配。如果不匹配，则该主机将不会响应该ARP Request报文。如果匹配，则该主机会将ARP请求报文中的发送端MAC地址和发送端IP地址信息记录到自己的ARP缓存表中，然后通过ARP Reply报文进行响应。

ARP的工作原理 (4)



- 主机2会向主机1回应ARP Reply报文。
- ARP Reply报文中的发送端IP地址是主机2自己的IP地址，目的端IP地址是主机1的IP地址，目的端MAC地址是主机1的MAC地址，发送端MAC地址是自己的MAC地址，同时操作类型被设置为Reply。
- ARP Reply报文通过单播传送。

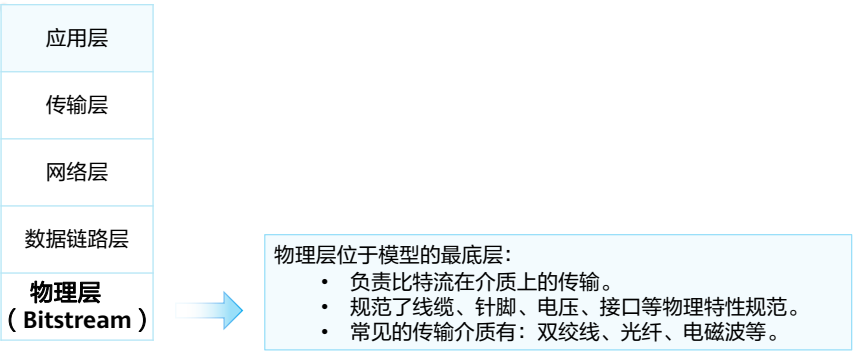
ARP的工作原理 (5)



- 主机1收到ARP Reply以后，会检查ARP报文中目的端IP地址字段与自身的IP地址是否匹配。如果匹配，ARP报文中的发送端MAC地址和发送端IP地址会被记录到主机1的ARP缓存表中。

物理层

- 数据到达物理层之后，物理层会根据物理介质的不同，将数字信号转换成光信号、电信号或者是电磁波信号。
- 物理层的PDU被称为比特流（Bitstream）。



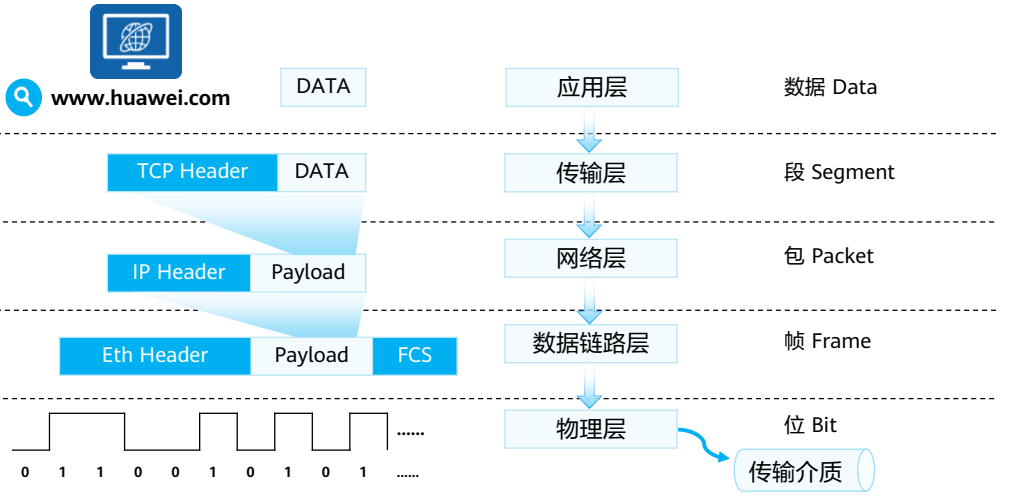
HUAWEI

常见传输介质

- 双绞线：当今以太网最常见的传输介质，按照抗电磁干扰能力还可以分为：
  - STP-屏蔽双绞线
  - UTP-非屏蔽双绞线
- 光纤传输，按照功能部件可分为：
  - 光纤：光传输介质，简单的说，就是一根玻璃纤维，用于约束光传输的通道。
  - 光模块：将电信号与光信号互转的器件，产生光信号。
- 串口电缆在WAN（Wide Area Network，广域网）中大规模使用，根据WAN线路类型不同，串口电缆在设备上连接的接口类型也不同：异/同步串口、ATM接口、POS接口、CE1/PRI接口等。
- 无线信号的传输可以通过电磁波进行，例如：无线路由器将数据通过调制以电磁波发送出去，移动终端的无线网卡将电磁波解调，得到数据，完成从无线路由器到移动终端的数据传输。

目录

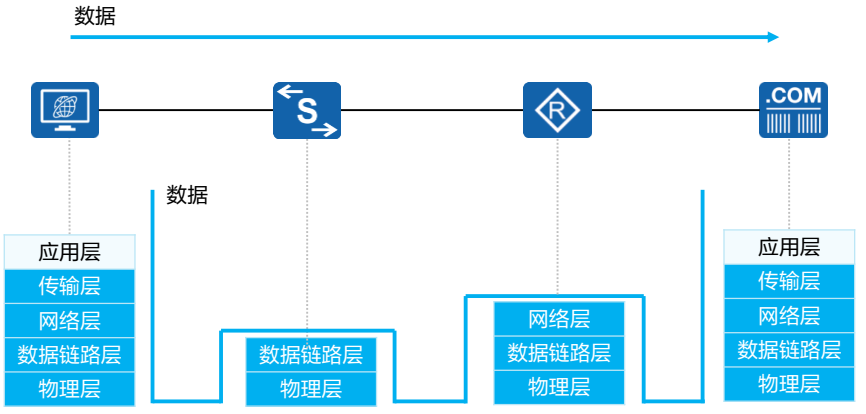
发送方数据封装



- 假设你正在通过网页浏览器访问华为官网，当你输入完网址，敲下回车后，计算机内部会发生下列事情：
  - 1. IE浏览器(应用程序)调用HTTP(应用层协议)，完成应用层数据的封装(图中DATA还应包括HTTP头部，此处省略) 。
  - 2. HTTP依靠传输层的TCP进行数据的可靠性传输，将封装好的数据传递到TCP模块。
  - 3. TCP模块给应用层传递下来的Data添加上相应的TCP头部信息(源端口、目的端口等)。此时的PDU被称作Segment(段)。
  - 4. 在IPv4网络中，TCP模块会将封装好的Segment传递给网络层的IPv4模块(若在IPv6环境，会交给IPv6模块进行处理)。
  - 5. IPv4模块在收到TCP模块传递来的Segment之后，完成IPv4头部的封装，此时的PDU被称为Packet(包)。
  - 6. 由于使用了Ethernet作为数据链路层协议，故在IPv4模块完成封装之后，会将Packet交由数据链路层的Ethernet模块(例如以太网卡)处理。
  - 7. Ethernet模块在收到IPv4模块传递来的Packet之后，添加上相应的Ethernet头部信息和FCS帧尾，此时的PDU被称为Frame(帧)。
  - 8. 在Ethernet模块封装完毕之后，会将数据传递到物理层。
  - 9. 根据物理介质的不同，物理层负责将数字信号转换成电信号，光信号，电磁波(无线)信号等。
  - 10. 转换完成的信号在网络中开始传递。

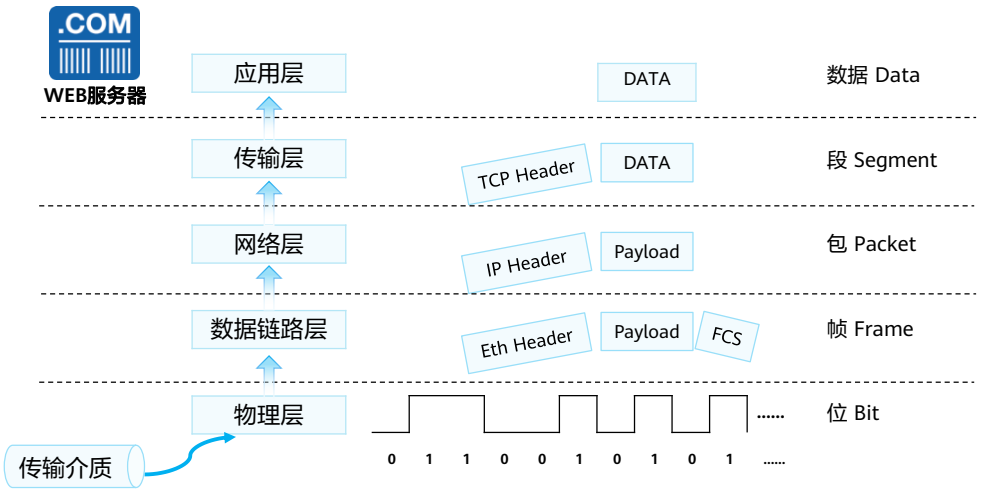
中间网络数据传输

- 封装好的完整数据，将会在网络中被传递。



- 一般情况下：
  - 网络中的二层设备（如以太网交换机）只会解封装数据的二层头部，根据二层头部的信息进行相应的“交换”操作。
  - 网络中的三层设备（如路由器）只会解封装到三层头部，并且根据三层头部的信息进行相应的“路由”操作。
  - 注：“交换”和“路由”的详细细节和原则，将会在后面的课程中详细介绍。

接收方数据解封装



- 经过中间网络传递之后，数据最终到达目的服务器。根据不同的协议头部的信息，数据将被一层层的解封装并做相应的处理和传递，最终交由WEB服务器上的应用程序进行处理。

本章总结

- 不论是OSI参考模型还是TCP/IP参考模型，都采用了分层的设计理念。
  - 各个层次之间分工、界限明确，有助于各个部件的开发、设计和故障排除
  - 通过定义在模型的每一层实现什么功能，鼓励产业的标准化
  - 通过提供接口的方式，使得各种类型的网络硬件和软件能够相互通信，提高兼容性
- 数据的产生与传递，需要各模块之间相互协作，同时每个模块又需要“各司其职”。

思考题

1. 分层模型的概念有什么好处?
2. 常见的应用层、传输层、网络层、数据链路层有哪些协议?

HUAWEI

1. 答案:
   - 各个层次之间分工、界限明确，有助于各个部件的开发、设计和故障排除。
   - 通过定义在模型的每一层实现什么功能,鼓励产业的标准化。
   - 通过提供接口的方式，使得各种类型的网络硬件和软件能够相互通信，提高兼容性。
2. 答案:
   - 应用层：HTTP、FTP、Telnet等
   - 传输层：UDP、TCP
   - 网络层：IP、ICMP等
   - 数据链路层：Ethernet、PPP、PPPoE等


谢 谢
www.huawei.com

华为VRP系统基础

前言

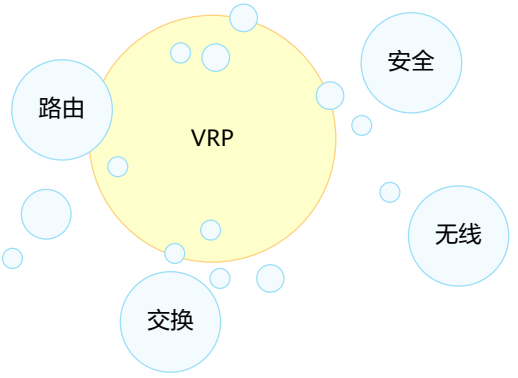
- 通用路由平台VRP（Versatile Routing Platform）是华为公司数据通信产品的通用操作系统平台。它以IP业务为核心，采用组件化的体系结构，在实现丰富功能特性的同时，还提供了基于应用的可裁剪和可扩展的功能，使得路由器和交换机的运行效率大大增加。熟悉VRP操作系统并且熟练掌握VRP配置是高效管理华为网络设备的必备基础。
- 本课程主要介绍VRP的基本概念、常用命令和CLI界面的使用。

目标

- 学完本课程后，您将能够：
  - 了解VRP的基础知识
  - 掌握CLI界面的使用
  - 掌握命令行的基本命令

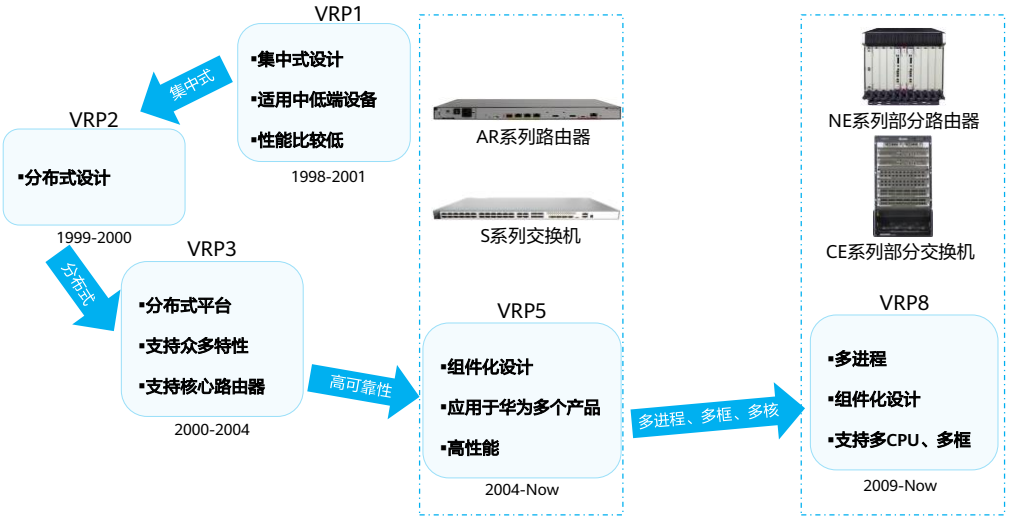
目录

什么是VRP?



- VRP是华为公司数据通信产品的通用操作系统平台，作为华为公司从低端到核心的全系列路由器、以太网交换机、业务网关等产品的软件核心引擎。
- VRP提供以下功能：
  - 实现统一的用户界面和管理界面
  - 实现控制平面功能，并定义转发平面接口规范
  - 实现各产品转发平面与VRP控制平面之间的交互
  - 屏蔽各产品链路层对于网络层的差异

VRP的发展



文件系统

- 文件系统是指对存储器中文件、目录的管理，功能包括查看、创建、重命名和删除目录，拷贝、移动、重命名和删除文件等。
- 掌握文件系统的基本操作，对于网络工程师高效管理设备的配置文件和VRP系统文件至关重要。

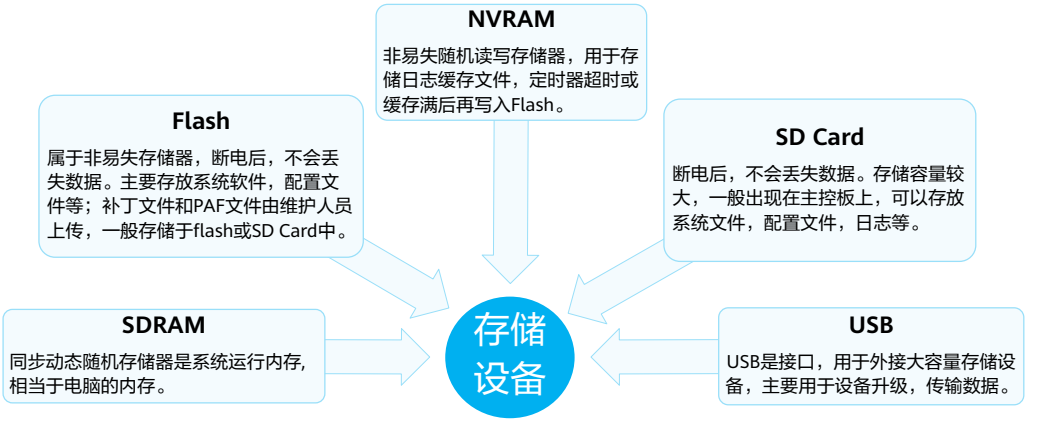
常见文件类型

HUAWEI

- 配置文件是命令行的集合。用户将当前配置保存到配置文件中，以便设备重启后，这些配置能够继续生效。另外，通过配置文件，用户可以非常方便地查阅配置信息，也可以将配置文件上传到别的设备，来实现设备的批量配置。
- 补丁是一种与设备系统软件兼容的软件，用于解决设备系统软件少量且急需解决的问题。在设备的运行过程中，有时需要对设备系统软件进行一些适应性和排错性的修改，如改正系统中存在的缺陷、优化某功能以适应业务需求等。
- 文件的管理方式包括：
  - 通过Console或者telnet等直接登陆系统管理
  - 通过FTP、TFTP或SFTP登录设备进行管理

存储设备

- 存储设备包括：SDRAM、Flash、NVRAM 、SD Card、USB。



HUAWEI

- 存储设备包括：SDRAM、Flash、NVRAM 、SD Card、USB
  - SDRAM是系统运行内存，相当于电脑的内存；
  - NVRAM非易失存储器，日志写入FLASH操作是耗时耗CPU的操作，因此采用缓存机制，即日志产生后，先存入缓存，定时器超时或缓存满后再写入Flash；
  - Flash与SD Card属于非易失存储器，配置文件与系统文件存放于flash或SD Card中，具体设备参考产品文档；
  - SD Card是外置的SD存储卡，用来扩展。USB是接口，用于外接大容量存储设备，主要用于设备升级，传输数据；
  - 补丁文件和PAF文件由维护人员上传可自行指定放置。

设备初始化过程

- 设备上电后，首先运行BootROM软件，初始化硬件并显示设备的硬件参数，然后运行系统软件，最后从默认存储路径中读取配置文件进行设备的初始化操作。

```
BIOS Creation Date : Jan  5 2020, 18:00:24
DDR DRAM init : OK
Start Memory Test ? ('t' or 'T' is test):skip
Copying Data : Done
Uncompressing : Done
......
Press Ctrl+B to break auto startup ... 1
Now boot from flash:/AR2220E-V200R007C00SPC600.cc,
......
```


- BootROM（Boot Read-Only Memory）是一组固化到设备内主板上ROM芯片中的程序，它保存着设备最重要的基本输入输出的程序、系统设置信息、开机后自检程序和系统自启动程序。
- 开机界面提供了系统启动的运行程序和正在运行的VRP版本及其加载路径等信息。

设备管理

- 用户对设备的常见管理方式主要有命令行方式和Web网管方式两种。
- 用户需要通过相应的方式登录到设备后才能对设备进行管理。

Web网管方式	命令行方式
• Web网管方式通过图形化的操作界面，实现对设备直观方便地管理与维护，但是此方式仅可实现对设备部分功能的管理与维护。 • Web网管方式可以通过HTTP和HTTPS方式登录设备。	• 命令行方式需要用户使用设备提供的命令行对设备进行管理与维护，此方式可实现对设备的精细化管理，但是要求用户熟悉命令行。 • 命令行方式可以通过Console口、Telnet或SSH方式登录设备。

VRP用户界面

- 用户通过命令行方式登录设备时，系统会分配一个用户界面用来管理、监控设备和用户间的当前会话。
- 设备系统支持的用户界面有Console用户界面和虚拟类型终端VTY（ Virtual Type Terminal ）用户界面。

Console用户界面

- Console用户界面用来管理和监控通过Console口登录的用户。
- 用户终端的串行口可以与设备Console口直接连接，实现对设备的本地访问。

VTY用户界面

- VTY用户界面用来管理和监控通过VTY方式登录的用户。
- 用户通过终端与设备建立Telnet或STelnet连接后，即建立了一条VTY通道，通过VTY通道实现对设备的远程访问。

VRP用户级别

- VRP提供基本的权限控制，可以实现不同级别的用户能够执行不同级别的命令，用以限制不同用户对设备的操作。

用户等级	命令等级	名称	说明
0	0	参观级	可使用网络诊断工具命令（ping、tracert）、从本设备出发访问外部设备的命令（Telnet客户端命令）、部分display命令等。
1	0 and 1	监控级	用于系统维护，可使用display等命令。
2	0,1 and 2	配置级	可使用业务配置命令，包括路由、各个网络层次的命令，向用户提供直接网络服务。
3-15	0,1,2 and 3	管理级	可使用用于系统基本运行的命令，对业务提供支撑作用，包括文件系统、FTP、TFTP下载、命令级别设置命令以及用于业务故障诊断的debugging命令等。

- 为了限制不同用户对设备的访问权限，系统对用户也进行了分级管理。用户的级别与命令级别对应，不同级别的用户登录后，只能使用等于或低于自己级别的命令。缺省情况下，命令级别按0～3级进行注册，用户级别按0～15级进行注册，用户级别和命令级别对应关系如表所示。

WEB网管方式登录

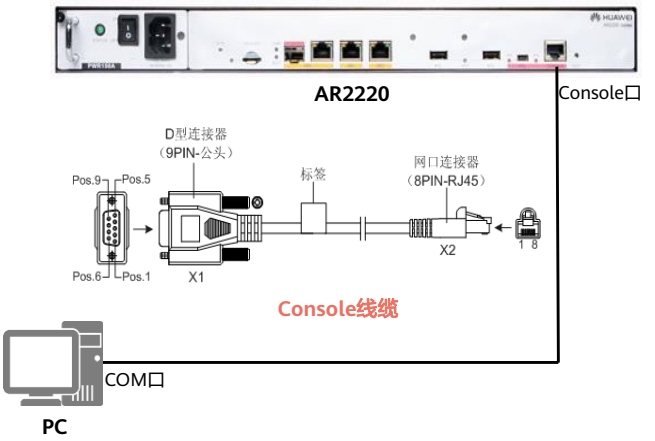
以华为AR系列路由器为例，PC终端打开浏览器软件，在地址栏中输入“https://192.168.1.1”，按下回车键，显示AR Web管理平台登录界面。

- 注意：不同设备登录界面、登录方式和登录的IP可能不同，具体参考产品文档。

命令行方式 - 本地登录 (1)

设备登录方式分为两种：本地登录和远程登录。其中本地登录包括：

- 当用户需为第一次上电的设备进行配置时，可通过Console口本地登录设备。
- 控制口（Console Port）是一种通信串行端口，由设备的主控板提供。
- 用户终端的串行端口可以与设备Console口直接连接，然后通过PuTTY工具本地登录实现对设备的本地配置。

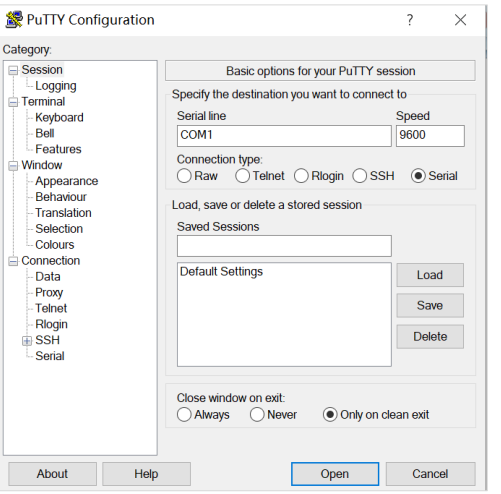


- 使用Console线缆来连接交换机或路由器的Console口与计算机的COM口，这样就可以通过计算机通过PuTTY工具实现本地调试和维护。Console口是一种符合RS232串口标准的RJ45接口。目前大多数台式电脑提供的COM口都可以与Console口连接。笔记本电脑一般不提供COM口，需要使用USB到RS232的转换接口。
- Console口登录是设备默认开启的功能，不需要对设备做预配置。

命令行方式 - 本地登录 (2)

PuTTY工具是一个Telnet、SSH、串行接口等的连接软件。

本地登录时，终端设备采用串口与华为设备Console口连接，所以采用“Serial”连接类型，COM端口根据终端设备实际端口选取，速率固定为9600。



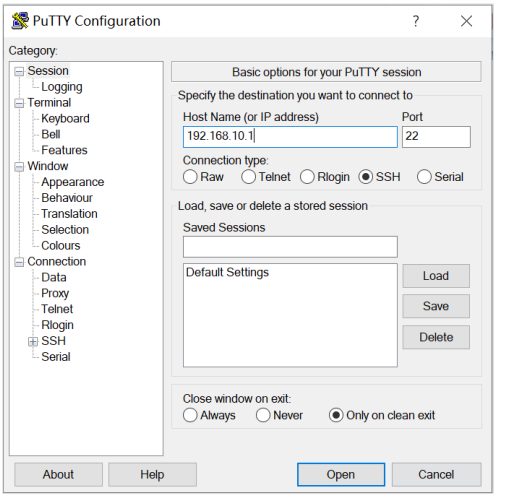
- 很多终端模拟程序都能发起Console连接，例如，可以使用PuTTY连接到VRP操作系统。使用PuTTY连接VRP时，必须设置端口参数。上图是端口参数设置的示例，如果对参数值做了修改，需要恢复默认参数值。
- 完成设置以后，点击“Open”按钮即可与VRP建立连接。

命令行方式 - 远程登录

远程登录允许终端远程登录到任何可以充当远程登录服务器的设备，对这些网络设备进行集中的管理和维护。远程登录方法包括：Telnet和SSH。

- 如果通过SSH远程登录，连接类型为“SSH”，需要输入远程登录服务器的IP地址，端口号缺省为22。
- 如果通过Telnet远程登录，连接类型为“Telnet”，需要输入远程登录服务器的IP地址，端口号缺省为23。

192.168.10.1

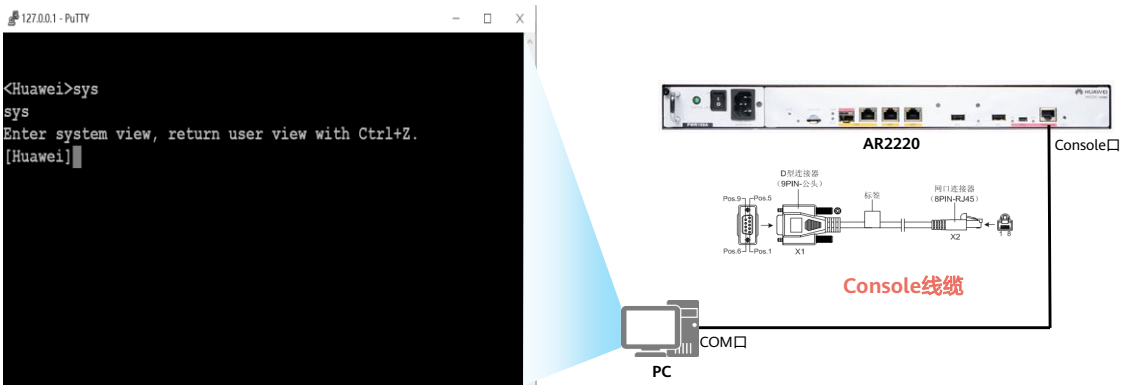


HUAWEI

- 设备默认不开启SSH登录功能，需要用户先通过Console口登录，配置上SSH登录必须的参数之后，才可以使用SSH登录功能。

命令行界面

- 登录成功后即进入命令行界面CLI（Command Line Interface）。
- 命令行界面是工程师与网络设备进行交互的常用工具。

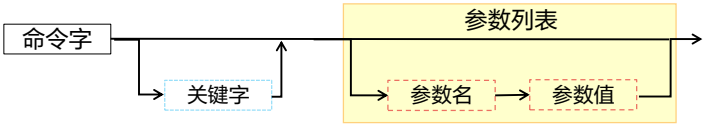


- 命令行界面CLI（Command Line Interface）是用户与路由器进行交互的常用工具。用户登录到路由器出现命令行提示符后，即进入命令行界面。

目录

基本命令结构

- 华为提供的命令按照一定的格式设计，用户可以通过命令行界面输入命令，由命令行界面对命令进行解析，实现用户对路由器的配置和管理。



- **命令字**：规定了系统应该执行的功能，如display（查询设备状态），reboot（重启设备）等命令字。
- **关键字**：特殊的字符构成，用于进一步约束命令，是对命令的拓展，也可用于表达命令构成逻辑而增设的补充字符串。
- **参数列表**：是对命令执行功能的进一步约束。包括一对或多对参数名和参数值。

举例1:
display ip interface GE0/0/0，查看接口信息的命令
命令字：display
关键字：ip
参数名：interface
参数值：GE0/0/0

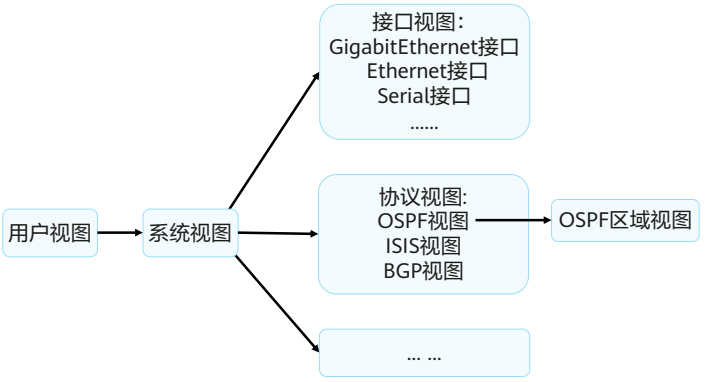
举例2:
Reboot，重启设备的命令
命令字：reboot，操作命令都要用命令字，并且必须从规范的命令字集合中选取。

HUAWEI

- 每条命令有最多一个命令字，若干个关键字和参数，形成一条命令，参数必须由参数名和参数值组成。
- 命令字、关键字、参数名、参数值之间，需要用空格分隔开。

命令行视图 (1)

- 设备提供了多样的配置和查询命令，为便于用户使用这些命令，VRP系统按功能分类将命令分别注册在不同的命令行视图下。



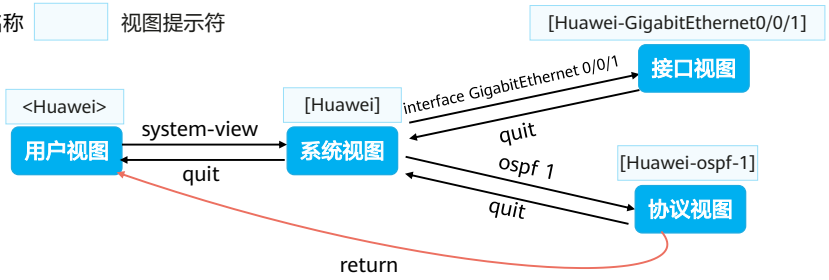
- 用户视图：用户可以完成查看运行状态和统计信息等功能。
- 系统视图：用户可以配置系统参数以及通过该视图进入其他的功能配置视图。
- 其他视图：比如接口视图，协议视图，用户可以进行接口参数和协议参数配置。

HUAWEI

- 用户视图应为登录系统后的第一个视图。在用户视图中不提供除查询和工具命令之外的其他命令。
- 用户视图中，唯一可进入的视图是系统视图；系统视图中提供全局的配置命令；如果系统存在下一级配置视图，则在系统视图中须提供进入下一级配置视图的命令。

命令行视图 (2)

视图名称　　视图提示符



命令行举例：

```
<Huawei>system-view                                  #用户首先进入用户视图，通过命令进入系统视图
[Huawei]interface GigabitEthernet 0/0/1              #在系统视图进入接口视图
[Huawei-GigabitEthernet0/0/1]ip address 192.168.1.1 24   #配置IP地址
[Huawei-GigabitEthernet0/0/1]quit                    #退回到上一个视图
[Huawei]ospf 1                                       #在系统视图进入协议视图
[Huawei-ospf-1]area 0                                #在协议视图进入OSPF区域视图
[Huawei-ospf-1-area-0.0.0.0]return                   #返回用户视图
<Huawei>
```


- 登录到系统后，首先进入的是用户视图，这里仅提供查询，以及ping，telnet等工具命令，不提供任何配置命令。
- 用户视图下通过system-view命令可以进入到系统视图，系统视图提供一些简单的全局配置功能。
- 一些复杂的配置功能，如配置一个以太网接口，需要的参数比较多，因此系统提供以太网接口的配置视图，在系统视图中，使用命令interface GigabitEthernet X（此处X表示一个具体接口的编号）进入GE接口配置视图，这个配置视图当前针对GE接口X，所有的命令仅对GE接口X生效。

编辑命令行 (1)

- 设备的命令行界面提供基本的命令行编辑功能，以下为常用的编辑功能：

1、功能键

- 退格键Backspace：删除光标位置的前一个字符，光标左移，若已经到达命令首，则响铃告警。
- 左光标键←或<Ctrl+B>：光标向左移动一个字符位置，若已经到达命令首，则响铃告警。
- 右光标键→或<Ctrl+F>：光标向右移动一个字符位置，若已经到达命令尾，则响铃告警。

2、不完整关键字输入

- 设备支持不完整关键字输入，即在当前视图下，当输入的字符能够匹配唯一的关键字时，可以不必输入完整的关键字，例如：

```
<Huawei>d cu
<Huawei>di cu
<Huawei>dis cu
<Huawei>d c
        ^
Error:Ambiguous command found at '^' position.
<Huawei>dis c
           ^
Error:Ambiguous command found at '^' position.
```

“display current-configuration”命令，可以输入d cu、di cu或dis cu等都可以执行此命令，但不能输入d c或dis c等，因为以d c、dis c开头的命令不唯一。

- 注：此处的关键字与命令行格式中的“关键字”不同，一条命令中除“参数值”外都可以被叫做关键字。

编辑命令行 (2)

3、Tab键的使用：

- 如果与之匹配的关键字唯一，按下<Tab>键，系统自动补全关键字，补全后，反复按<Tab>关键字不变。

```
[Huawei] info-                    #按下Tab键
[Huawei] info-center
```

- 如果与之匹配的关键字不唯一，反复按<Tab>键可循环显示所有以输入字符串开头的关键字。

```
[Huawei] info-center log          #按下Tab键
[Huawei] info-center logbuffer    #继续按Tab键循环翻词
[Huawei] info-center logfile
[Huawei] info-center loghost
```

- 如果没有与之匹配的关键字，按Tab键后，关键字不变。

```
[Huawei] info-center loglog       #输入错误的关键字，按下Tab键
[Huawei] info-center loglog
```

使用命令行在线帮助

- 用户在使用命令行时，可以使用在线帮助以获取实时帮助，从而无需记忆大量的复杂的命令。
- 命令行在线帮助可分为完全帮助和部分帮助，可通过输入“？”实现。

完全帮助

- 当用户输入命令时，可以使用命令行的完全帮助获取全部关键字和参数的提示。

```
<Huawei> ?
User view commands:
  arp-ping       ARP-ping
  autosave       <Group> autosave command group
  backup         Backup  information
  cd             Change current directory
  clear          Clear
  clock          Specify the system clock
...
```

部分帮助

- 当用户输入命令时，如果只记得此命令关键字的开头一个或几个字符，可以使用命令行的部分帮助获取以该字符串开头的所有关键字的提示。

```
<Huawei> d?
  debugging  <Group> debugging command group
  delete     Delete a file
  dialer     Dialer
  dir        List files on a filesystem
  display    Display information
```


HUAWEI

- 以上获取的在线帮助的显示信息仅为示意，请以设备实际显示为准。

解读命令行的错误信息

- 用户键入的命令，如果通过语法检查，则正确执行，否则系统将会向用户报告错误信息。

```
[Huawei] sysname
                ^
Error:Incomplete command found at '^' position.        #箭头所指地方提示命令不完整，需要进一步补齐

[Huawei] router if 1.1.1.1
                ^
Error: Unrecognized command found at '^' position.     #箭头所指地方提示该命令不能识别，需要确认命令正确性

[Huawei] a
         ^
Error:Ambiguous command found at '^' position.         #箭头所指的命令不明确，有多个a开头的命令字

[Huawei-GigabitEthernet0/0/0]ospf cost 800000          #箭头所指的参数值越界
                             ^
Error: Wrong parameter found at '^' position.
```

使用undo命令行

- 在命令前加undo关键字，即为undo命令行。undo命令行一般用来恢复缺省情况、禁用某个功能或者删除某项配置。以下为参考案例：
  - 使用undo命令行恢复缺省情况

```
<Huawei> system-view
[Huawei] sysname Server
[Server] undo sysname
[Huawei]
```

  - 使用undo命令禁用某个功能

```
<Huawei> system-view
[Huawei] ftp server enable
[Huawei] undo ftp server
```

  - 使用undo命令删除某项设置

```
[Huawei]interface g0/0/1
[Huawei-GigabitEthernet0/0/1]ip address 192.168.1.1 24
[Huawei-GigabitEthernet0/0/1]undo ip address
```

使用命令行的快捷键

- 用户可以使用设备中的快捷键，完成对命令的快速输入，从而简化操作。
- 系统中的快捷键分成两类，自定义快捷键和系统快捷键。

自定义快捷键

- 自定义快捷键：共有4个，<Ctrl+G>、<Ctrl+L>、<Ctrl+O>和<Ctrl+U>。
- 用户可以根据自己的需要将这4个快捷键与任意命令进行关联，当使用快捷键时，系统自动执行它所对应的命令。

```
<Huawei> system-view
[Huawei] hotkey ctrl_l "display tcp status"
```

系统快捷键

- CTRL_A：将光标移动到当前行的开头
- CTRL_B：将光标向左移动一个字符
- CTRL_C：停止当前命令的运行
- CTRL_E：将光标移动到当前行的末尾
- CTRL_X：删除光标左侧所有的字符
- CTRL_Y：删除光标所在位置及其右侧所有的字符
- CTRL_Z：返回到用户视图
- CTRL+]：终止当前连接或切换连接

目录

- VRP基于文件系统来管理设备上的文件和目录。在管理文件和目录时，经常会使用一些基本命令来查询文件或者目录的信息，常用的命令包括**pwd**，**dir** [**/all**] [filename | directory]和**more** [**/binary**] filename [offset] [**all**]。
  - **pwd**命令用来显示当前工作目录。
  - **dir** [**/all**] [filename | directory]命令用来查看当前目录下的文件信息。
  - **more** [**/binary**] filename [offset] [**all**]命令用来查看文本文件的具体内容。
  - 本例中，在用户视图中使用**dir**命令，可以查看flash中的文件信息。
- 目录操作常用的命令包括：**cd** *directory*，**mkdir** *directory*和**rmdir** *directory*。
  - **cd** *directory*命令用来修改用户当前的工作目录。
  - **mkdir** *directory*命令能够创建一个新的目录。目录名称可以包含1-64个字符。

常见文件系统操作命令 (2)

6.删除目录

```
<Huawei>rmdir
```

7.复制文件

```
<Huawei>copy
```

8.移动文件

```
<Huawei>move
```

9.重命名文件

```
<Huawei>rename
```

10.删除文件

```
<Huawei>delete
```


- **rmdir** *directory*命令能够删除文件系统中的目录，此处需要注意的是，只有空目录才能被删除。
- **copy** *source-filename destination-filename*命令可以复制文件。如果目标文件已存在，系统会提示此文件将被替换。目标文件名不能与系统启动文件同名，否则系统将会出现错误提示。
- **move** *source-filename destination-filename*命令可以用来将文件移动到其他目录下。**move**命令只适用于在同一储存设备中移动文件。
- **rename** *old-name new-name*命令可以用来对目录或文件进行重命名。
- **delete** [/**unreserved**] [/**force**] { *filename* | *devicename* }命令可以用来删除文件。不带unreserved参数的情况下，被删除的文件将直接被移动到回收站。回收站中的文件也可以通过执行**undelete**命令进行恢复，但是如果执行delete命令时指定了unreserved参数，则文件将被永久删除。在删除文件时，系统会提示“是否确定删除文件”，但如果命令中指定了/force 参数，系统将不会给出任何提示信息。*filename*参数指的是需要删除的文件的名称，*devicename*参数指定了储存设备的名称。

常见文件系统操作命令 (3)

11.恢复删除的文件

```
<Huawei>undelete
```

12.彻底删除回收站中的文件

```
<Huawei>reset  recycle-bin
```


- **reset recycle-bin** [*filename* | *devicename*]可以用来永久删除回收站中的文件，*filename*参数指定了需要永久删除的文件的名称，*device-name*参数指定了储存设备的名称。

基本配置命令 (1)

1.配置设备名称

```
[Huawei] sysname name
```

2.设置系统时钟

```
<Huawei> clock timezone time-zone-name { add | minus } offset
```

用来对本地时区信息进行设置。

```
<Huawei> clock datetime [ utc ] HH:MM:SS YYYY-MM-DD
```

用来设置设备当前或UTC日期和时间。

```
<Huawei> clock daylight-saving-time
```

用来设置设备的夏令时。

- 网络上一般都会部署不止一台设备，管理员需要对这些设备进行统一管理。在进行设备调试的时候，首要任务是设置设备名。设备名用来唯一地标识一台设备。AR系列路由器默认的设备名是Huawei，而S系列交换机默认的设备名是HUAWEI。设备名称一旦设置，立刻生效。
- 为了保证与其他设备协调工作，需要准确设置系统时钟。系统时钟的=UTC（Coordinated Universal Time）+当前时区与UTC的时间差，一般设备上都会有内置的UTC和时间差配置。
  - 可以通过clock datetime命令直接设置设备的系统时钟，格式为HH:MM:SS YYYY-MM-DD，此时UTC等于系统时钟-时间差。
  - 也可以通过修改UTC和系统当前时区来修改系统时钟
    - **clock datetime** [utc] *HH:MM:SS YYYY-MM-DD*用来修改UTC时间。
    - **clock timezone** *time-zone-name* { **add** | **minus** } *offset* 用来配置本地时区信息。本地时间加上或减去*offset*即为UTC。
  - 有的地区实行夏令时制，因此当进入夏令时实施区间的一刻，系统时间要根据用户的设定进行夏令时时间的调整。VRP支持夏令时功能。

基本配置命令 (2)

3.配置命令等级

[Huawei] **command-privilege level** *level* **view** *view-name command-key*

用来设置指定视图内的命令的级别。命令级别分为参观、监控、配置、管理4个级别，分别对应标识0、1、2、3。

4.配置用户通过Password方式登录设备

[Huawei]**user-interface vty 0 4**
[Huawei-ui-vty0-4]**set authentication password cipher** *information*

用来进入指定的用户视图并配置用户认证方式为password。系统支持的用户界面包括Console用户界面和VTY用户界面，Console界面用于本地登录，VTY界面用于远程登录。默认情况下，设备一般最多支持15个用户同时通过VTY方式访问。

5.配置用户连接的超时时间

[Huawei] **idle-timeout** *minutes [seconds]*

用来设置用户界面断开连接的超时时间。如果用户在一段时间内没有输入命令，系统将断开连接。缺省情况下，超时时间是10分钟。

- 每类用户界面都有对应的用户界面视图。用户界面（User-interface）视图是系统提供的一种命令行视图，用来配置和管理所有工作在异步交互方式下的物理接口和逻辑接口，从而达到统一管理各种用户界面的目的。在连接到设备前，用户要设置用户界面参数。系统支持的用户界面包括Console用户界面和VTY用户界面。控制口（Console Port）是一种通信串行端口，由设备的主控板提供。虚拟类型终端（Virtual Type Terminal）是一种虚拟线路端口，用户通过终端与设备建立Telnet或SSH连接后，也就建立了一条VTY，即用户可以通过VTY方式登录设备。设备一般最多支持15个用户同时通过VTY方式访问。执行user-interface maximum-vty number 命令可以配置同时登录到设备的VTY类型用户界面的最大个数。如果将最大登录用户数设为0，则任何用户都不能通过Telnet或者SSH登录到路由器。display user-interface 命令用来查看用户界面信息。
- 不同的设备，或使用不同版本的VRP软件系统，具体可以被使用的VTY接口的最大数量可能不同。

基本配置命令 (3)

6.配置接口IP地址

```
[Huawei]interface interface-number
[Huawei-interface-number]ip address ip address
```

用来给设备上的物理或逻辑接口配置IP地址。

7.查看当前运行的配置文件

```
<Huawei>display current-configuration
```

8.配置文件保存

```
<Huawei>save
```

9.查看保存的配置

```
<Huawei>display saved-configuration
```


- 要在接口运行IP服务，必须为接口配置一个IP地址。一个接口一般只需要一个IP地址,如果接口配置了新的主IP地址，那么新的主IP地址就替代了原来的主IP地址。
- 用户可以利用**ip address** *ip-address* { *mask* | *mask-length* } 命令为接口配置IP地址，这个命令中，*mask*代表子网掩码，如255.255.255.0，*mask-length* 代表的是掩码长度，如24。这两者任取其一均可。
- Loopback接口是一个逻辑接口，可用来虚拟一个网络或者一个IP主机。在运行多种协议的时候，由于Loopback接口稳定可靠，所以也可以用来做管理接口。
- 在给物理接口配置IP地址时，需要关注该接口的物理状态。默认情况下，华为路由器和交换机的接口状态为up；如果该接口曾被手动关闭，则在配置完IP地址后，应使用undo shutdown打开该接口。

基本配置命令 (4)

10.清除已保存的配置

```
<Huawei>reset saved-configuration
```

11.查看系统启动配置参数

```
<Huawei> display startup
```

用来查看设备本次及下次启动相关的系统软件、备份系统软件、配置文件、License文件、补丁文件以及语音文件。

12.配置系统下次启动时使用的配置文件

```
<Huawei>startup saved-configuration configuration-file
```

设备升级时，可以通过此命令让设备下次启动时加载指定的配置文件

13.配置设备重启

```
<Huawei>reboot
```


- **reset saved-configuration**命令用来清除配置文件或配置文件中的内容。执行该命令后，如果不使用命令**startup saved-configuration**重新指定设备下次启动时使用的配置文件，也不使用save命令保存当前配置，则设备下次启动时会采用缺省的配置参数进行初始化。
- **display startup**命令用来查看设备本次及下次启动相关的系统软件、备份系统软件、配置文件、License文件、补丁文件以及语音文件。
- **startup saved-configuration** *configuration-file* 命令用来指定系统下次启动时使用的配置文件，*configuration-file*参数为系统启动配置文件的名称。
- **reboot**命令用来重启设备，重启前提示用户是否保存配置。

目录

案例一：文件查询命令、目录操作

需求说明：

- 查看路由器RTA当前目录下的文件和目录的信息；
- 创建一个新目录test，然后删除该目录。

RTA

```
<RTA>pwd
flash:
<RTA>dir
Directory of flash:/
  Idx  Attr     Size(Byte)  Date        Time(LMT)  FileName
    0  drw-      -          Dec 27 2019  02:54:09   dhcp
    1  -rw-      121,802    May 26 2014 09:20:58   portalpage.zip
    2  -rw-      2,263      Dec 27 2019  02:53:59   statemach.efs
    3  -rw-      828,482    May 26 2014 09:20:58   sslvpn.zip

1,090,732 KB total (784,464 KB free)
<RTA>mkdir test
<RTA>dir
Directory of flash:/
  Idx  Attr     Size(Byte)  Date        Time(LMT)  FileName
    0  drw-      -          Dec 27 2019 02:54:39    test
    1  drw-      -          Dec 27 2019 02:54:09    dhcp
    2  -rw-      121,802    May 26 2014 09:20:58    portalpage.zip
    3  -rw-      2,263      Dec 27 2019 02:53:59    statemach.efs
    4  -rw-      828,482    May 26 2014 09:20:58    sslvpn.zip

1,090,732 KB total (784,460 KB free)
<RTA>rmdir test
```

案例二：文件操作 (1)

需求说明：

- 将文件huawei.txt重命名为save.zip；
- 将文件save.zip复制并命名为file.txt；
- 将文件file.txt移动到dhcp目录下；
- 删除文件file.txt；
- 恢复删除文件file.txt。

RTA

```
<RTA>rename huawei.txt save.zip
<RTA>dir
Directory of flash:/
Idx  Attr     Size(Byte)    Date        Time(LMT)    FileName
  0  drw-              -    Mar 04 2020 04:39:52     dhcp
  1  -rw-        121,802    May 26 2014 09:20:58     portalpage.zip
  2  -rw-        828,482    Mar 04 2020 04:51:45     save.zip
  3  -rw-          2,263    Mar 04 2020 04:39:45     statemach.efs
  4  -rw-        828,482    May 26 2014 09:20:58     sslvpn.zip

1,090,732 KB total (784,464 KB free)
<RTA>copy save.zip file.txt
<RTA>dir
Directory of flash:/
Idx  Attr     Size(Byte)    Date        Time(LMT)    FileName
  0  drw-              -    Mar 04 2020 04:39:52     dhcp
  1  -rw-        121,802    May 26 2014 09:20:58     portalpage.zip
  2  -rw-        828,482    Mar 04 2020 04:51:45     save.zip
  3  -rw-          2,263    Mar 04 2020 04:39:45     statemach.efs
  4  -rw-        828,482    May 26 2014 09:20:58     sslvpn.zip
  5  -rw-        828,482    Mar 04 2020 04:56:05     file.txt

1,090,732 KB total (784,340 KB free)
```

案例二：文件操作 (2)

需求说明：

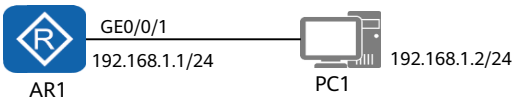
- 将文件huawei.txt重命名为save.zip；
- 将文件save.zip复制为file.txt；
- 将文件file.txt移动到dhcp目录下；
- 删除文件file.txt；
- 恢复删除文件file.txt。

RTA

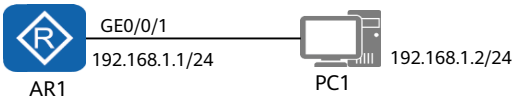
```
<RTA>move file.txt flash:/dhcp/
<RTA>cd dhcp
<RTA>dir
Directory of flash:/dhcp/
  Idx  Attr     Size(Byte)  Date        Time(LMT)    FileName
    0  -rw-          98     Dec 27 2019 02:54:09     dhcp-duid.txt
    1  -rw-       121,802   Dec 27 2019 03:13:50     file.txt

1,090,732 KB total (784,344 KB free)
<RTA>delete file.txt
<RTA>dir
Directory of flash:/dhcp/
  Idx  Attr     Size(Byte)  Date        Time(LMT)    FileName
    0  -rw-          98     Dec 27 2019 02:54:09     dhcp-duid.txt

1,090,732 KB total (784,340 KB free)
<RTA>undelete file.txt
<RTA>dir
Directory of flash:/dhcp/
  Idx  Attr     Size(Byte)  Date        Time(LMT)    FileName
    0  -rw-          98     Dec 27 2019 02:54:09     dhcp-duid.txt
    1  -rw-       121,802   Dec 27 2019 03:13:50     file.txt
1,090,732 KB total (784,340 KB free)
```

案列三：VRP基本配置命令

- 如图，某工程师需要为公司配置路由器，需求如下：
  - 路由器与PC互通，地址规划如图；
  - 公司其他人员可以通过PC远程登录访问路由器，密码是huawei123，但是只能查看配置不能随意修改配置命令；
  - 将当前配置保存为huawei.zip文件，并配置系统下次启动时使用该配置文件。



配置步骤（1）



配置接口地址

```
<Huawei>system-view
[Huawei]sysname AR1
[AR1]interface GigabitEthernet 0/0/1
[AR1-GigabitEthernet0/0/1]ip address 192.168.1.1 24
[AR1-GigabitEthernet0/0/1]quit
```

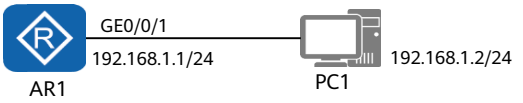
配置用户权限和用户认证

```
[AR1]user-interface vty 0 4
[AR1-ui-vty0-4]authentication-mode password
Please configure the login password (maximum length
16):huawei123
[AR1-ui-vty0-4]user privilege level 1
[AR1-ui-vty0-4]quit
```

不同版本设备在配置设备密码时，命令会有稍微不同，具体操作命令可以参考产品文档。

- 部分型号设备配置密码时只需要输入” **authentication-mode password** “命令，便会自动跳出输入密码的页面，将密码输入即可；
- 有的设备密码设置命令为” **set authentication-mode password** 密码”，密码需要手动输入。

配置步骤 (2)



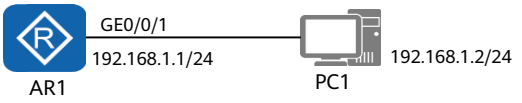
配置系统下次启动文件

```
<AR1>save huawei.zip
 Are you sure to save the configuration to huawei.zip? (y/n)[n]:y
  It will take several minutes to save configuration file, please wait.....
  Configuration file had been saved successfully
  Note: The configuration file will take effect after being activated
<AR1>startup saved-configuration huawei.zip
```

配置默认保存在vrpcfg.cfg文件中，也可以创建保存文件名称，华为VRPv5与VRPv8操作系统指定启动文件的命令是相同的，不同在于保存的目录不同。

- 设备中直接使用命令”save”进行保存，默认保存在vrpcfg.cfg文件中，当然也可以更改保存文件名,VRP5操作系统默认文件放置在flash:目录下。

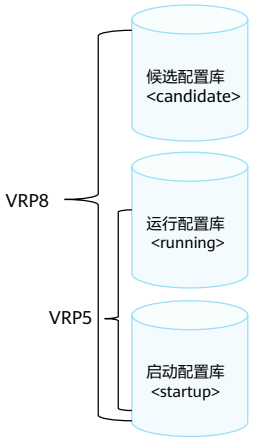
查看配置结果




```
<AR1>display startup
MainBoard:
  Startup system software:                   null
  Next startup system software:              null
  Backup system software for next startup:   null
  Startup saved-configuration file:          flash:/vrpcfg.zip
  Next startup saved-configuration file:     flash:/huawei.zip
  Startup license file:                      null
  Next startup license file:                 null
  Startup patch package:                     null
  Next startup patch package:                null
  Startup voice-files:                       null
  Next startup voice-files:                  null
```


- **display startup**命令用来查看设备本次及下次启动相关的系统软件、备份系统软件、配置文件、License文件、补丁文件以及语音文件。
  - Startup system software表示的是本次系统启动所使用的VRP文件。
  - Next startup system software表示的是下次系统启动所使用的VRP文件。
  - Startup saved-configuration file表示的是本次系统启动所使用的配置文件。
  - Next startup saved-configuration file表示的是下次系统启动所使用的配置文件。
  - 设备启动时，会从存储设备中加载配置文件并进行初始化。如果存储设备中没有配置文件，设备将会使用默认参数进行初始化。
- **startup saved-configuration** [*configuration-file*] 命令用来指定系统下次启动时使用的配置文件，*configuration-file*参数为系统启动配置文件的名称。

更多信息



<Huawei>display configuration candidate
命令用来显示设备当前未提交的命令行信息。

当用户完成一系列配置后还未提交时,这些命令会放置在候选配置库。

<Huawei>display current-configuration
命令用来查看路由器当前生效的配置参数。

当用户将配置命令提交后,这些命令会放置在运行配置库。

<Huawei>display startup
命令用来显示与本次及下次启动相关的系统软件、配置文件名、PAF文件名和补丁文件名。

当用户将配置保存后,这些命令会放置在启动配置库。

VRP5操作系统只有运行配置库和启动配置库，配置命令后直接生效无需提交，而VRP8操作系统配置命令后需要提交才能生效。

思考题

1. 华为数通设备目前使用的VRP版本是多少?
2. 华为网络设备支持多少个用户同时使用Console口登录?
3. 如果设备中有多个配置文件，如何指定下次启动时使用的配置文件?

1. 目前，大多数华为数通产品使用的是VRP5版本，少数产品如NE系列路由器使用的是VRP8版本。
2. 华为网络设备同时只能有一个用户登录Console界面，因此Console用户的编号固定为0。
3. 需要指定某一配置文件为下次启动时使用的配置文件，可以执行startup saved-configuration configuration-file 命令，这里的配置文件名包括文件名称和扩展名。

本章总结

- VRP是华为公司具有完全自主知识产权的网络操作系统，可以运行在多种硬件平台之上。VRP拥有一致的网络界面、用户界面和管理界面，熟悉VRP命令行并且熟练掌握VRP配置是高效管理华为网络设备的必备基础。
- 在此基础上需要了解一些常用命令和快捷键的使用，快速掌握这些命令和快捷键。
- 学完本章节后，可以掌握VRP的基本概念，常用命令的作用和CLI界面的使用。


谢 谢
www.huawei.com

网络层协议及IP编址

前言

- IPv4 (Internet Protocol Version 4)协议族是TCP/IP协议族中最为核心的协议族。它工作在TCP/IP协议栈的网络层，该层与OSI参考模型的网络层相对应。
- 网络层提供了无连接数据传输服务，即网络在发送数据报文时不需要先建立连接，每一个IP数据报文独立发送。
- 在本章节中，将介绍IPv4地址的基本概念，介绍如何进行子网划分，并且会介绍网络IP地址规划和IP地址的基本配置。

目标

- 学完本课程后，您将能够：
  - 描述网络层的主要协议
  - 描述IPv4地址的概念、分类及特殊IP地址
  - 计算IP网络以及IP子网
  - 掌握IP网络地址规划方式

目录

网络层协议

- 网络层经常被称为IP层。但网络层协议并不只是IP协议，还包括ICMP（ Internet Control Message Protocol)协议、IPX（ Internet Packet Exchange ）协议等。

TCP/IP对等模型

-

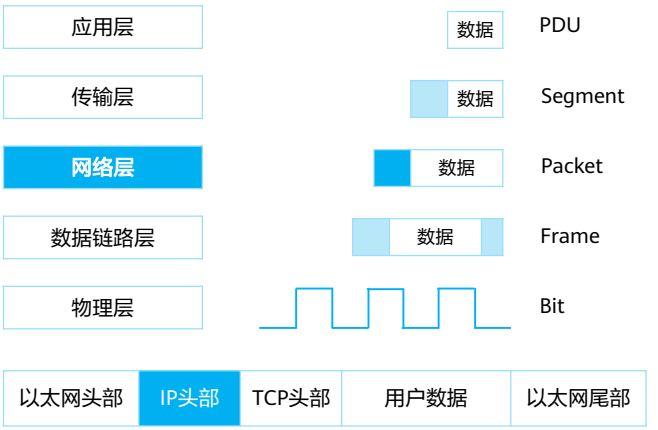
IP协议

- IP是Internet Protocol的缩写。 Internet Protocol本身是一个协议文件的名称，该协议文件的内容非常少，主要是定义并阐述了IP报文的格式。
- 经常被提及的IP，一般不是特指Internet Protocol这个协议文件本身，而是泛指直接或间接与IP协议相关的任何内容。

作用	版本
• 为网络层的设备提供逻辑地址 • 负责数据包的寻址和转发	• IPv4 (IP Version 4) • IPv6 (IP Version 6)

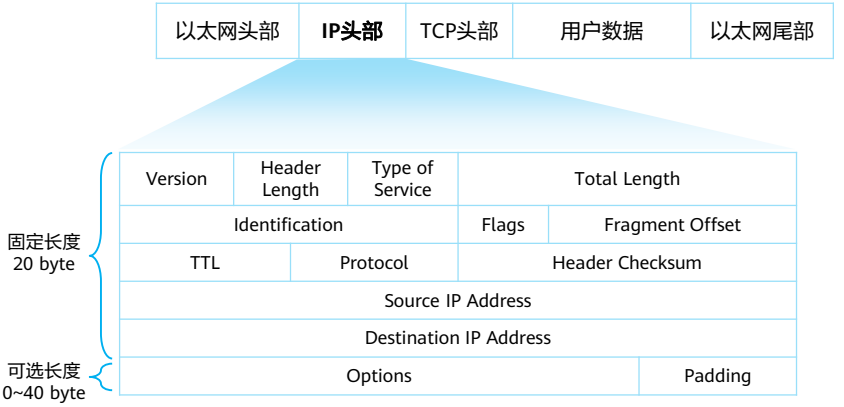
- IP协议有版本之分，分别是IPv4和IPv6。目前，Internet上的IP报文主要都是IPv4报文，但是逐步在向IPv6过渡。若无特别声明，本章所提及的IP均指IPv4。
  - IPv4（Internet Protocol Version 4）协议族是TCP/IP协议族中最为核心的协议族。它工作在TCP/IP协议栈的网络层，该层与OSI参考模型的网络层相对应。
  - IPv6（Internet Protocol Version 6）是网络层协议的第二代标准协议，也被称为IPng（IP Next Generation）。它是Internet工程任务组IETF（Internet Engineering Task Force）设计的一套规范，是IPv4（Internet Protocol Version 4）的升级版本。

数据封装



- 应用数据需要经过TCP/IP每一层处理之后才能通过网络传输到目的端，每一层上都使用该层的协议数据单元PDU（Protocol Data Unit）彼此交换信息。不同层的PDU中包含有不同的信息，因此PDU在不同层被赋予了不同的名称。
  - 如上层数据在传输层添加TCP报头后得到的PDU被称为Segment（数据段）；数据段被传递给网络层，网络层添加IP报头得到的PDU被称为Packet（数据包）；数据包被传递到数据链路层，封装数据链路层报头和尾部得到的PDU被称为Frame（数据帧）；最后，帧被转换为比特，通过网络介质传输。
  - 这种协议栈逐层向下传递数据，并添加报头和报尾的过程称为封装。
- 本章节我们主要讨论数据在网络层的封装，如果封装为IP协议，则被称为IP Packet（IP数据包）。

IPv4报文格式



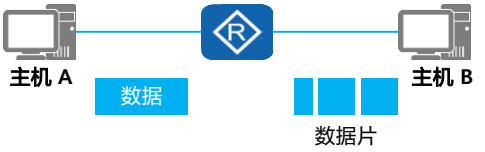
HUAWEI

- IP Packet（IP数据包），其包头主要内容如下：
  - Version：4 bit，4：表示为IPv4；6：表示为IPv6。
  - Header Length：4 bit，首部长度，如果不带Option字段，则为20，最长为60。
  - Type of Service：8 bit，服务类型。只有在有QoS差分服务要求时，这个字段才起作用。
  - Total Length：16 bit，总长度，整个IP数据包的长度。
  - Identification：16 bit，标识，分片重组时会用到该字段。
  - Flags：3 bit，标志位。
  - Fragment Offset：13 bit，片偏移，分片重组时会用到该字段。
  - Time to Live：8 bit，生存时间。
  - Protocol：8 bit，协议：下一层协议。指出此数据包携带的数据使用何种协议，以便目的主机的IP层将数据部分上交给哪个进程处理。
    - 常见值：
      - 1: ICMP, Internet Control Message；
      - 2: IGMP, Internet Group Management；
      - 6: TCP , Transmission Control Protocol；
      - 17: UDP, User Datagram Protocol。
  - Header Checksum：16 bit，首部检验和。
  - Source IP Address：32 bit，源IP地址。
  - Destination IP Address：32 bit，目的IP地址。
  - Options：可变，选项字段。
  - Padding：可变，填充字段，全填0。

数据包分片

- 将报文分割成多个片段的过程叫做分片。
- 网络中转发的IP报文的长度可以不同，但如果报文长度超过了数据链路所支持的最大长度，则报文就需要分割成若干个较小的片段才能够在链路上传输。

Version	Header Length	Type of Service	Total Length	
Identification			**Flags**	**Fragment Offset**
TTL	Protocol	Header Checksum		
Source IP Address				
Destination IP Address				
Options				Padding



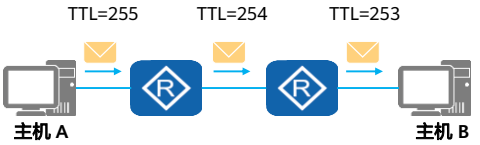
HUAWEI

- Identification：16 bit，发送主机赋予的标识，分片重组时会用到该字段。
- Flags：3 bit，标志位。
  - 保留段位：0，保留。
  - 不分段位：1，表示“不能分片”；0，表示“能分片”。
  - 更多段位：1，表示“后面还有分片”；0，表示“最后一个数据片”。
- Fragment Offset：13 bit，片偏移，分片重组时会用到该字段。指出较长的分组在分片后，该片在原分组中的相对位置，与更多段位组合，帮助接收方组合分段的报文。

生存时间 (Time to Live, TTL)

- TTL字段设置了数据包可以经过的路由器数目。
- 一旦经过一个路由器，TTL值就会减1，当该字段值为0时，数据包将被丢弃。

Version	Header Length	Type of Service	Total Length	
Identification			Flags	Fragment Offset
TTL	Protocol		Header Checksum	
Source IP Address				
Destination IP Address				
Options				Padding

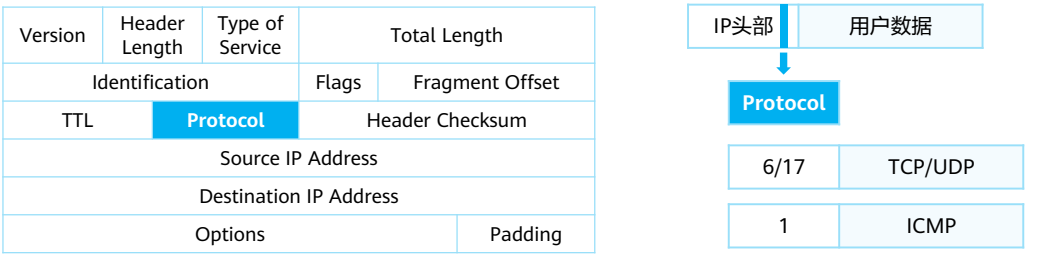


HUAWEI

- Time to Live：8 bit，生存时间。可经过的最多路由数，即数据包在网络中可通过的路由器数的最大值。
  - 报文在网段间转发时，如果网络设备上的路由规划不合理，就可能会出现环路，导致报文在网络中无限循环，无法到达目的端。环路发生后，所有发往这个目的地的报文都会被循环转发，随着这种报文逐渐增多，网络将会发生拥塞。
  - 为避免环路导致的网络拥塞，IP报文头中包含一个生存时间TTL（Time To Live）字段。报文每经过一台三层设备，TTL值减1。初始TTL值由源端设备设置。当报文中的TTL降为0时，报文会被丢弃。同时，丢弃报文的设备会根据报文头中的源IP地址向源端发送ICMP错误消息。（注意：网络设备也可被配置为不向源端发送ICMP错误消息。）

协议号 (Protocol)

- IP报文头中的协议号字段标识了将会继续处理该报文的协议。
- 即指出此数据包携带的数据使用何种协议，以便目的主机的IP层将数据部分上报给哪个进程处理。



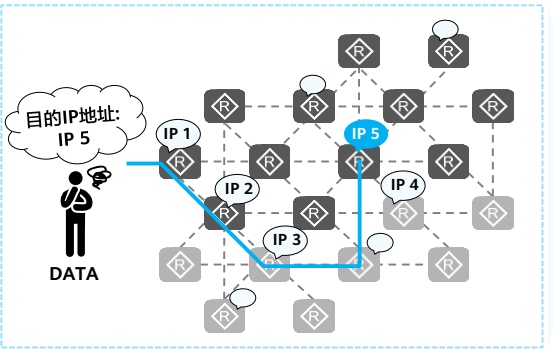
HUAWEI

- 目的端的网络层在接收并处理报文以后，需要决定下一步对报文如何处理。IP报文头中的协议字段标识了将会继续处理报文的协议。
- 该字段可以标识网络层协议，如ICMP（Internet Control Message Protocol，因特网控制报文协议，对应值0x01）；也可以标识上层协议，如TCP（Transmission Control Protocol，传输控制协议，对应值0x06）、UDP（User Datagram Protocol，用户数据包协议，对应值0x11）。

目录

什么是IP地址

- IP地址在网络中用于标识一个节点（或者网络设备的接口）。
- IP地址用于IP报文在网络中的寻址。



IP地址

IP地址就像现实中的地址，可以标识网络中的一个节点，数据就是通过它来找到目的地。

HUAWEI

- 在IP网络上，如果用户要将一台计算机连接到Internet上，就需要申请一个IP地址。IP地址就像现实中的地址，可以标识网络中的一个节点，数据就是通过它来找到目的地的。即我们通过IP地址实现全球范围内的网络通信。
- IP地址是网络设备接口的属性，不是网络设备本身的属性。当我们说给某台设备分配一个IP地址时，实质上是指给这台设备的某个接口分配一个IP地址。如果设备有多个接口，通常每个接口都至少需要一个IP地址。

- 注：需要使用IP地址的接口，通常是路由器和计算机的接口。

概念 地址分类 地址计算 特殊地址 IPv4 vs IPv6

IP地址表示

- 一个IPv4地址有32 bit。
- IPv4地址通常采用“点分十进制”表示。

点分十进制表示法

十进制	**192.**	**168.**	**10.**	**1**	4 byte
二进制	**11000000**	10101000	00001010	00000001	32 bit（32位）

十进制与二进制的转换

幂	2^7	2^6	2^5	2^4	2^3	2^2	2^1	2^0
	128	64	32	16	8	4	2	1
位	1	1	0	0	0	0	0	0

= 128 + 64 = 192

- IPv4地址范围：0.0.0.0~255.255.255.255。

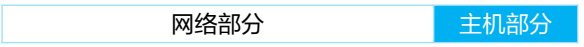
HUAWEI

- IP地址表示
  - IP地址是长度是32 bit，由4个字节组成。为了阅读和书写方便，IP地址通常采用点分十进制数来表示。
- 点分十进制表示法
  - IP地址表现形式能够帮助我们更好的使用和配置网络，但通信设备在对IP地址进行计算时使用的是二进制的操作方式，因此掌握十进制、二进制的转换运算非常有必要。
- IPv4地址范围
  - 00000000.00000000.00000000.00000000~11111111.11111111.11111111.11111111，即0.0.0.0~255.255.255.255。

概念 地址分类 地址计算 特殊地址 IPv4 vs IPv6

IP地址构成

- **网络部分：**用来标识一个网络。
- **主机部分：**用来区分一个网络内的不同主机。



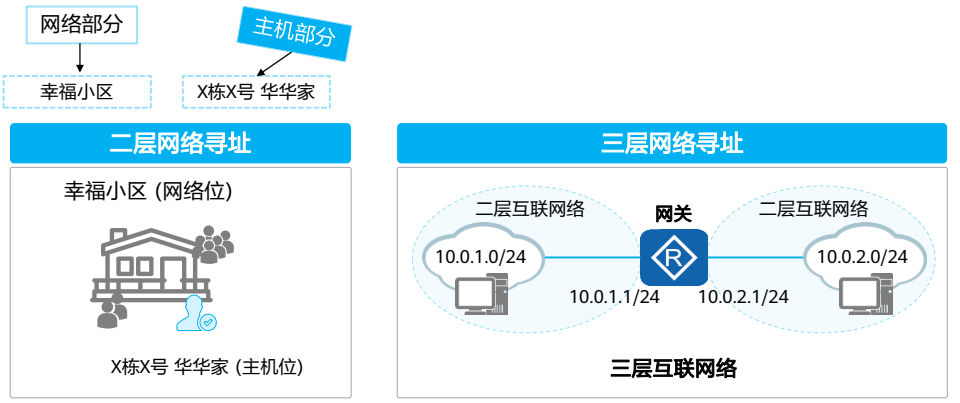
- **网络掩码：**区分一个IP地址中的网络部分及主机部分。

HUAWEI

- IPv4地址由如下两部分组成：
  - 网络部分 (网络号)：用来标识一个网络。
    - IP地址不能反映任何有关主机位置的地理信息，只能通过网络号码字段判断出主机属于哪个网络。
    - 对于网络号相同的设备，无论实际所处的物理位置如何，它们都是处在同一个网络中。
  - 主机部分 (主机号)：用来区分一个网络内的不同主机。
- 网络掩码 (Netmask)，又称子网掩码 (Subnet Mask)：
  - 网络掩码为32 bit，与IP地址的位数一样，通常也以点分十进制数来表示。
  - 网络掩码不是一个IP地址，在二进制的表示上是一堆连续的1、后面接一堆连续的0。
  - 通常将网络掩码中1的个数称为这个网络掩码的长度。如：掩码0.0.0.0的长度是0，掩码252.0.0.0的长度是6。
  - 网络掩码一般与IP地址结合使用，其中值为1的比特对应IP地址中的网络位；值为0的比特对应IP地址中的主机位，以此来辅助我们识别一个IP地址中的网络位与主机位。即网络掩码中1的个数就是IP地址的网络号的位数，0的个数就是IP地址的主机号的位数。

IP地址寻址

- **网络部分：**用来标识一个网络，代表IP地址所属网络。
- **主机部分：**用来区分一个网络内的不同主机，能唯一标识网段上的某台设备。



- 网络号用于表示主机所在的网络，类似于“XX省XX市XX区XX小区”的作用。
- 主机号用于表示网络号所定义的网络范围内某个特定的主机接口，类似于门牌号“XX栋XX号”的作用。
- 网络寻址：
  - 二层网络寻址：可直接通过IP地址，找到对应的主机接口。
  - 三层网络寻址：利用网关转发来自不同网段之间的数据包。
- 网关：
  - 报文转发过程中，首先需要确定转发路径以及通往目的网段的接口。如果目的主机与源主机不在同一网段，报文需要先转发到网关，然后通过网关将报文转发到目的网段。
  - 网关是指接收并处理本地网段主机发送的报文并转发到目的网段的设备。为实现此功能，网关必须知道目的网段的路由。网关设备上连接本地网段的接口地址即为该网段的网关地址。

IP地址分类 (有类编址)

- 为了方便IP地址的管理及组网，IP地址分成五类：

- A/B/C类默认网络掩码
  - A类：8 bit， 0.0.0.0~127.255.255.255/8
  - B类：16 bit，128.0.0.0~191.255.255.255/16
  - C类：24 bit，192.0.0.0~223.255.255.255/24

网络部分
主机部分

- 为了方便IP地址的管理及组网，IP地址分成五类：
  - A、B、C、D、E类的类别字段分别是二进制数0、10、110、1110、1111，通过网络号码字段的前几个比特就可以判断IP地址属于哪一类，这是区分各类地址最简单的方法。
  - A、B、C三类地址是单播IP地址 (除一些特殊地址外)，只有这三类地址才能分配给主机接口使用。
  - D类地址属于组播IP地址。
  - E类地址专门用于特殊的实验目的。
  - 本节内容，只关注A、B、C三类地址。
- A、B、C类地址比较：
  - 使用A类地址的网络称为A类网络；使用B类地址的网络称为B类网络；使用C类地址的网络称为C类网络。
  - A类网络的网络号为8 bit，个数很少，但所允许的主机接口的个数很多；首位恒定为0，地址空间为：0.0.0.0~127.255.255.255。
  - B类网络的网络号为16 bit，介于A类和C类网络之间；首两位恒定为10，地址空间为：128.0.0.0~191.255.255.255。
  - C类网络的网络号为24 bit，个数很多，但所允许的主机接口的个数就很少；首三位恒定为110，地址空间为：192.0.0.0~223.255.255.255。
- 注：
  - 主机 (Host)，通常指路由器和计算机的统称。并且常把主机的某个接口的IP地址简称为主机IP地址。
  - 组播地址：组播能实现一对多传递消息。

概念 地址分类 地址计算 特殊地址 IPv4 vs IPv6

IP地址类型

- 我们通常把一个网络号所定义的网络范围称为一个网段。
- **网络地址**：用于标识一个网络。

例如：192.168.10.0/24

192.	168.	10.	00000000

- **广播地址**：用于向该网络中的所有主机发送数据的特殊地址。

例如：192.168.10.255/24

192.	168.	10.	11111111

- **可用地址**：可分配给网络中的节点或网络设备接口的地址。

例如：192.168.10.1/24

192.	168.	10.	00000001

注意

- 网络地址和广播地址不能直接被节点或网络设备所使用。
- 一个网段可用地址数量为：2^n-2（n：主机部分的比特位数）

HUAWEI

- 网络地址
  - 网络号为X，主机号的每个比特都为0。
  - 不能分配给具体的主机接口使用。
- 广播地址
  - 网络号为X，主机号的每个比特都为1。
  - 不能分配给具体的主机接口使用。
- 可用地址
  - 又称主机地址，可用分配给具体的主机接口使用。
- 一个网段可用地址数量计算：
  - 一个网段的主机位为n位，则IP地址数为：2^n，可用IP地址数为：2^n-2 (减去网络地址和广播地址)。

IP地址计算

- 例：172.16.10.1/16这个B类地址的网络地址、广播地址以及可用地址数分别是?

	172.	16.	00001010.	00000001
IP地址	10101100	00010000	00001010	00000001
网络掩码	11111111	11111111	00000000	00000000
网络地址	10101100	00010000	00000000	00000000
广播地址	10101100	00010000	11111111	11111111

主机位全为0，得出网络地址 172.16.0.0

主机位全为1，得出广播地址 172.16.255.255

IP地址数 2^{16}=65536

可用IP地址数 2^{16}-2=65534

可用IP地址范围 172.16.0.1~172.16.255.254

练习

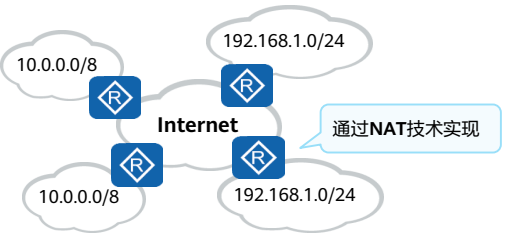
请计算10.128.20.10/8这个A类地址的网络地址、广播地址以及可用地址数。

- 网络地址：将IP地址的主机位全设为0，所得结果是该IP地址所在网络的网络地址。
- 广播地址：将IP地址的主机位全设为1，所得结果是该IP地址所在网络的广播地址。
- IP地址数：2^n，n为主机位位数。
- 可用IP地址数：2^n-2，n为主机位位数。

- 练习题答案：
  - 网络地址：10.0.0.0
  - 广播地址：10.255.255.255
  - IP地址数：2^{24}
  - 可用IP地址数：2^{24}-2
  - 可用IP地址范围：10.0.0.1~10.255.255.254

私网IP地址

- **公网IP地址：**IP地址是由IANA统一分配的，以保证任何一个IP地址在Internet上的唯一性。这里的IP地址是指公网IP地址。
- **私网IP地址：**实际上一些网络不需要连接到Internet，比如一个大学的封闭实验室内的网络，只要同一网络中的网络设备的IP地址不冲突即可。在IP地址空间里，A、B、C三类地址中各预留了一些地址专门用于上述情况，称为私网IP地址。
  - A类：10.0.0.0~10.255.255.255
  - B类：172.16.0.0~172.31.255.255
  - C类：192.168.0.0~192.168.255.255



私有网络连接到Internet

- 为了解决IP地址短缺的问题，提出了私有地址的概念。私有地址是指内部网络或主机地址，这些地址只能用于某个内部网络，不能用于公共网络。
  - 公网IP地址：连接到Internet的网络设备必须具有由ICANN分配的公网IP地址。
  - 私网IP地址：私网IP地址的使用使得网络可以得到更为自由地扩展，因为同一个私网IP地址是可以在不同的私有网络中重复使用的。
- 私有网络连接到Internet：私有网络由于使用了私网IP地址，是不允许连接到Internet的。后来在实际需求的驱动下，许多私有网络也希望能够连接到Internet上，从而实现私网与Internet之间的通信，以及通过Internet实现私网与私网之间的通信。私网与Internet的互联，必须使用网络地址转换 (NAT)技术实现。
- 注：
  - NAT (Network Address Translation)，网络地址转换，其基本作用是实现私网IP地址与公网IP地址之间的转换。
  - IANA (Internet Assigned Numbers Authority)，因特网地址分配组织。

特殊IP地址

- IP地址空间中，有一些特殊的IP地址，这些IP地址有特殊的含义和作用，举例如下。

特殊IP地址	地址范围	作用
有限广播地址	255.255.255.255	可作为目的地址，发往该网段所有主机（受限于网关）
任意地址	0.0.0.0	“任何网络”的网络地址； “这个网络上这个主机接口”的IP地址
环回地址	127.0.0.0/8	测试设备自身的软件系统
本地链路地址	169.254.0.0/24	当主机自动获取地址失败后，可使用该网段中的某个地址进行临时通信

- 255.255.255.255
  - 这个地址称为有限广播地址，它可以作为一个IP报文的目的IP地址使用。
  - 路由器接收到目的IP地址为有限广播地址的IP报文后，会停止对该IP报文的转发。
- 0.0.0.0
  - 如果把这个地址作为网络地址，它的意思就是“任何网络”的网络地址；如果把这个地址作为主机接口地址，它的意思就是“这个网络上主机接口”的IP地址。
  - 例如：当一个主机接口在启动过程中尚未获得自己的IP地址时，就可以向网络发送目的IP地址为有限广播地址、源IP地址为0.0.0.0的DHCP请求报文，希望DHCP服务器在收到自己的请求后，能够给自己分配一个可用的IP地址。
- 127.0.0.0/8
  - 这个地址为环回地址，它可以作为一个IP报文的目的IP地址使用。其作用是测试设备自身的软件系统。
  - 一个设备产生的、目的IP地址为环回地址的IP报文是不可能离开这个设备本身的。
- 169.254.0.0/16
  - 如果一个网络设备获取IP地址的方式被设置成了自动获取方式，但是该设备在网络上又没有找到可用的DHCP服务器，那么该设备就会使用169.254.0.0/16网段的某个地址来进行临时通信。
- 注：DHCP (Dynamic Host Configuration Protocol)，动态主机配置协议，用于动态分配网络配置参数，如IP地址。

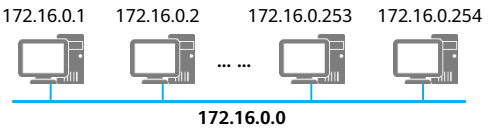
IPv4 vs IPv6

- 由全球IP地址分配机构，IANA (Internet Assigned Numbers Authority)管理的IPv4地址，于2011年完全用尽。随着最后一个IPv4公网地址分配完毕，加上接入公网的用户及设备越来越多，IPv4地址枯竭的问题日益严重，这是当前IPv6替代IPv4的最大源动力。

IPv4	IPv6
• 地址长度：32 bit • 地址分类：单播地址、广播地址、组播地址 • 特点： ▫ 地址枯竭 ▫ 包头设计不合理 ▫ 对ARP的依赖，导致广播泛滥 ▫	• 地址长度：128 bit • 地址分类：单播地址、组播地址、任播地址 • 特点： ▫ 无限地址 ▫ 简化的报文头部 ▫ IPv6地址自动部署 ▫ .....

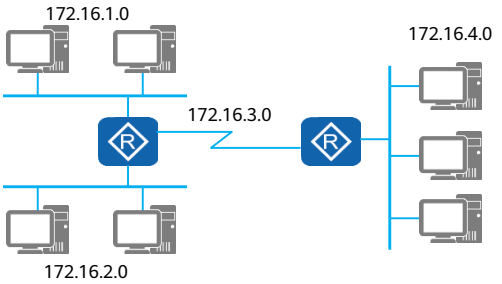
目录

为什么要划分子网



2^{16}=65536个IP地址

- 一个B类地址用于一个广播域，地址浪费。
- 广播域太庞大，一旦发生广播，内网不堪重负。



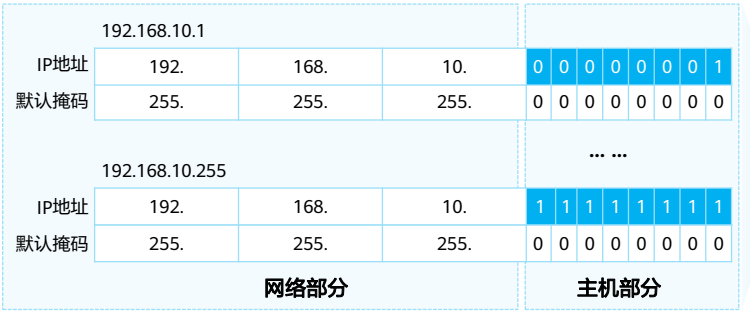
- 将一个网络号划分成多个子网，每个子网分配给一个独立的广播域。
- 如此一来广播域的规模更小、网络规划更加合理。
- IP地址得到了合理利用。

HUAWEI

- “有类编址”的地址划分过于死板，划分的颗粒度太大，会有大量的主机号不能被充分利用，从而造成了大量的IP地址资源浪费。
- 因此可以利用子网划分来减少地址浪费，即VLSM (Variable Length Subnet Mask)，可变长子网掩码。将一个大的有类网络，划分成若干个小的子网，使得IP地址的使用更为科学。

如何进行子网划分 - 原网段分析

- 例如：192.168.10.0/24网段



1个C类网络：192.168.10.0/24
默认掩码：255.255.255.0

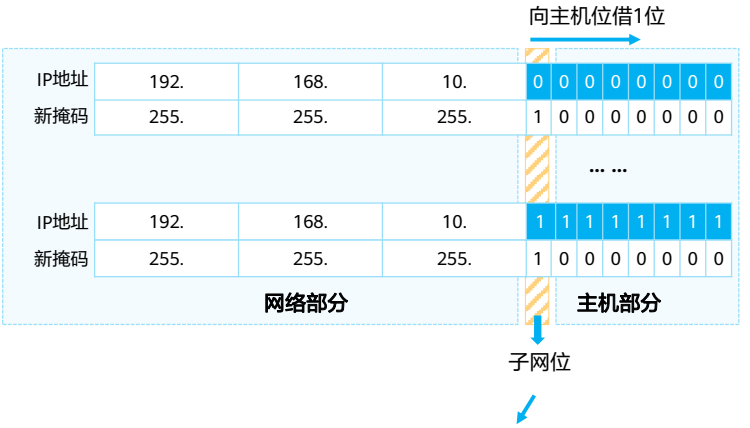
网络地址：192.168.10.0/24
广播地址：192.168.10.255
IP地址数： 2^8=256个
可用IP地址数：2^8-2=254个

HUAWEI

- 假设有一个C类网段地址：192.168.10.0；默认情况下，网络掩码为24位，包括24位网络位，8位主机位。
- 通过计算可知，这样的网络中，有256个IP地址。

如何进行子网划分 - 向主机借位

- 向主机借位，形成子网。



2个子网：
子网1：192.168.10.0/25
子网2：192.168.10.128/25
新掩码：255.255.255.128

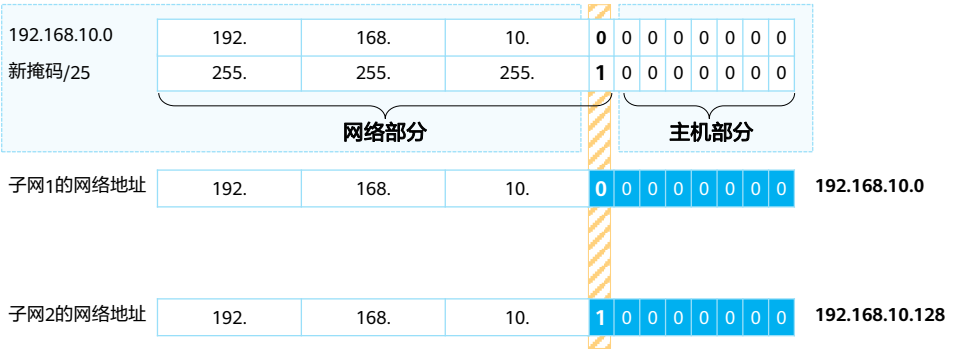
IP地址数： 2^7=128个
可用IP地址数：2^7-2=126个

- 可变长子网掩码，VLSM (Variable Length Subnet Mask)

- 现在，将原有的24位网络位向主机位去“借”1位，这样网络位就扩充到了25位，相对的主机位就减少到了7位，而借过来的这1位就是子网位，此时网络掩码就变成了25位，即255.255.255.128，或/25。
- 子网位：可取值0或取值1，则得到了两个新的子网。
- 通过计算可知，现在网络中，有128个IP地址。

如何进行子网划分 - 计算子网网络地址

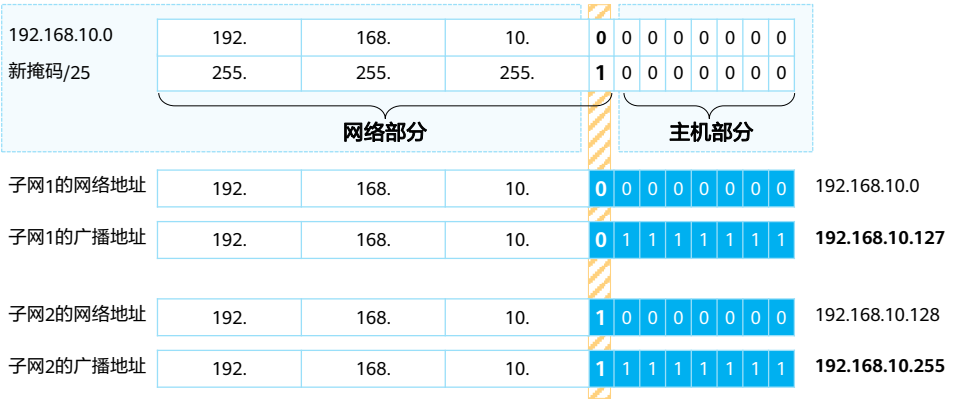
- 主机位全为0，计算子网网络地址。



- 计算网络地址，主机位全为0：
  - 如果子网位取值0，则网络地址为192.168.10.0。
  - 如果子网位取值1，则网络地址为192.168.10.128。

如何进行子网划分 - 计算子网广播地址

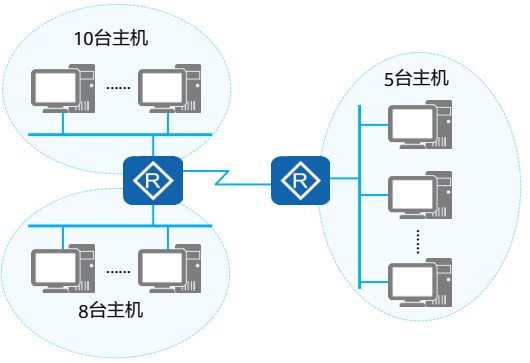
- 主机位全为1，计算子网广播地址。



HUAWEI

- 计算广播地址，主机位全为1：
  - 如果子网位取值0，则广播地址为192.168.10.127。
  - 如果子网位取值1，则广播地址为192.168.10.255。

练习：计算子网 (1)



- **问题：**现有一个C类网络地址段192.168.1.0/24，请使用可变长子网掩码给三个子网分别分配IP地址。

- **计算：**（以10台主机为例）

  步骤1：计算所需主机位

  $2^n-2 \geq 10$

  $n \geq 4$位，主机位

  步骤2：向主机位借位

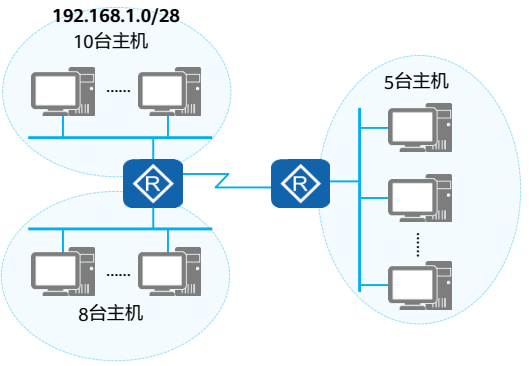
向主机位借4位

IP地址	192.	168.	1.	**0**	**0**	**0**	**0**	0	0	0	0
子网掩码	255.	255.	255.	1	1	1	1	0	0	0	0

子网位　子网数：2^4=16个子网

- 实际网络规划中，会先规划主机多的子网络。

练习：计算子网 (2)



- **问题：**现有一个C类网络地址段192.168.1.0/24，请使用可变长子网掩码给三个子网分别分配IP地址。
- **计算：**（以10台主机为例）

步骤3：计算子网网络地址

IP地址	192.	168.	1.	**0**	**0**	**0**	**0**	0	0	0	0
新掩码	255.	255.	255.	1	1	1	1	0	0	0	0
											网络地址
子网1	192.	168.	1.	0	0	0	0	0	0	0	0	192.168.1.0/28
子网2	192.	168.	1.	0	0	0	1	0	0	0	0	192.168.1.16/28
子网3	192.	168.	1.	0	0	1	0	0	0	0	0	192.168.1.32/28
⋮											
子网16	192.	168.	1.	1	1	1	1	0	0	0	0	192.168.1.240/28

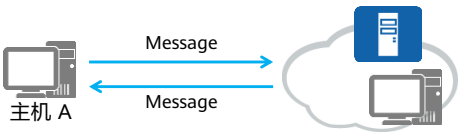
HUAWEI

- 子网的网络地址分别为：
  - 192.168.1.0
  - 192.168.1.16
  - 192.168.1.32
  - 192.168.1.48
  - 192.168.1.64
  - 192.168.1.80
  - 192.168.1.96
  - 192.168.1.112
  - 192.168.1.128
  - 192.168.1.144
  - 192.168.1.160
  - 192.168.1.176
  - 192.168.1.192
  - 192.168.1.208
  - 192.168.1.224
  - 192.168.1.240

目录

ICMP协议

- Internet控制消息协议ICMP (Internet Control Message Protocol)是IP协议的辅助协议。
- ICMP协议用来在网络设备间传递各种差错和控制信息，对于收集各种网络信息、诊断和排除各种网络故障等方面起着至关重要的作用。



以太网头部	IP头部	**ICMP报文**	以太网尾部

Type	**Code**	Checksum
ICMP的报文内容		

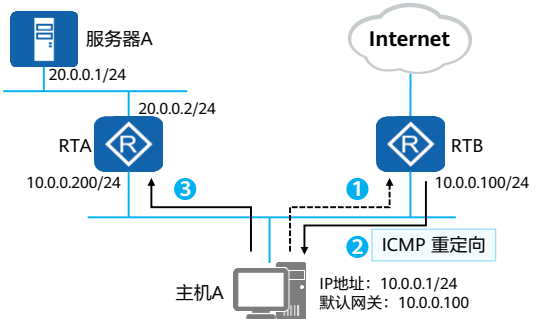
Type	Code	描述
0	0	Echo Reply
3	0	网络不可达
3	1	主机不可达
3	2	协议不可达
3	3	端口不可达
5	0	重定向
8	0	Echo Request

HUAWEI

- 为了更有效地转发IP数据报文和提高数据报文交互成功的机会，在网络层使用ICMP协议。ICMP允许主机或设备报告差错情况和提供有关异常情况的报告。
- ICMP消息：
  - ICMP消息封装在IP报文中，IP报文头部Protocol值为1时表示ICMP协议。
  - 字段解释：
    - ICMP消息的格式取决于Type和Code字段，其中Type字段为消息类型，Code字段包含该消息类型的具体参数。
    - 校验和字段用于检查消息是否完整。
    - 消息中包含32 bit的可变参数，这个字段一般不使用，通常设置为0。
      - 在ICMP重定向消息中，这个字段用来指定网关IP地址，主机根据这个地址将报文重定向到指定网关。
      - 在Echo请求消息中，这个字段包含标识符和序号，源端根据这两个参数将收到的回复消息与本端发送的Echo请求消息进行关联。尤其是当源端向目的端发送了多个Echo请求消息时，需要根据标识符和序号将Echo请求和回复消息进行一一对应。

ICMP重定向

- ICMP重定向报文是ICMP控制报文中的一种。在特定的情况下，当路由器检测到一台机器使用非最优路由的时候，它会向该主机发送一个ICMP重定向报文，请求主机改变路由。

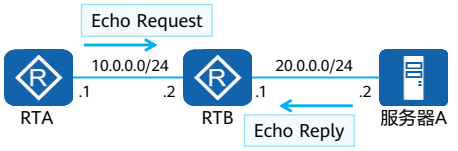


HUAWEI

- ICMP重定向过程：
  1. 主机A希望发送报文到服务器A，于是根据配置的默认网关地址向网关RTB发送报文。
  2. 网关RTB收到报文后，检查报文信息，发现报文应该转发到与源主机在同一网段的另一个网关设备RTA，此转发路径是更优的路径，所以RTB会向主机发送一个Redirect消息，通知主机直接向另一个网关RTA发送该报文。
  3. 主机收到Redirect消息后，会向RTA发送报文，然后RTA会将该报文再转发给服务器A。

ICMP差错检测

- ICMP Echo消息常用于诊断源和目的地之间的网络连通性，同时还可以提供其他信息，如报文往返时间等。



功能：Ping

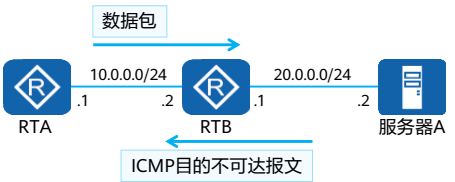
Ping是网络设备、Windows、Unix和Linux平台上的一个命令，其实是一个小巧而实用的应用程序，该应用基于ICMP协议。Ping常用于探测到达目的节点的网络可达性。

```
[RTA]ping 20.0.0.2
  PING 20.0.0.2: 56  data bytes, press CTRL_C to break
    Reply from 20.0.0.2: bytes=56 Sequence=1 ttl=254 time=70 ms
    Reply from 20.0.0.2: bytes=56 Sequence=2 ttl=254 time=30 ms
    Reply from 20.0.0.2: bytes=56 Sequence=3 ttl=254 time=30 ms
    Reply from 20.0.0.2: bytes=56 Sequence=4 ttl=254 time=40 ms
    Reply from 20.0.0.2: bytes=56 Sequence=5 ttl=254 time=30 ms

  --- 20.0.0.2 ping statistics ---
    5 packet(s) transmitted
    5 packet(s) received
    0.00% packet loss
    round-trip min/avg/max = 30/40/70 ms
```


- ICMP的一个典型应用是Ping。Ping是检测网络连通性的常用工具，同时也能够收集其他相关信息。用户可以在Ping命令中指定不同参数，如ICMP报文长度、发送的ICMP报文个数、等待回复响应的超时时间等，设备根据配置的参数来构造并发送ICMP报文，进行Ping测试。

ICMP错误报告

- ICMP定义了各种错误消息，用于诊断网络连接性问题；根据这些错误消息，源设备可以判断出数据传输失败的原因。如：当网络设备无法访问目标网络时，会自动发送ICMP目的不可达报文到发送端设备。




```
[RTA]tracert 20.0.0.2

 traceroute to  20.0.0.2(20.0.0.2), max hops: 30 ,packet length:
40,press CTRL_C
 to break

 1 10.0.0.2              80 ms     10 ms     10 ms

 2 20.0.0.2              30 ms     30 ms     20 ms
```

功能：Tracert

Tracert基于报文头中的TTL值来逐跳跟踪报文的转发路径。Tracert是检测网络丢包和时延的有效手段，同时可以帮助管理员发现网络中的路由环路。

- ICMP定义了各种错误消息，用于诊断网络连接性问题；根据这些错误消息，源设备可以判断出数据传输失败的原因。
  - 如果网络中发生了环路，导致报文在网络中循环，且最终TTL超时，这种情况下网络设备会发送TTL超时消息给发送端设备。
  - 如果目的地不可达，则中间的网络设备会发送目的不可达消息给发送端设备。目的不可达的情况有多种，如果是网络设备无法找到目的网络，则发送目的网络不可达消息；如果网络设备无法找到目的网络中的目的主机，则发送目的主机不可达消息。
- ICMP的另一个典型应用是Tracert。Tracert基于报文头中的TTL值来逐跳跟踪报文的转发路径。为了跟踪到达某特定目的地址的路径，源端首先将报文的TTL值设置为1。该报文到达第一个节点后，TTL超时，于是该节点向源端发送TTL超时消息，消息中携带时间戳。然后源端将报文的TTL值设置为2，报文到达第二个节点后超时，该节点同样返回TTL超时消息，以此类推，直到报文到达目的地。这样，源端根据返回的报文中的信息可以跟踪到报文经过的每一个节点，并根据时间戳信息计算往返时间。

目录

IP地址的基础配置命令

1. 进入接口视图

```
[Huawei] interface interface-type interface-number
```

通过此命令可以进入指定的接口视图，配置接口的相关属性。

- *interface-type interface-number*：指定接口类型和接口编号。接口类型和接口编号之间可以输入空格也可以不输入空格。

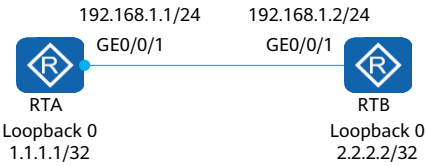
2. 配置接口的IP地址

```
[Huawei-GigabitEthernet0/0/1] ip address ip-address { mask | mask-length }
```

在接口视图下，通过此命令来给网络设备上的接口配置IP地址，实现网络的互连。

- *ip-address*：指定接口的IP地址，点分十进制形式。
- *mask*：指定子网掩码，点分十进制形式。
- *mask-length*：指定掩码长度，整数形式，取值范围是0～32。

案例：配置接口IP地址



在上述两台路由器互联的网络中，配置设备的互联物理接口地址以及各自的逻辑地址。

配置物理接口地址：

```
[RTA] interface gigabitethernet 0/0/1
[RTA-GigabitEthernet0/0/1] ip address 192.168.1.1 255.255.255.0
                或
[RTA-GigabitEthernet0/0/1] ip address 192.168.1.1 24
```

配置逻辑接口地址：

```
[RTA] interface LoopBack 0
[RTA-LoopBack0] ip address 1.1.1.1 255.255.255.255
                或
[RTA-LoopBack0] ip address 1.1.1.1 32
```


HUAWEI

- 物理接口：物理接口是指网络设备上实际存在的接口，分为负责承担业务传输的业务接口和负责管理设备的管理接口，例如GE业务接口和MEth管理接口。
- 逻辑接口：逻辑接口是指能够实现数据交换功能但物理上不存在、需要通过配置建立的接口，需要承担业务传输，例如VLANIF接口、Loopback接口。
  - Loopback接口：用户需要一个接口状态永远是Up的接口的IP地址时，可以选择Loopback接口的IP地址。
    - Loopback接口一旦被创建，其物理状态和链路协议状态永远是Up，即使该接口上没有配置IP地址。
    - Loopback接口配置IP地址后，就可以对外发布。Loopback接口上可以配置32位掩码的IP地址，达到节省地址空间的目的。
    - Loopback接口不能封装任何链路层协议，数据链路层也就不存在协商问题，其协议状态永远都是Up。
    - 对于目的地址不是本地IP地址，出接口是本地Loopback接口的报文，设备会将其直接丢弃。

网络IP地址规划

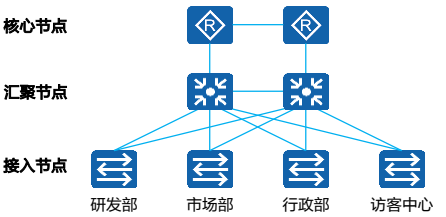
- IP地址规划要和网络结构、路由协议、流量规划、业务规则等结合起来考虑。IP地址的规划应尽可能和网络层次相对应，应该是自顶向下的一种规划。
- 总得来说： IP地址规划的目标是：易管理、易扩展、利用率高。

参考规划原则

唯一性、连续性、扩展性
结构化、业务相关性

- **IP地址规划范例**

背景	地址类型	地址范围
例如： 某公司被分配了192.168.0.0/16网段地址	研发部所属网段	192.168.1.0/24
	市场部所属网段	192.168.2.0/24
	行政部所属网段	192.168.3.0/24
	访客中心所属网段	192.168.4.0/24
	其他	... ...



HUAWEI

- 规划原则：
  - 唯一性：一个IP网络中不能有两个主机采用相同的IP地址。
  - 连续性：连续地址在层次结构网络中易于进行路由汇总，大大缩减路由表，提高路由计算的效率、加速路由收敛。
  - 扩展性：地址分配在每一层次上都要有合理的预留，在网络规模扩展时能保证路由汇总所需的连续性。避免网络扩展造成的地址、路由重新规划。
  - 结构化、业务相关性：地址规划与网络拓扑结构和网络承载业务结合起来，便于路由规划和QoS部署。好的IP地址规划使得每个地址都具有实际含义，看到一个地址就可以大致判断出该地址所属的设备和对应的业务。

思考题

1. （单选）201.222.5.64是第几类IP地址。（　）
   A. A类
   B. B类
   C. C类
   D. D类
2. （多选）某公司被分到了一个C类网络地址段192.168.20.0/24，现有一个部门有40台主机，请问以下哪些子网可被分配(　)?
   A. 192.168.20.64/26
   B. 192.168.20.64/27
   C. 192.168.20.128/26
   D. 192.168.20.190/26

1. C
2. AC

本章总结

- 在IP网络上，如果用户要将一台计算机连接到Internet上，就需要向因特网服务提供方ISP（Internet Service Provider）申请一个IP地址。
- 在本章节中，我们介绍了IP协议的基本概况，并介绍了IPv4地址的相关概念以及如何进行子网划分。
- 在本章节中，我们还介绍了网络IP地址规划以及IP地址的基本配置。


谢 谢
www.huawei.com

IP路由基础

前言

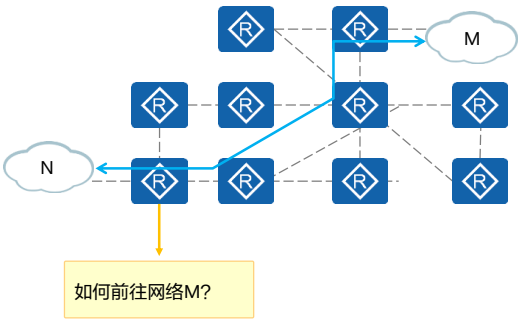
- 在一个典型的数据通信网络中，往往存在多个不同的IP网段，数据在不同的IP网段之间交互是需要借助三层设备的，这些设备具备路由能力，能够实现数据的跨网段转发。
- 路由是数据通信网络中最基本的要素。路由信息是指导报文转发的路径信息，路由过程就是报文转发的过程。
- 本课程将会向读者介绍路由的基本概念。

目标

- 学完本课程后，您将能够：
  - 了解路由器的基本工作原理
  - 掌握路由器选择最优路由的方法
  - 了解路由表的具体内容
  - 掌握路由转发高级特性

目录

背景：网段间通信



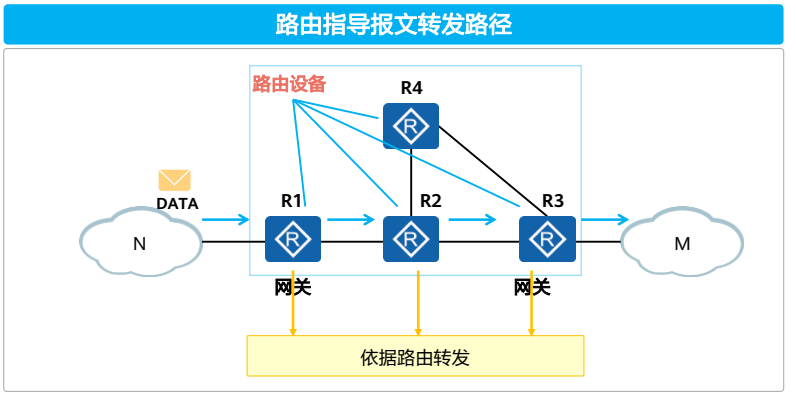
- IP地址唯一标识了网络中的一个节点，每个IP地址都拥有自己的网段，各个网段可能分布在网络的不同区域。
- 为实现IP寻址，分布在不同区域的网段之间要能够相互通信。

HUAWEI

- 通过IP地址能够寻找到一个唯一的网络节点，每个IP都有自己所属的网段，这些网络可能分布在世界各地，共同组成了全球的网络。
- 为了实现不同网段之间的相互通信，网络设备需要能够转发来自不同网段的IP报文，将其送达不同的IP网段。

路由

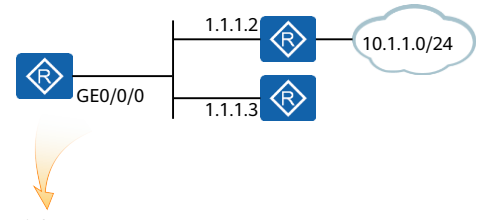
- 路由是指导报文转发的路径信息，通过路由可以确认转发IP报文的路径。
- 路由设备是依据路由转发报文到目的网段的网络设备，最常见的路由设备：路由器。
- 路由设备维护着一张路由表，保存着路由信息。



HUAWEI

- 网关以及中间节点（路由器）根据收到的IP报文其目的地址选择一条合适的路径，并将报文转发到下一个路由器。在路径中的最后一跳路由器二层寻址将报文转发给目的主机。这个过程被称为路由转发。
- 中间节点选择路径所依赖的表项为称为路由表。
- 路由条目包含明确的出接口以及下一跳，这两项信息指导IP报文转发到相应的下一跳设备上。

路由信息介绍



IP路由

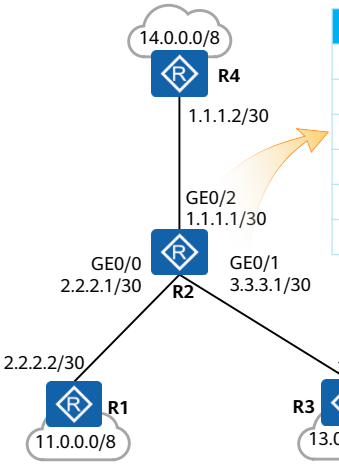
目的网络/掩码	出接口	下一跳
10.1.1.0/24	GE0/0/0	1.1.1.2

- 路由中包含以下信息：
  - 目的网络：标识目的网段
  - 掩码：与目的地址共同标识一个网段
  - 出接口：数据包被路由后离开本路由器的接口
  - 下一跳：路由器转发到达目的网段的数据包所使用的下一跳地址
- 这些信息标识了目的网段、明确了转发IP报文的路径。

HUAWEI

- 通过路由中包含的信息，路由设备可以转发IP报文到相应的路径。
- 目的地址、掩码用于识别IP报文目的地址，路由设备将IP报文匹配到相应的路由之后，根据路由的出接口、下一跳确认转发的路径。
- 只有出接口并不能够确认转发IP报文的下一跳设备，还需要明确的下一跳设备地址。

路由表



目的网络/掩码	下一跳	出接口
11.0.0.0/8	2.2.2.2	GE0/0
13.0.0.0/8	3.3.3.2	GE0/1
14.0.0.0/8	1.1.1.2	GE0/2
.....		
1.1.1.0/30	1.1.1.1	GE0/2
1.1.1.1/32	127.0.0.1	GE0/2

- 路由器通过各种方式发现路由
- 路由器选择最优的路由条目放入路由表中
- 路由表指导设备对IP报文的转发
- 路由器通过对路由表的管理实现对路径信息的管理

HUAWEI

- 路由器依据路由表转发报文。
- 路由表由一条条详细的路由条目组成。
- 路由表由路由条目组成，但不代表路由表中保存了所有路由，路由表中只会保存“最优的”路由。
- 对路由表中的路由条目的管理实际上就是路由器维护、管理路由信息的具体实现。

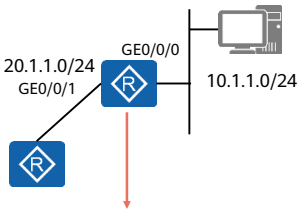
目录

路由信息获取方式

- 路由器依据路由表进行路由转发，为实现路由转发，路由器需要发现路由，以下为常见的路由获取方式。

直连路由

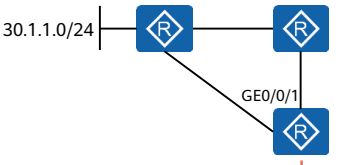
- 由设备自动生成指向本地直连网络



路由来源	目的网络/掩码	出接口
直连	10.1.1.0/24	GE0/0/0
直连	20.1.1.0/24	GE0/0/1

静态路由

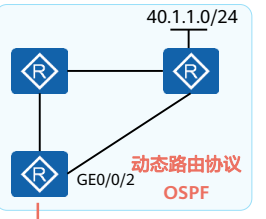
- 由网络管理员手工配置的路由条目



路由来源	目的网络/掩码	出接口
静态	30.1.1.0/24	GE0/0/1

动态路由

- 路由器运行动态路由协议学习到的路由

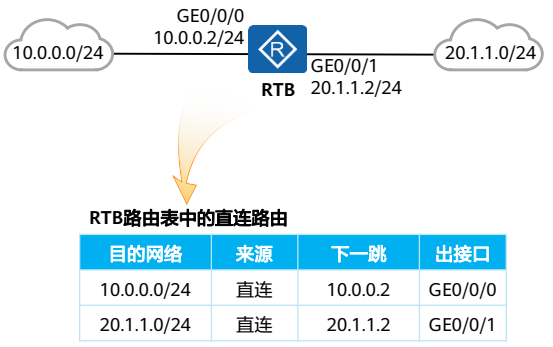


路由来源	目的网络/掩码	出接口
动态路由协议	40.1.1.0/24	GE0/0/2

- 直连路由：直连接口所在网段的路由，由设备自动生成。
- 静态路由：由网络管理员手工配置的路由条目
- 动态路由：路由器通过动态路由协议（如OSPF、IS-IS、BGP等）学习到的路由

直连路由（1）

直连路由



RTB路由表中的直连路由

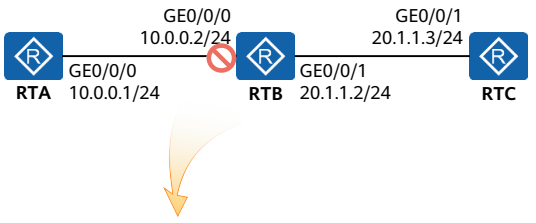
目的网络	来源	下一跳	出接口
10.0.0.0/24	直连	10.0.0.2	GE0/0/0
20.1.1.0/24	直连	20.1.1.2	GE0/0/1

- 直连路由指向本地直连网络的路由，由设备自动生成。
- 当路由器为路由转发的最后一跳路由器时，IP报文匹配直连路由，路由器转发IP报文到目的主机。
- 使用直连路由进行路由转发时，报文的目的IP和路由器接口IP在一个网段之中。

- 当匹配中直连路由进行转发时，此时路由器会查看ARP表项，将报文直接转到目的地址，此时该路由器为路由转发的最后一跳路由器。
- 直连路由的下一跳地址并不是其他设备上的接口地址，因为该路由的目的网段为接口所在网段，本接口就是最后一跳，不需要再转发给下一跳，所以在路由表中的下一跳地址就是接口自身地址。
- 使用直连路由进行路由转发时，转发的动作不是交给下一跳，而是查询ARP表项，根据ARP表项封装报文，将报文发送到目的IP。

直连路由 (2)

直连路由



RTB路由表中的直连路由

目的网络	来源	下一跳	出接口
20.1.1.0/24	直连	20.1.1.2	G0/0/1

- GE0/0/0接口down，该接口的直连路由未出现在路由表中

- 并不是所有接口生成的直连路由都会出现在路由表中，直连路由出现在路由表中的前提是该接口的物理状态、协议状态都为UP。

目录

查看IP路由表

```
<Huawei> display ip routing-table
Route Flags: R - relay, D - download to fib
------------------------------------------------------------------------------
Routing Tables: Public
         Destinations : 6        Routes : 6

Destination/Mask    Proto   Pre  Cost      Flags NextHop         Interface

        1.1.1.1/32  Static  60   0           D   0.0.0.0         NULL0
        2.2.2.2/32  Static  60   0           D   100.0.0.2       Vlanif100
      100.0.0.0/24  Direct  0    0           D   100.0.0.1       Vlanif100
      100.0.0.1/32  Direct  0    0           D   127.0.0.1       Vlanif100
      127.0.0.0/8   Direct  0    0           D   127.0.0.1       InLoopBack0
      127.0.0.1/32  Direct  0    0           D   127.0.0.1       InLoopBack0
```

目的网络地址/网络掩码长度　协议类型　路由优先级　开销（度量值）　标志　下一跳地址　出接口

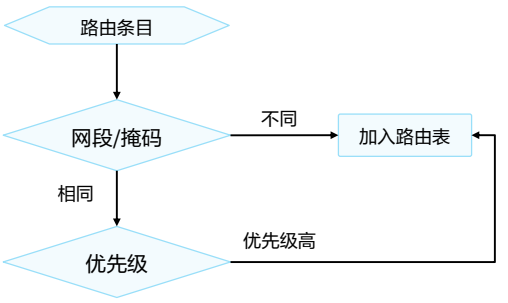
路由表中各个内容的含义

- Destination/Mask：表示此路由的目的网络地址与网络掩码。将目的地址和子网掩码“逻辑与”后可得到目的主机或路由器所在网段的地址。例如：目的地址为1.1.1.1，掩码为255.255.255.0的主机或路由器所在网段的地址为1.1.1.0。
- Proto（Protocol）：该路由的协议类型，也即路由器是通过什么协议获知该路由的。
- Pre（Preference）：表示此路由的路由协议优先级。针对同一目的地，可能存在不同下一跳、出接口等多条路由，这些不同的路由可能是由不同的路由协议发现的，也可以是手工配置的静态路由。优先级最高（数值最小）者将成为当前的最优路由。
- Cost：路由开销。当到达同一目的地的多条路由具有相同的路由优先级时，路由开销最小的将成为当前的最优路由。
- NextHop：表示对于本路由器而言，到达该路由指向的目的网络的下一跳地址。该字段指明了数据转发的下一个设备。
- Interface：表示此路由的出接口。指明数据将从本路由器的哪个接口转发出去。

- Preference用于不同路由协议间路由优先级的比较，Cost用于同一种路由协议内部不同路由的优先级的比较。在业界，Cost也被称为路由度量值（Metric）。

路由优先级 - 基本概念

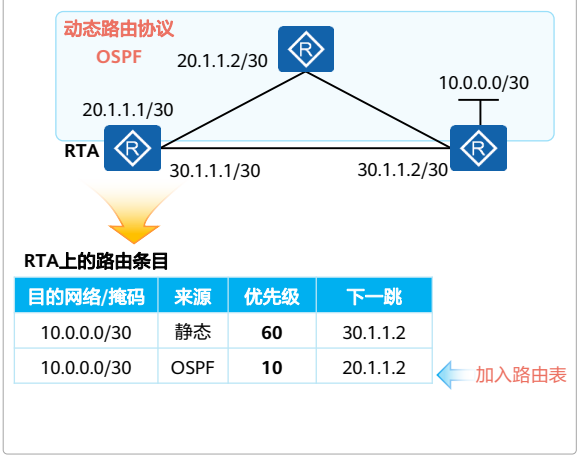
优先级比较



- 当路由器从多种不同的途径获知到达同一个目的网段的路由（这些路由的目的网络地址及网络掩码均相同）时，路由器会比较这些路由的优先级，优选优先级值最小的路由。
- 路由来源的优先级值（Preference）越小代表加入路由表的优先级越高。
- 拥有最高优先级的路由将被添加进路由表。

路由优先级 - 比较过程

优先级比较示例



RTA上的路由条目

目的网络/掩码	来源	优先级	下一跳
10.0.0.0/30	静态	**60**	30.1.1.2
10.0.0.0/30	OSPF	**10**	20.1.1.2

加入路由表

- RTA通过动态路由协议OSPF和手动配置的方式都发现了到达10.0.0.0/30的路由，此时会比较这两条路由的优先级，优选优先级值最小的路由。
- 每一种路由协议都有相应的优先级。
- OSPF拥有更优的优先级，因此通过OSPF学习到的路由被添加到路由表中。

- RTA通过静态、动态路由协议学习到相同的路由条目，比较路由协议优先级，OSPF优先。OSPF的路由条目被加入到路由表。

路由优先级 - 常见默认数值

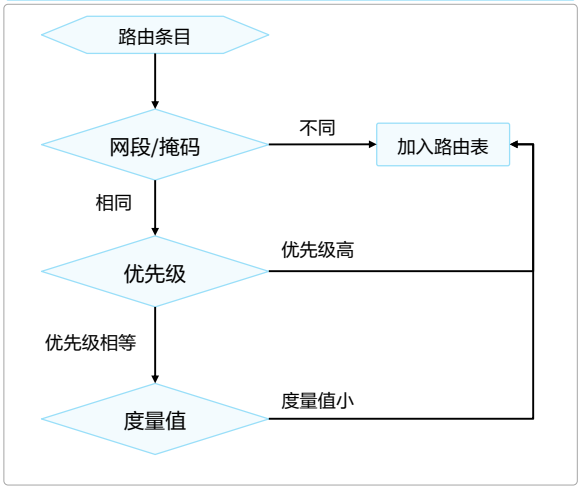
- 常见路由类型的默认优先级如下：

路由来源	路由类型	默认优先级
直连	直连路由	0
静态	静态路由	60
动态路由	OSPF内部路由	10
	OSPF外部路由	150

- 这里例举了一些常见的路由协议优先级，实际上动态路由的类型存在多种，我们将会在后续的学习中详细地了解它们，上表中只展示了OSPF的路由优先级。

度量值 - 基本概念

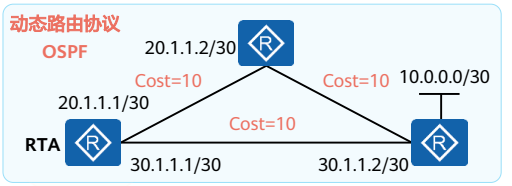
度量值比较



- 当路由器通过某种路由协议发现了多条到达同一个目的网络的路由时（拥有相同的路由优先级），度量值将作为路由优选的依据之一。
- 路由度量值表示到达这条路由所指目的地址的代价。
- 一些常用的度量值有：跳数、带宽、时延、代价、负载、可靠性等。
- 度量值数值越小越优先，度量值最小路由将会被添加到路由表中。
- 度量值很多时候被称为开销（Cost）。

度量值 - 比较过程

度量值比较示例



RTA上的路由条目

目的网络/掩码	来源	Cost	下一跳
10.0.0.0/30	OSPF	**20**	20.1.1.2
10.0.0.0/30	OSPF	**10**	30.1.1.2

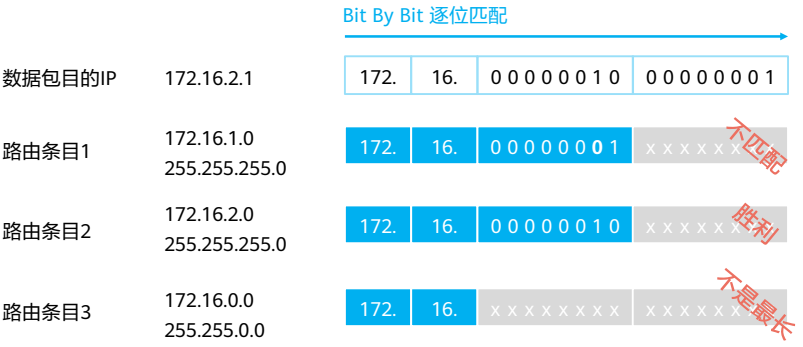
加入路由表

- RTA通过动态路由协议OSPF学习到了两条目的地为10.0.0.0/30的路由，学习自同一路由协议、优先级相同，因此需要继续比较度量值。
- 两条路由拥有不同的度量值，下一跳为30.1.1.2的OSPF的路由条目拥有更小的度量值，因此被加入到路由表中。

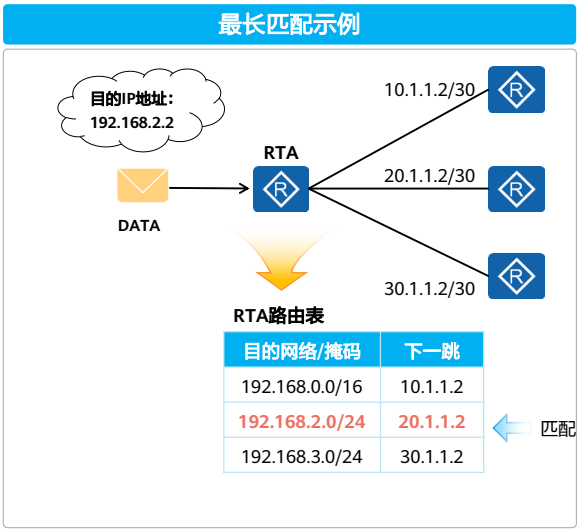
目录

最长匹配原则

- 当路由器收到一个IP数据包时，会将数据包的目的IP地址与自己本地路由表中的所有路由表项进行逐位（Bit-By-Bit）比对，直到找到匹配度最长的条目，这就是最长前缀匹配机制。

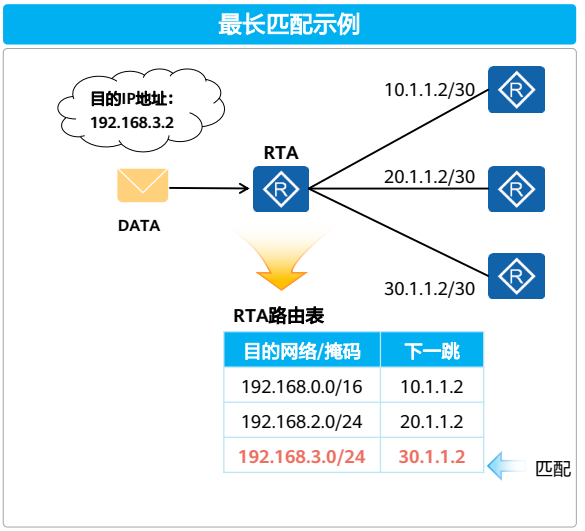


最长匹配示例 (1)



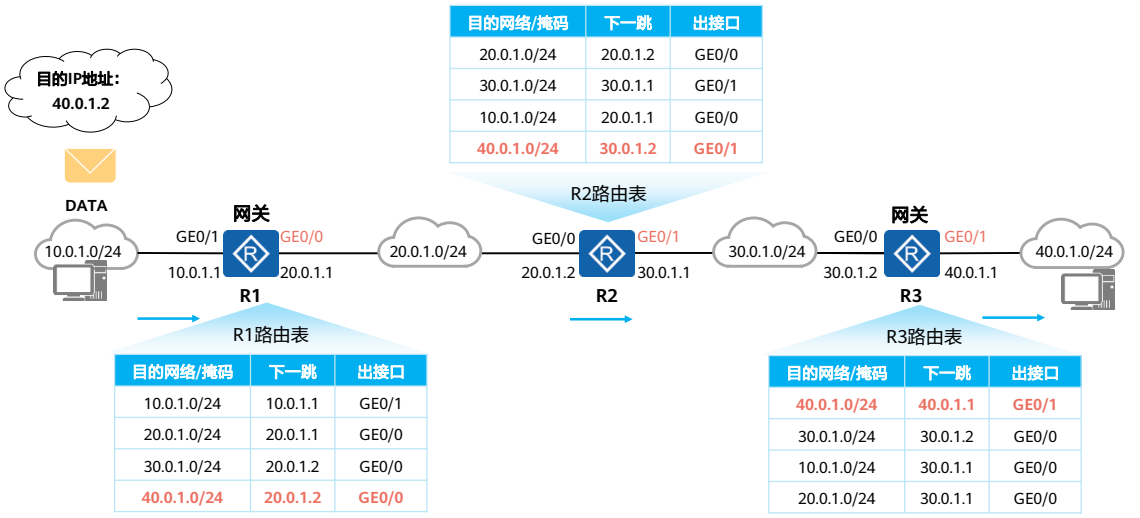
- 根据最长匹配原则进行匹配，能够匹配192.168.2.2的路由存在两条，但是路由的掩码长度中，一个为16 bit，另一个为24 bit，掩码长度为24 bit的路由满足最长匹配原则，因此被选择来指导发往192.168.2.2的报文转发。

最长匹配示例 (2)



- 根据最长匹配原则匹配，能够匹配到192.168.3.2的路由只有一条，此路由为最终转发依据。

路由转发流程



- 来自10.0.1.0/24网段的IP报文想要去往40.0.1.0/24网段，首先到达网关，网关查找路由表项，确定转发的下一跳、出接口，之后报文转发给R2。报文到达R2之后，R2通过查找路由表项转发给R3，R3收到后查找路由表项，发现IP报文目的IP属于本地接口所在网段，直接本地转发。

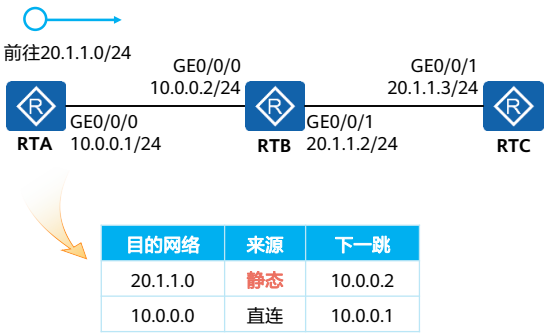
IP路由表小结

- 当路由器从多种不同的途径获知到达同一个目的网段的路由（这些路由的目的网络地址及网络掩码均相同）时，会选择路由优先级值最小的路由；如果这些路由学习自相同的路由协议，则优选度量值最优的。总之，最优的路由加入路由表。
- 当路由器收到一个数据包时，会在自己的路由表中查询数据包的目的IP地址。如果能够找到匹配的路由表项，则依据表项所指示的出接口及下一跳来转发数据；如果没有匹配的表项，则丢弃该数据包。
- 路由器的行为是逐跳的，数据包从源到目的地沿路径每个路由器都必须有关于目标网段的路由，否则就会造成丢包。
- 数据通信往往是双向的，因此要关注流量的往返（往返路由）。

目录

静态路由应用场景

静态路由



目的网络	来源	下一跳
20.1.1.0	静态	10.0.0.2
10.0.0.0	直连	10.0.0.1

- 静态路由由网络管理员手动配置，配置方便，对系统要求低，适用于拓扑结构简单并且稳定的小型网络。
- 缺点是不能自动适应网络拓扑的变化，需要人工干预。
- RTA上转发目的地址属于20.1.1.0/24的报文，在只有直连路由的情况下没有路由匹配。此时可以通过手动配置静态路由，使RTA发送前往20.1.1.0/24网段的报文交给下一跳10.0.0.2转发。

静态路由配置

1. 关联下一跳IP的方式

   [Huawei] **ip route-static** *ip-address* { *mask* | *mask-length* } *nexthop-address*

2. 关联出接口的方式

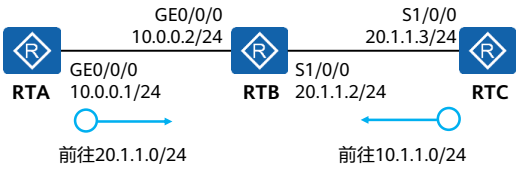
   [Huawei] **ip route-static** *ip-address* { *mask* | *mask-length* } *interface-type interface-number*

3. 关联出接口和下一跳IP方式

   [Huawei] **ip route-static** *ip-address { mask | mask-length } interface-type interface-number [nexthop-address]*

   在创建静态路由时，可以同时指定出接口和下一跳。对于不同的出接口类型，也可以只指定出接口或只指定下一跳。
   对于点到点接口（如串口），只需指定出接口。
   对于广播接口（如以太网接口）和VT（Virtual-template）接口，必须指定下一跳。

配置举例



- RTA与RTC上配置静态路由，实现10.0.0.0/24与20.1.1.0/24的互通。
- 因为报文是逐跳转发的，所以每一跳路由设备上都需要配置到达相应目的地址的路由。
- 另外需要注意通信是双向的，针对通信过程中的往返流量，都需关注途径设备上的路由配置。

RTA的配置如下：

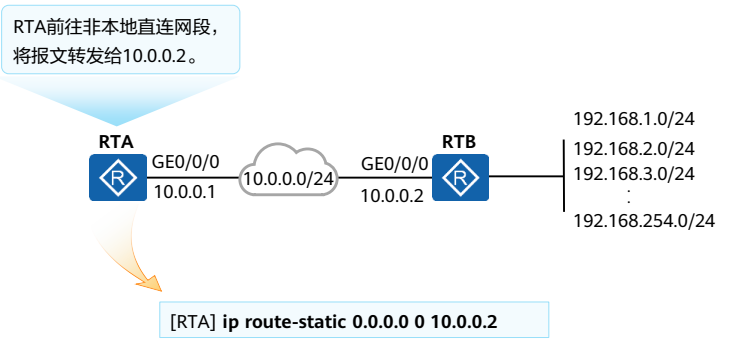
```
[RTA] ip route-static 20.1.1.0 255.255.255.0 10.0.0.2
```

RTC的配置如下：

```
[RTC] ip route-static 10.0.0.0 255.255.255.0 S1/0/0
```

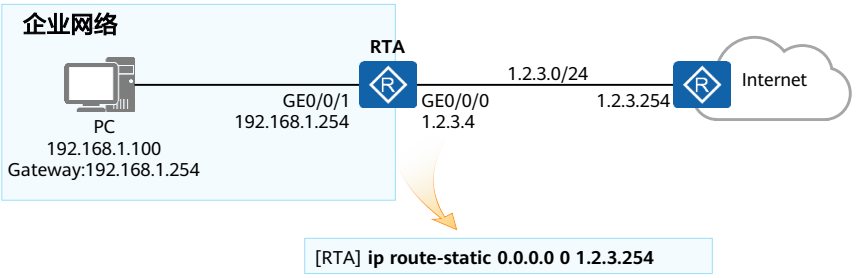
缺省路由

- 缺省路由是一种特殊的路由，当报文没有在路由表中找到匹配的具体路由表项时才使用的路由。如果报文的目的地址不能与路由表的任何目的地址相匹配，那么该报文将选取缺省路由进行转发。
- 缺省路由在路由表中的形式为0.0.0.0/0，缺省路由也被叫做默认路由。



缺省路由应用场景

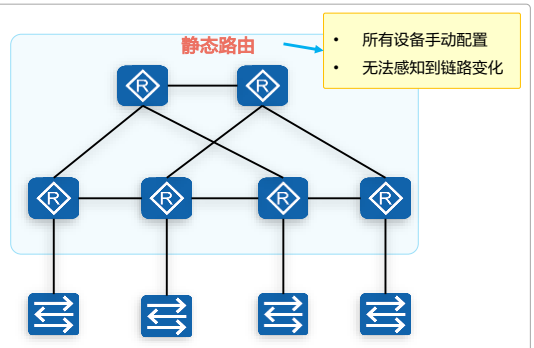
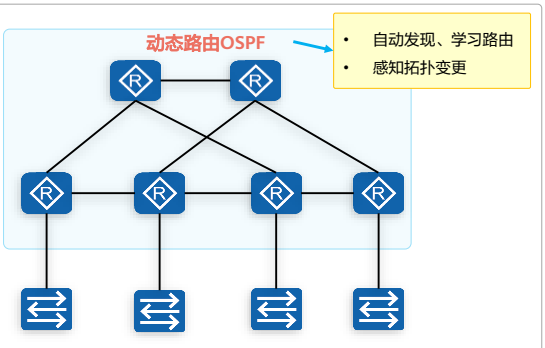
- 缺省路由一般用于企业网络出口，配置一条缺省路由让出口设备能够转发前往Internet上任意地址的IP报文。



HUAWEI

目录

动态路由概述

静态路由	动态路由
 • 当网络规模越来越大时，使用手动配置静态路由的方式获取路由条目将变得越发复杂，同时在拓扑发生变化时不能及时、灵活响应。	 • 动态路由协议能够自动发现和生成路由，并在拓扑变化时及时更新路由，可以有效减少管理人员工作量，更适用于大规模网络。

- 静态路由的缺点是不能自动适应网络拓扑的变化，需要人工干预。
- 动态路由协议有自己的路由算法，能够自动适应网络拓扑的变化，适用于具有一定数量三层设备的网络。

动态路由分类

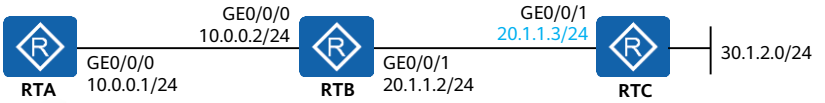
HUAWEI

- 根据路由信息传递的内容、计算路由的算法，可以将动态路由协议分为两大类
  - 距离矢量协议（Distance-Vector Protocol）
    - RIP
  - 链路状态协议（Link-State Protocol）
    - OSPF
    - IS-IS
  - BGP使用一种基于距离矢量算法修改后的算法，该算法被称为路径矢量（Path Vector）算法。因此在某些场合下，BGP也被称为路径矢量路由协议。
- 根据工作范围不同，又可以分为
  - 内部网关协议IGP（Interior Gateway Protocol）:在一个自治系统内部运行。RIP、OSPF、ISIS为常见的IGP协议。
  - 外部网关协议EGP（Exterior Gateway Protocol）：运行于不同自治系统之间。BGP是目前最常用的EGP协议。

目录

路由递归 (1)

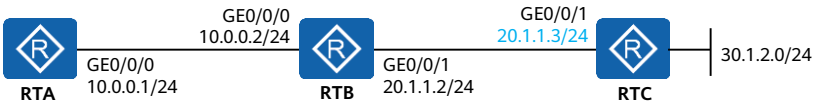
- 路由必须有直连的下一跳才能够指导转发，但是路由生成时下一跳可能不是直连的，因此需要计算出一个直连的下一跳和对应的出接口，这个过程就叫做路由递归。
- 路由递归也被称为路由迭代。



[RTA] **ip route-static 30.1.2.0 24 20.1.1.3**

去往30.1.2.0/24的路由，下一跳为20.1.1.3，非本地直连网络，如果路由表中没有去往20.1.1.3的路由，该静态路由将不会生效，无法作为有效路由条目，并不会出现在路由表。

路由递归 (2)



[RTA] **ip route-static 30.1.2.0 24 20.1.1.3**

递归

[RTA] **ip route-static 20.1.1.0 24 10.0.0.2**

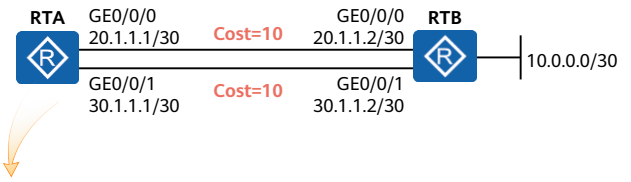
目的网络	下一跳	出接口
30.1.2.0/24	20.1.1.3	GE0/0/0
20.1.1.0/24	10.0.0.2	GE0/0/0

添加一条去往20.1.1.3的路由，下一跳为直连网络内的IP地址10.0.0.2。
去往30.1.2.0/24的路由通过递归查询得到一个直连的下一跳，该路由因此生效。

HUAWEI

等价路由 (1)

等价路由



RTA路由表

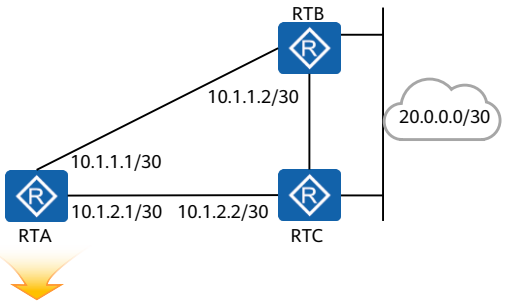
目的网络/掩码	下一跳
10.0.0.0/30	20.1.1.2
	30.1.1.2

来源相同、开销相同的路由都会被加入路由表，形成的路由为等价路由（两个路由条目指向的目的网段相同，但是具有不同的下一跳地址），路由转发会将流量分布到多条路径上。

- 路由表中存在等价路由之后，前往该目的网段的IP报文路由器会通过所有有效的接口、下一跳转发，这种转发行为被称为负载分担。

浮动路由 - 基本概念

浮动路由

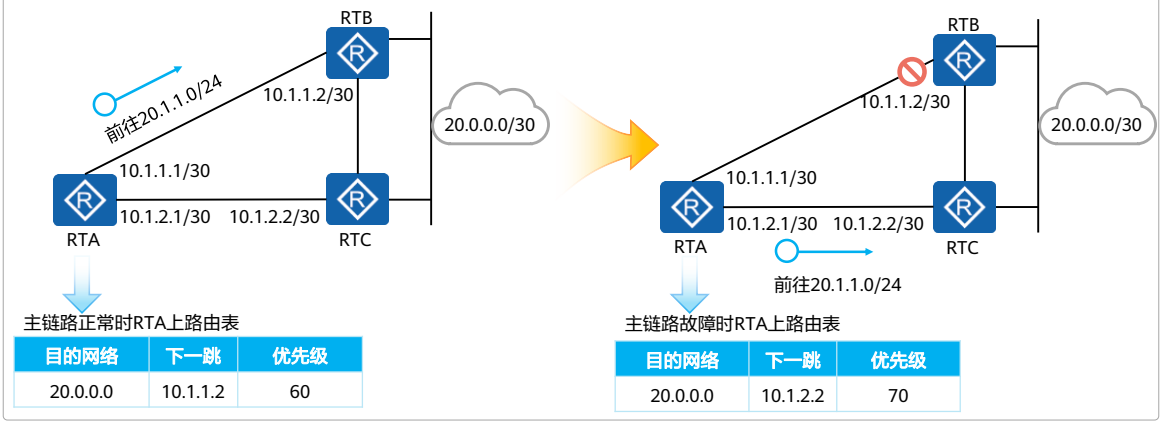


RTA上配置浮动路由

```
[RTA] ip route-static 20.0.0.0 30 10.1.1.2
[RTA] ip route-static 20.0.0.0 30 10.1.2.2  preference 70
```

- 静态路由支持配置时手动指定优先级，可以通过配置目的地址/掩码相同、优先级不同、下一跳不同的静态路由，实现转发路径的备份。
- 浮动路由是主用路由的备份，保证链路故障时提供备份路由。主用路由下一跳可达时该备份路由不会出现在路由表。

浮动路由 - 示例

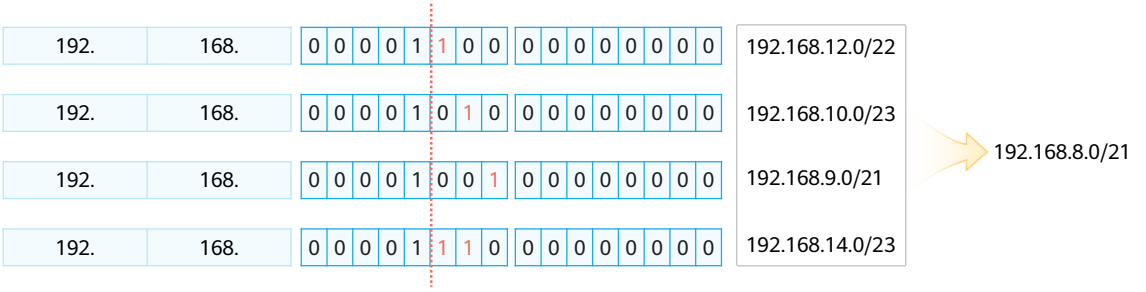


  HUAWEI

- RTA-RTB之间的链路正常时，20.0.0.0/30的两条路由条目都是有效的条目，此时比较优先级，下一跳为10.1.1.2的优先级60，下一跳为10.1.2.2的优先级70，因此下一跳为10.1.1.2的加入路由表。
- RTA-RTB之间的链路故障时，10.1.1.2不可达，因此下一跳为10.1.1.2的路由失效，此时前往20.0.0.0/30的路由就只存在一条，该条路由将会被选入路由表。前往20.0.0.1的流量将会被转发到10.1.2.2。

CIDR

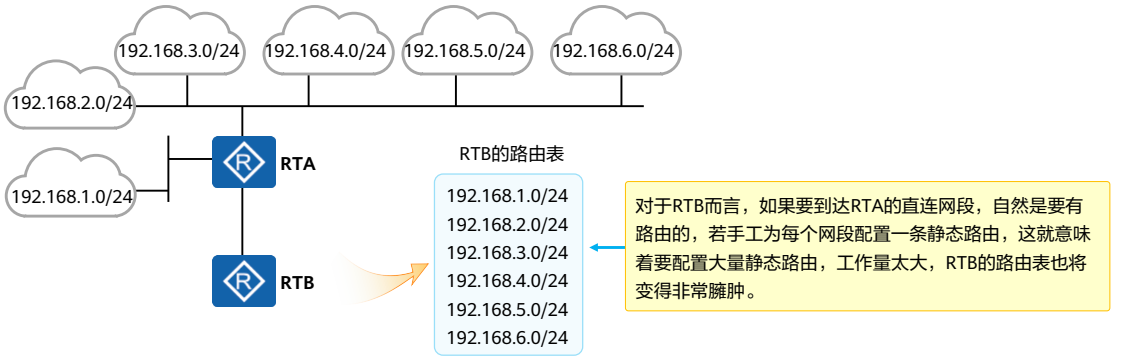
- CIDR（classless inter-domain routing，无类别域间路由）采用IP地址加掩码长度来标识网络和子网，而不是按照传统A、B、C等类型对网络地址进行划分。
- CIDR容许任意长度的掩码长度，将IP地址看成连续的地址空间，可以使用任意长度的前缀分配，多个连续的前缀可以聚合成一个网络，该特性可以有效减少路由表条目数量。



HUAWEI

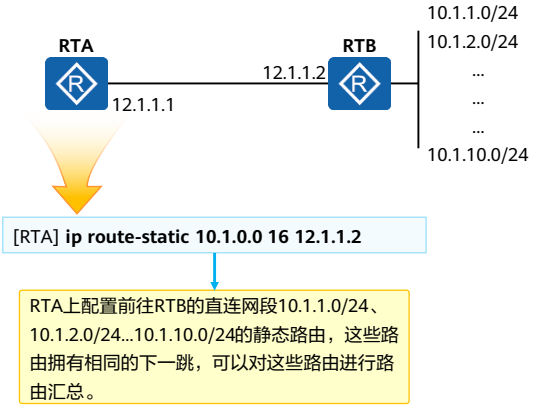
路由汇总需求

- 子网划分、VLSM解决了地址空间浪费的问题，但同时也带了新的问题：路由表中的路由条目数量增加。
- 为减少路由条目数量可以使用路由汇总。



- 对于一个大规模的网络来说，路由器或其他具备路由功能的设备势必需要维护大量的路由表项，为了维护臃肿的路由表，这些设备就不得不耗费大量的资源。同时，由于路由表的规模变大，会导致路由器在查表转发时效率降低。因此在保证网络中的路由器到各网段都具备IP可达性的同时，需要减小设备的路由表规模。一个网络如果具备科学的IP编址，并且进行合理的规划，是可以利用多种手段减小设备路由表规模的。一个非常常见而又有效的办法就是使用路由汇总（Route Summarization）。路由汇总又被称为路由聚合（Route Aggregation），是将一组有规律的路由汇聚成一条路由，从而达到减小路由表规模以及优化设备资源利用率的目的，我们把汇聚之前的这组路由称为精细路由或明细路由，把汇聚之后的这条路由称为汇总路由或聚合路由。

路由汇总简介

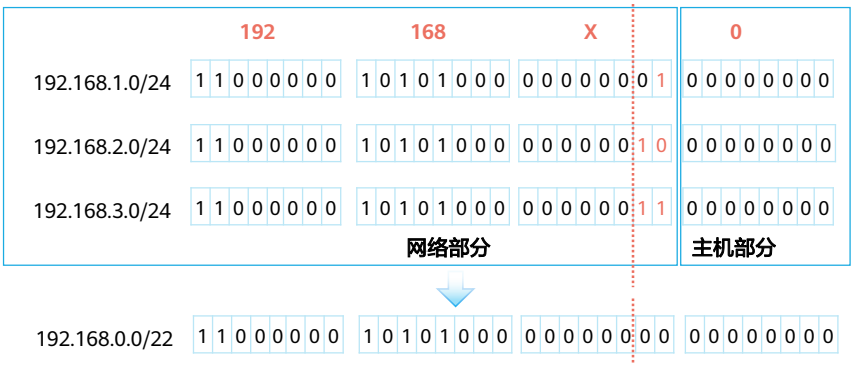


- 路由汇总将一组具有相同前缀的路由汇聚成一条路由，从而达到减小路由表规模以及优化设备资源利用率的目的。
- 路由汇总采用了CIDR的思想：将相同前缀的地址聚合成一个。
- 我们把汇聚之前的这组路由称为精细路由或明细路由，把汇聚之后的这条路由称为汇总路由或聚合路由。

  HUAWEI

- RTA上为了能够前往远端地址，需要为每一个远端网段配置一条明细路由。去往10.1.1.0/24、10.1.2.0/24、10.1.3.0/24...拥有相同下一跳。将拥有相同下一跳，一组有规律的路由汇总成一条路由，这叫做路由汇总。
- 路由汇总可以有效减少路由表项大小。

汇总计算

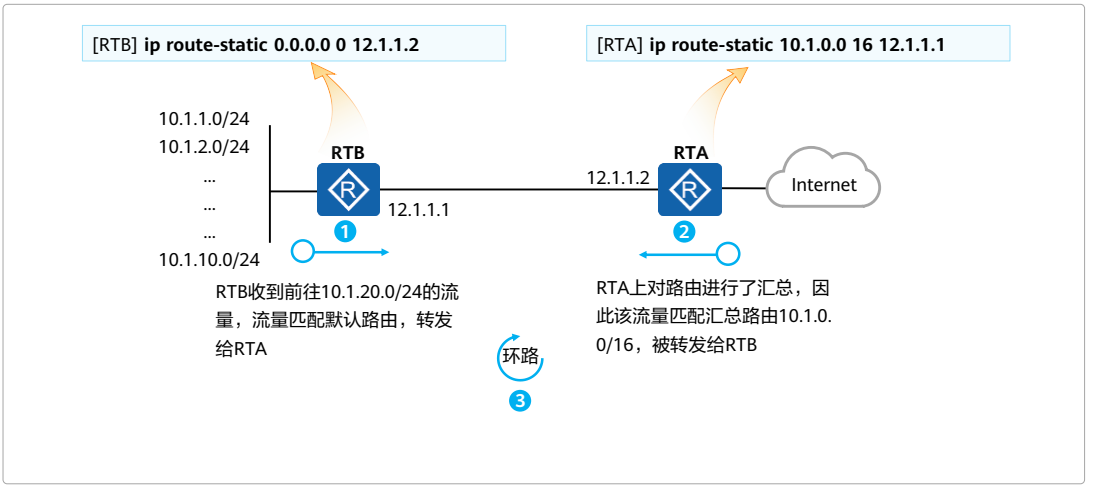


- 基于一系列连续的、有规律的IP网段，如果需计算相应的汇总路由，且确保得出的汇总路由刚好“囊括”上述IP网段，则需保证汇总路由的掩码长度尽可能长。
- 诀窍在于：将明细路由的目的网络地址都换算成二进制，然后排列起来，找出所有目的网络地址中“相同的比特位”。

HUAWEI

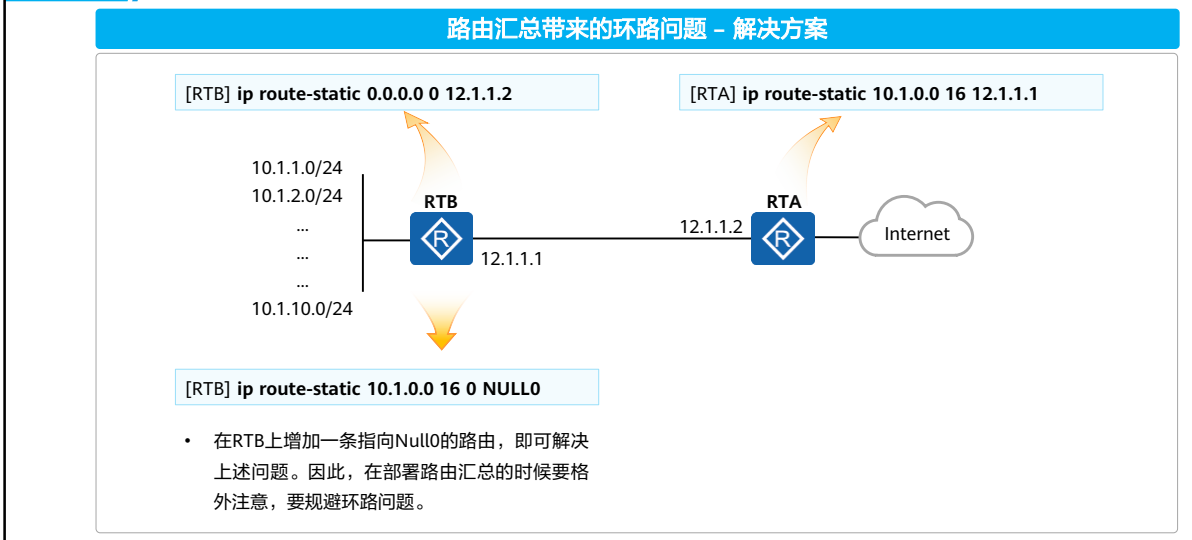
汇总引发的问题 (1)

路由汇总带来的环路问题



HUAWEI

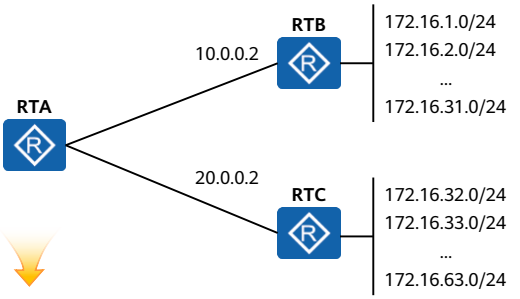
汇总引发的问题 (2)



- 一般来说一条路由，无论是静态的或者是动态的，都需要关联到一个出接口，路由的出接口指的是设备要到达一个目的网络时的出站接口。路由的出接口可以是该设备的物理接口，例如百兆、千兆以太网接口，也可以是逻辑接口，例如VLAN接口（VLAN Interface），或者隧道（Tunnel）接口等。在众多类型的出接口中，有一种接口非常特殊，那就是Null（无效）接口，这种类型的接口只有一个编号，也就是0。Null0是一个系统保留的逻辑接口，当网络设备在转发某些数据包时，如果使用出接口为Null0的路由，那么这些报文将被直接丢弃，就像被扔进了一个黑洞里，因此出接口为Null0的路由又被称为黑洞路由。

精确汇总（1）

精确进行路由汇总

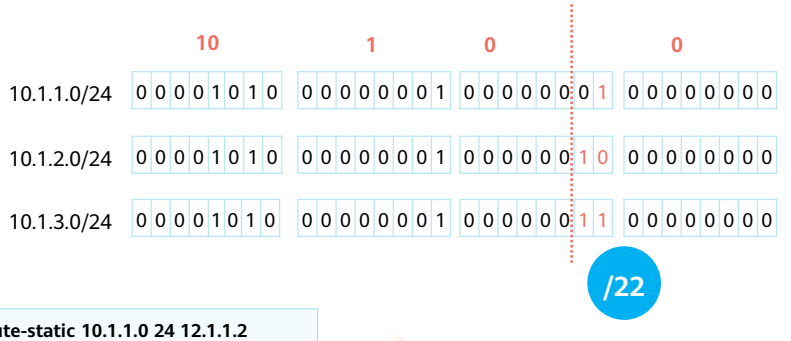


[RTA] **ip route-static 172.16.0 16 10.0.0.2**

- 为了让RTA能够到达RTB上的172.16.1.0/24-172.16.31.0/24网段，配置了一条静态的汇总路由，这条网段虽然优化了网络配置，但是汇总的范围太广，将RTC上的网段也包括在内，导致前往RTC上网段的流量到达RTA之后会被发往RTB，造成数据包的丢失，这种路由为不精确的路由。为此配置汇总路由时要尽量精确，刚好包括所有明细路由。

精确汇总 (2)

精确进行路由汇总




```
ip route-static 10.1.1.0 24 12.1.1.2
ip route-static 10.1.2.0 24 12.1.1.2
ip route-static 10.1.3.0 24 12.1.1.2
```

```
ip route-static 10.1.1.0 22 12.1.1.2
```

精确计算汇总后的网络号、掩码，避免汇总后掩码过小。

思考题

1. 路由器如何优选路由条目?
2. 如何配置实现浮动路由?
3. 将10.1.1.0/24、10.1.3.0/24、10.1.9.0/24汇总之后的网段是?

1. 首先根据preference选择，如果preference相同则继续比较度量值，如果度量值也相同，则都会被加入路由表形成等价路由。
2. 配置一条和被备份的路由目的网段、掩码相同的静态路由，但下一跳不同，最后加上一个preference xx，xx值大于被备份路由的preference值即可实现备份路由。
3. 10.1.0.0/20

本章总结

- 本章节学习了路由的基本概念，了解了路由如何指导路由器对IP报文进行转发，同时还了解了常见的路由属性。
- 特殊的静态路由：缺省路由，此外本章节展现了一些路由转发的高级特性，包括路由递归、浮动路由、等价路由，这些都在现网中有着广泛地应用。


谢 谢
www.huawei.com

OSPF基础

前言

- 由于静态路由由网络管理员手工配置，因此当网络发生变化时，静态路由需要手动调整，这制约了静态路由在现网大规模的应用。
- 动态路由协议因其灵活性高、可靠性好、易于扩展等特点被广泛应用于现网。在动态路由协议之中，OSPF（Open Shortest Path First，开放式最短路径优先）协议是使用场景非常广泛的动态路由协议之一。
- OSPF在RFC2328中定义，是一种基于链路状态算法的路由协议。
- 本课程将初步介绍OSPF基本概念、工作原理和基础配置。

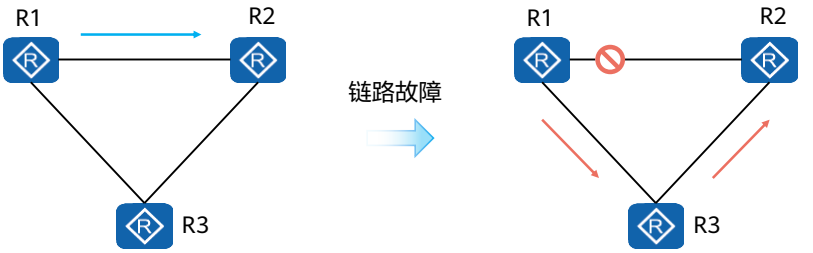
目标

- 学完本课程后，您将能够：
  - 描述动态路由协议的优势和它的分类
  - 描述OSPF的基本概念和适用的组网场景
  - 阐明OSPF协议的工作原理
  - 了解OSPF协议的基础配置

目录

为什么需要动态路由协议?

- 静态路由是由工程师手动配置和维护的路由条目，命令行简单明确，适用于小型或稳定的网络。静态路由有以下问题：
  - 无法适应规模较大的网络：随着设备数量增加，配置量急剧增加。
  - 无法动态响应网络变化：网络发生变化，无法自动收敛网络，需要工程师手动修改。



R1-R2静态路由　　手动配置R1-R3-R2静态路由

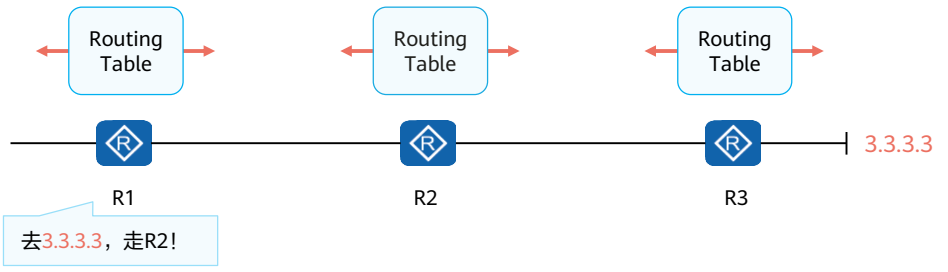
动态路由协议的分类

HUAWEI

- BGP使用一种基于距离矢量算法修改后的算法，该算法被称为路径矢量（Path Vector）算法。因此在某些场合下，BGP也被称为路径矢量路由协议。

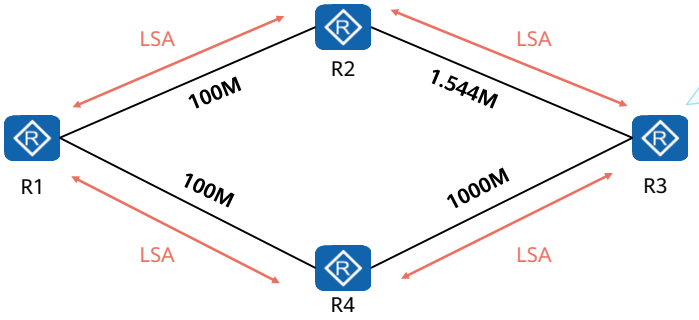
距离矢量路由协议

- 运行距离矢量路由协议的路由器周期性的泛洪自己的路由表。通过路由的交互，每台路由器都从相邻的路由器学习到路由，并且加载进自己的路由表中。
- 对于网络中的所有路由器而言，路由器并不清楚网络的拓扑，只是简单的知道要去往某个目的方向在哪里，距离有多远。这即是距离矢量算法的本质。



链路状态路由协议 - LSA泛洪

- 与距离矢量路由协议不同，链路状态路由协议通告的的是链路状态而不是路由表。运行链路状态路由协议的路由器之间首先会建立一个协议的邻居关系，然后彼此之间开始交互LSA（Link State Advertisement，链路状态通告）。

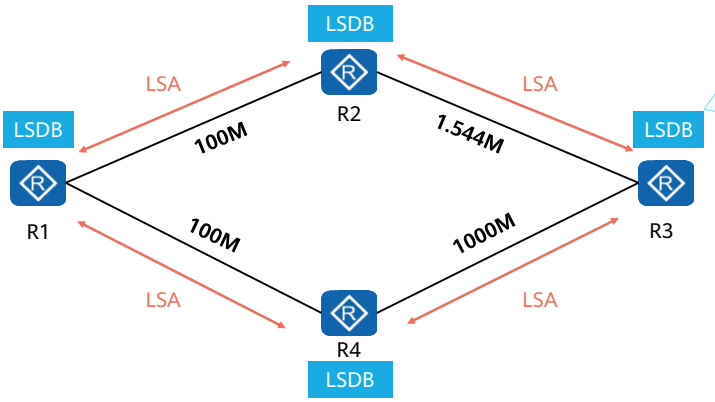


- 不再通告路由信息，而是LSA。
- LSA描述了路由器接口的状态信息，例如接口的开销、连接的对象等。

- 链路状态通告，可以简单的理解为每台路由器都产生一个描述自己直连接口状态（包括接口的开销、与邻居路由器之间的关系等）的通告。

链路状态路由协议 - LSDB组建

- 每台路由器都会产生LSAs，路由器将接收到的LSAs放入自己的LSDB（Link State DataBase，链路状态数据库）。路由器通过LSDB，掌握了全网的拓扑。

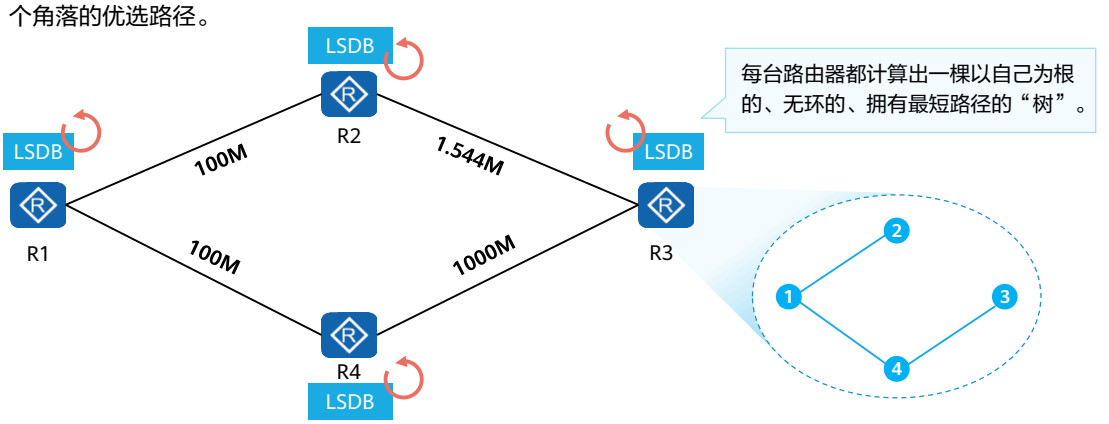


- 路由器将LSA存放在LSDB中
- LSDB汇总了网络中路由器对于自己接口的描述
- LSDB包含全网拓扑的描述

HUAWEI

链路状态路由协议 - SPF计算

- 每台路由器基于LSDB，使用SPF（Shortest Path First，最短路径优先）算法进行计算。每台路由器都计算出一棵以自己为根的、无环的、拥有最短路径的“树”。有了这棵“树”，路由器就已经知道了到达网络各个角落的优选路径。

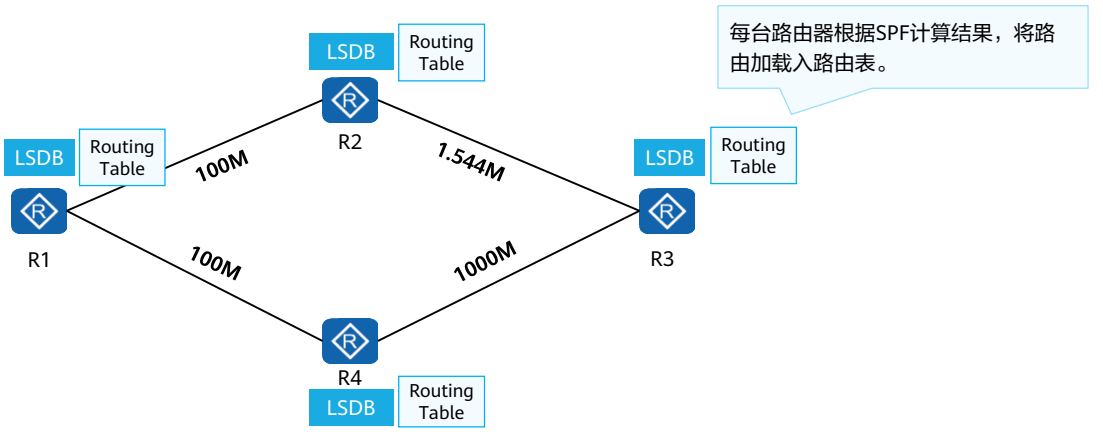


HUAWEI

- SPF是OSPF路由协议的一个核心算法，用来在一个复杂的网络中做出路由优选的决策。

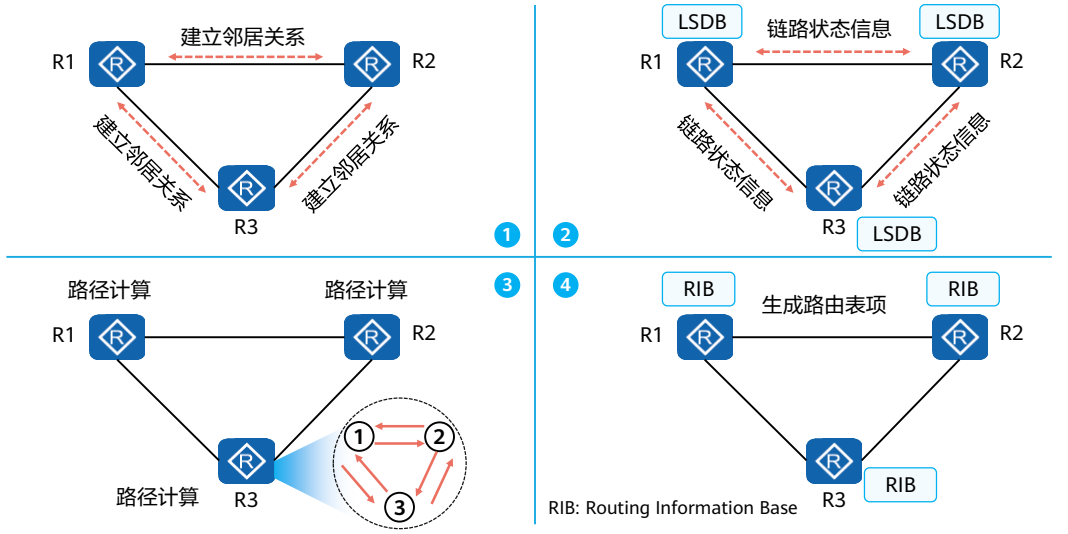
链路状态路由协议 - 路由表生成

- 最后，路由器将计算出来的优选路径，加载进自己的路由表（Routing Table）。



HUAWEI

链路状态路由协议总结

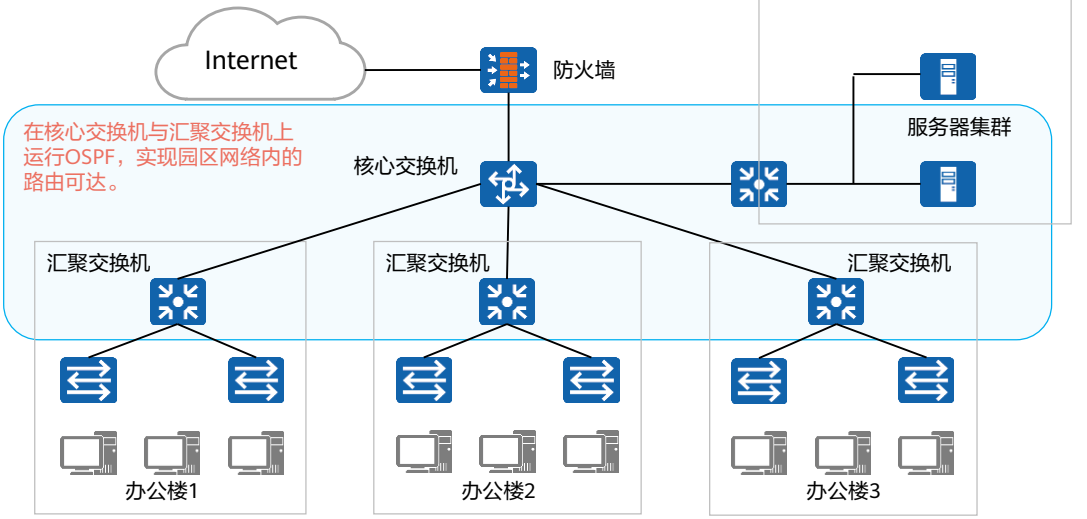


- 链路状态路由协议有四个步骤：
  - 第一步是建立相邻路由器之间的邻居关系。
  - 第二步是邻居之间交互链路状态信息和同步LSDB。
  - 第三步是进行优选路径计算。
  - 第四步是根据最短路径树生成路由表项加载到路由表。

OSPF简介

- OSPF是典型的链路状态路由协议，是目前业内使用非常广泛的IGP协议之一。
- 目前针对IPv4协议使用的是OSPF Version 2（RFC2328）；针对IPv6协议使用OSPF Version 3（RFC2740）。如无特殊说明本章后续所指的OSPF均为OSPF Version 2。
- 运行OSPF路由器之间交互的是LS（Link State，链路状态）信息，而不是直接交互路由。LS信息是OSPF能够正常进行拓扑及路由计算的关键信息。
- OSPF路由器将网络中的LS信息收集起来，存储在LSDB中。路由器都清楚区域内的网络拓扑结构，这有助于路由器计算无环路径。
- 每台OSPF路由器都采用SPF算法计算达到目的地的最短路径。路由器依据这些路径形成路由加载到路由表中。
- OSPF支持VLSM（Variable Length Subnet Mask，可变长子网掩码），支持手工路由汇总。
- 多区域的设计使得OSPF能够支持更大规模的网络。

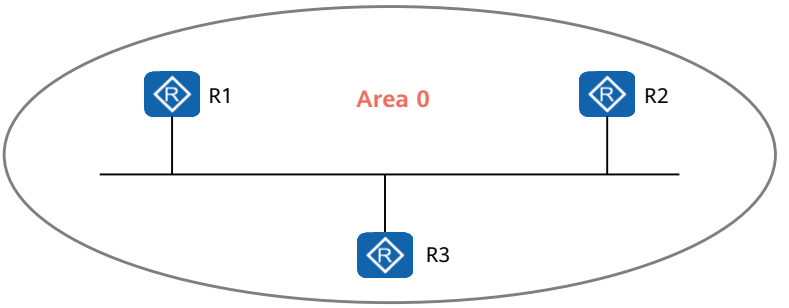
OSPF在园区网络中的应用



HUAWEI

OSPF基础术语：区域

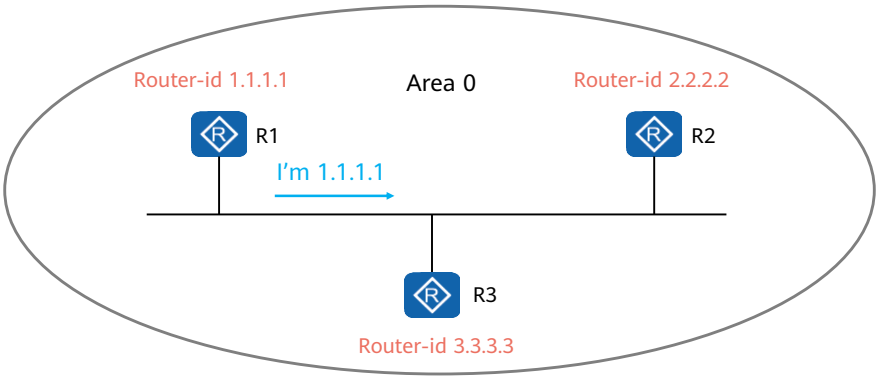
- OSPF Area用于标识一个OSPF的区域。
- 区域是从逻辑上将设备划分为不同的组，每个组用区域号（Area ID）来标识。



HUAWEI

OSPF基础术语：Router-ID

- Router-ID（Router Identifier，路由器标识符），用于在一个OSPF域中唯一地标识一台路由器。
- Router-ID的设定可以通过手工配置的方式，或使用系统自动配置的方式。

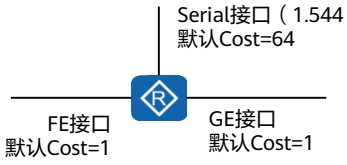


- 在实际项目中，通常会通过手工配置方式为设备指定OSPF Router-ID。请注意必须保证在OSPF域中任意两台设备的Router-ID都不相同。通常的做法是将Router-ID配置为与该设备某个接口（通常为Loopback接口）的IP地址一致。

OSPF的基础术语：度量值

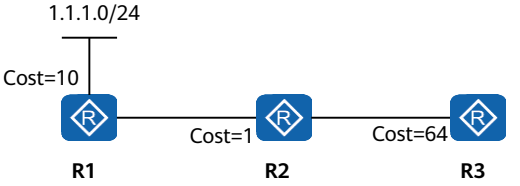
- OSPF使用Cost（开销）作为路由的度量值。每一个激活了OSPF的接口都会维护一个接口Cost值，缺省时接口Cost值 = $\frac{100\ \text{Mbit/s}}{\text{接口带宽}}$。其中100 Mbit/s为OSPF指定的缺省参考值，该值是可配置的。
- 笼统地说，一条OSPF路由的Cost值可以理解为是从目的网段到本路由器沿途所有入接口的Cost值累加。

OSPF接口Cost值



- OSPF不同接口因其带宽不同，有不同的Cost。

OSPF路径累计Cost值



- 在R3的路由表中，到达1.1.1.0/24的OSPF路由的Cost值=10+1+64，即75。

OSPF协议报文类型

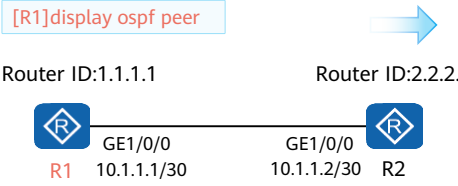
- OSPF有五种类型的协议报文。这些报文在OSPF路由器之间交互中起不同的作用。

报文名称	报文功能
Hello	周期性发送，用来发现和维护OSPF邻居关系。
Database Description	描述本地LSDB的摘要信息，用于两台设备进行数据库同步。
Link State Request	用于向对方请求所需要的LSA。设备只有在OSPF邻居双方成功交换DD报文后才会向对方发出LSR报文。
Link State Update	用于向对方发送其所需要的LSA。
Link State ACK	用来对收到的LSA进行确认。

OSPF三大表项 - 邻居表

- OSPF有三张重要的表项，OSPF邻居表、LSDB表和OSPF路由表。对于OSPF的邻居表，需要了解：
  - OSPF在传递链路状态信息之前，需先建立OSPF邻居关系。
  - OSPF的邻居关系通过交互Hello报文建立。
  - OSPF邻居表显示了OSPF路由器之间的邻居状态，使用display ospf peer查看。

[R1]display ospf peer



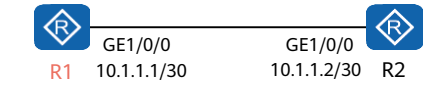

```
<R1> display ospf peer
    OSPF Process 1 with Router ID 1.1.1.1
                Neighbors
Area 0.0.0.0 interface 10.1.1.1(GigabitEthernet1/0/0)'s neighbors
Router ID: 2.2.2.2    Address: 10.1.1.2    GR State: Normal
    State: Full  Mode:Nbr is  Master  Priority: 1
    DR: 10.1.1.1  BDR: 10.1.1.2   MTU: 0
    Dead timer due in 35  sec
    Retrans timer interval: 5
    Neighbor is up for 00:00:05
    Authentication Sequence: [ 0 ]
```


- OSPF邻居表有很多关键信息，例如可以查看对端设备的Router ID和接口地址。更多详细信息在第二小结”OSPF协议工作原理”展开。

OSPF三大表项 - LSDB表

- 对于OSPF的LSDB表，需要了解：
  - LSDB会保存自己产生的及从邻居收到的LSA信息，本例中R1的LSDB包含了三条LSA。
  - Type标识LSA的类型，AdvRouter标识发送LSA的路由器。
  - 使用命令行display ospf lsdb查看LSDB表。

[R1]display ospf lsdb



<R1> display ospf lsdb

OSPF Process 1 with Router ID 1.1.1.1

Link State Database

Area: 0.0.0.0

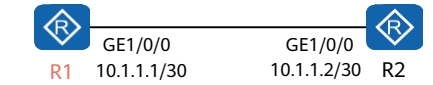
Type	LinkState ID	AdvRouter	Age	Len	Sequence	Metric
Router	2.2.2.2	2.2.2.2	98	36	8000000B	1
Router	1.1.1.1	1.1.1.1	92	36	80000005	1
Network	10.1.1.2	2.2.2.2	98	32	80000004	0

- 更多LSA相关内容请学习HCIP-DataCom。

OSPF三大表项 - OSPF路由表

- 对于OSPF的路由表，需要了解：
  - OSPF路由表和路由器路由表是两张不同的表项。本例中OSPF路由表有三条路由。
  - OSPF路由表包含Destination、Cost和NextHop等指导转发的信息。
  - 使用命令display ospf routing查看OSPF路由表。

[R1]display ospf routing




```
<R1> display ospf routing
OSPF Process 1 with Router ID 1.1.1.1
Routing Tables
Routing for Network
Destination      Cost  Type      NextHop      AdvRouter      Area
1.1.1.1/32       0     stub      1.1.1.1      1.1.1.1        0.0.0.0
10.1.1.0/20      1     Transit   10.1.1.1     1.1.1.1        0.0.0.0
2.2.2.2/32       1     stub      10.1.1.2     2.2.2.2        0.0.0.0

Total Nets: 3
Intra Area: 3  Inter Area: 0  ASE: 0  NSSA: 0
```


- 更多OSPF路由表相关内容请学习HCIP-DataCom。

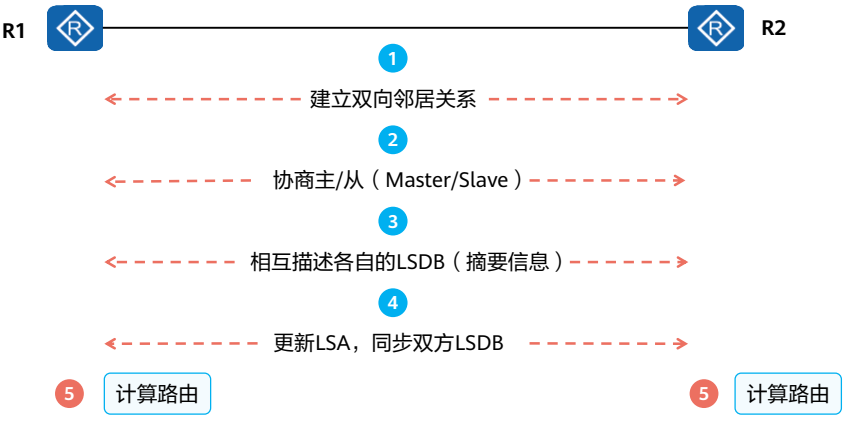
目录

OSPF路由器之间的关系

- 关于OSPF路由器之间的关系有两个重要的概念，邻居关系和邻接关系。
- 考虑一种简单的拓扑，两台路由器直连。在双方互联接口上激活OSPF，路由器开始发送及侦听Hello报文。在通过Hello报文发现彼此后，这两台路由器便形成了邻居关系。
- 邻居关系的建立只是一个开始，后续会进行一系列的报文交互，例如前文提到的DD、LSR、LSU和LS ACK等。当两台路由器LSDB同步完成，并开始独立计算路由时，这两台路由器形成了邻接关系。

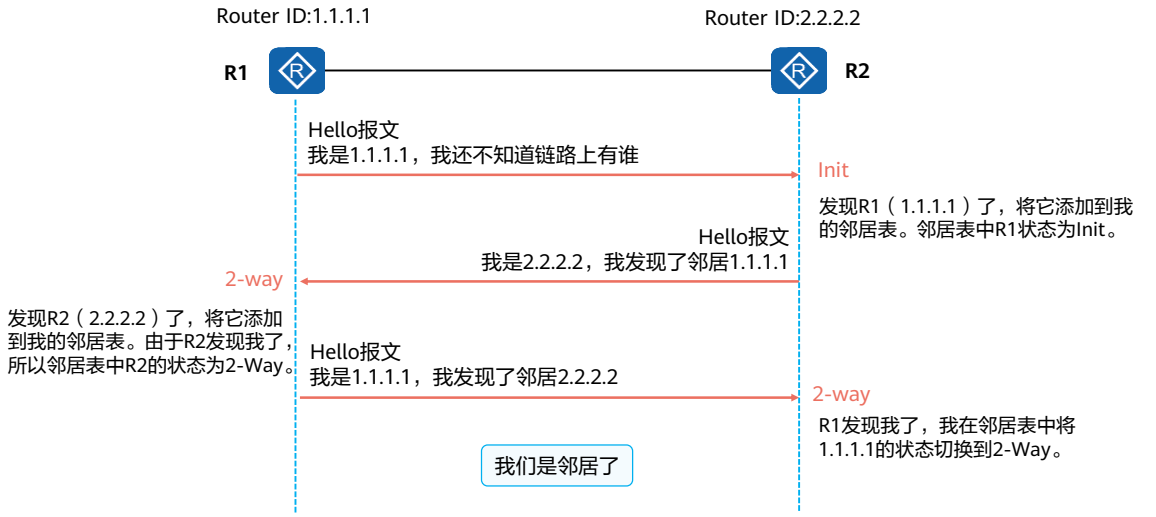
初识OSPF邻接关系建立过程

- OSPF完成邻接关系的建立有四个步骤，建立邻居关系、协商主/从、交互LSDB信息，同步LSDB。



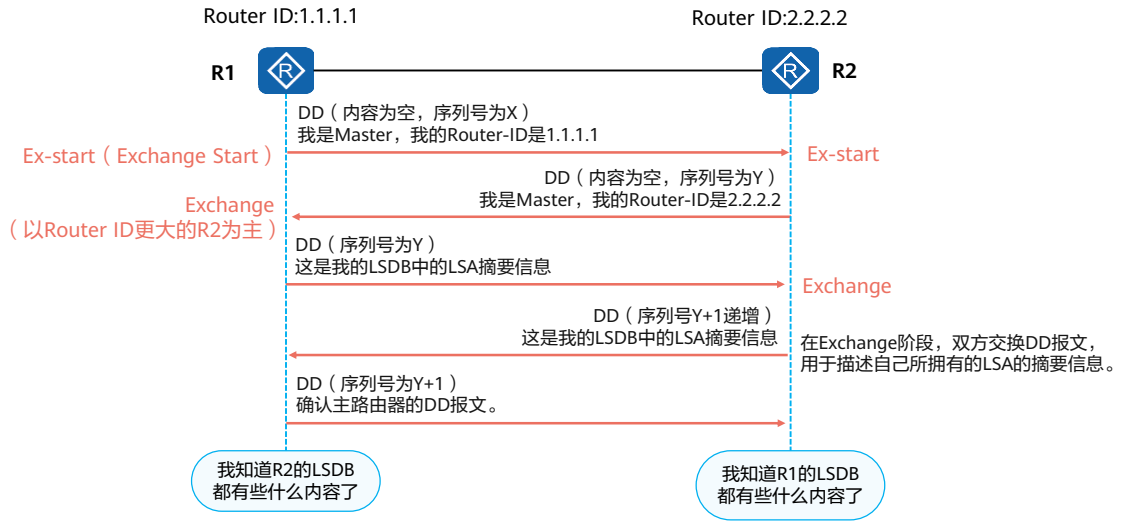
1-4过程由双方交互，5独立完成。

OSPF邻接关系建立流程 - 1



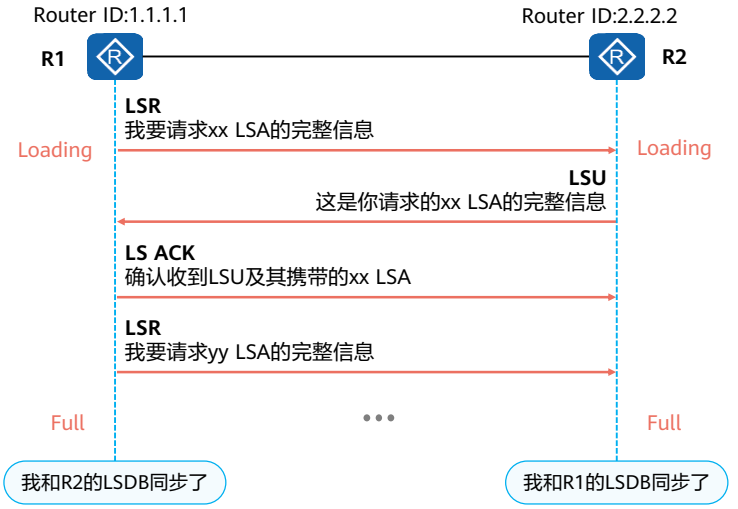
- 当一台OSPF路由器收到其他路由器发来的首个Hello报文时会从初始Down状态切换为Init状态。
- 当OSPF路由器收到的Hello报文中的邻居字段包含自己的Router ID时，从Init切换2-way状态。

OSPF邻接关系建立流程 - 2&3

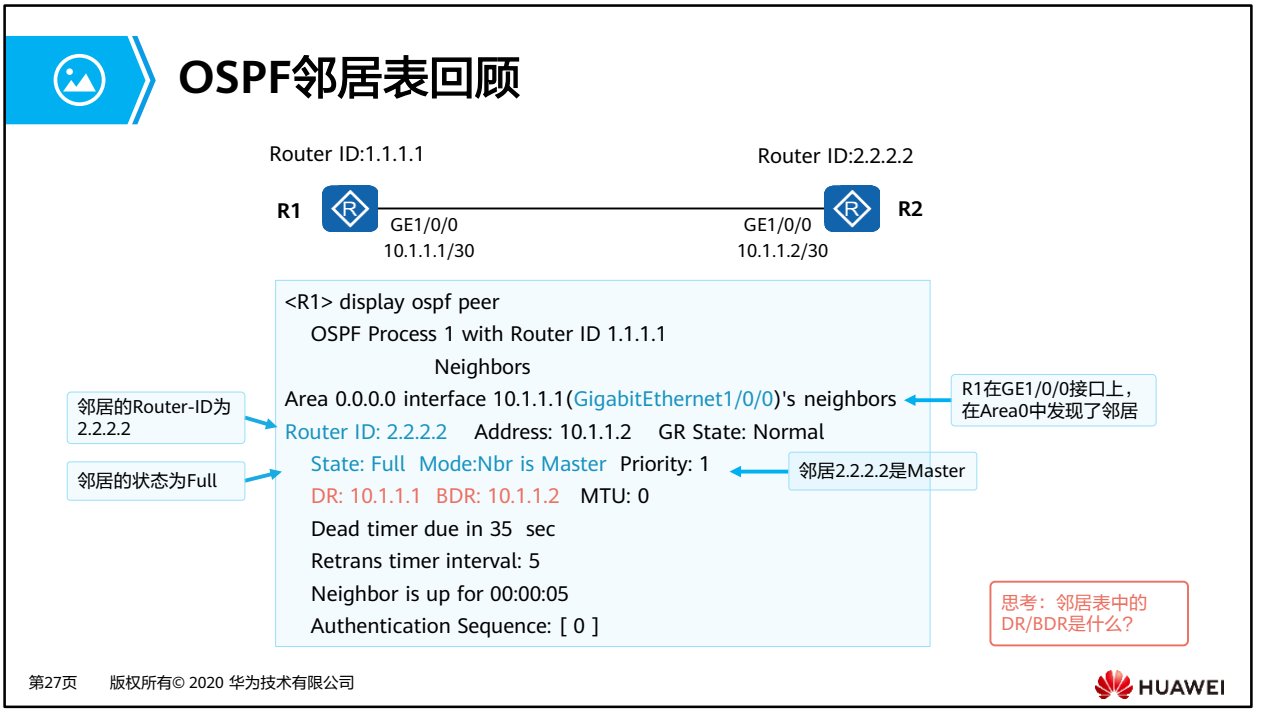


- 邻居状态机从2-way转为Exstart状态后开始主从关系选举：
  - R1向R2发送的第一个DD报文内容为空，其Seq序列号假设为X。
  - R2也向R1发出第一个DD报文，其Seq序列号假设为Y。
  - 选举主从关系的规则是比较Router ID，越大越优。R2的Router ID比R1大，因此R2成为真正的主设备。主从关系比较结束后，R1的状态从Exstart转变为Exchange。
- R1邻居状态变为Exchange后，R1发送一个新的DD报文，包含自己LSDB的描述信息，其序列号采用主设备R2的序列号。R2收到后邻居状态从Exstart转变为Exchange。
- R2向R1发送一个新的DD报文，包含自己LSDB的描述信息，序列号为Y+1。
- R1作为从路由器需要对主路由R2发送的每个DD报文进行确认，回复报文的序列号与主路由R2一致。
- 发送完最后一个DD报文后，R1将邻居状态切换为Loading。

OSPF邻接关系建立流程 - 4



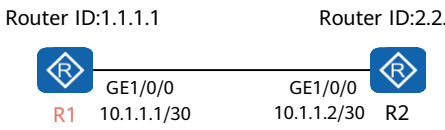
- 邻居状态转变为Loading后，R1向R2发送LSR报文，请求那些在Exchange状态下通过DD报文发现的，但是在本地LSDB中没有的LSA。
- R2收到后向R1回复LSU。在LSU报文中包含被请求的LSA的详细信息。
- R1收到LSU报文后，向R2回复LS ACK报文，确认已接收到，确保信息传输的可靠性。
- 此过程中R2也会向R1发送LSA请求。当两端LSDB完全一致时，邻居状态变为Full，表示成功建立邻接关系。



- 如图所示输入display ospf peer命令之后，各项参数含义如下：
  - OSPF Process 1 with Router ID 1.1.1.1：本地OSPF进程号为1与本端OSPF Router ID为1.1.1.1
  - Router ID：邻居OSPF路由器ID
  - Address：邻居接口地址
  - GR State：使能OSPF GR功能后显示GR的状态（GR为优化功能），默认为Normal
  - State：邻居状态，正常情况下LSDB同步完成之后，稳定停留状态为Full
  - Mode：用于标识本台设备在链路状态信息交互过程中的角色是Master还是Slave
  - Priority：用于标识邻居路由器的优先级（该优先级用于后续DR角色选举）
  - DR：指定路由器
  - BDR：备份指定路由器
  - MTU：邻居接口的MTU值
  - Retrans timer interval：重传LSA的时间间隔，单位为秒
  - Authentication Sequence：认证序列号

OSPF网络类型简介

- 在学习DR和BDR的概念之前，需要首先了解OSPF的网络类型。
- OSPF网络类型是一个非常重要的接口变量，这个变量将影响OSPF在接口上的操作，例如采用什么方式发送OSPF协议报文，以及是否需要选举DR、BDR等。
- 接口默认的OSPF网络类型取决于接口所使用的数据链路层封装。
- 如图所示，OSPF的有四种网络类型，Broadcast、NBMA、P2MP和P2P。

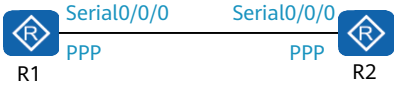



```
[R1-GigabitEthernet1/0/0] ospf network-type ?
  broadcast  Specify OSPF broadcast network
  nbma       Specify OSPF NBMA network
  p2mp       Specify OSPF point-to-multipoint network
  p2p        Specify OSPF point-to-point network
```

OSPF网络类型（1）

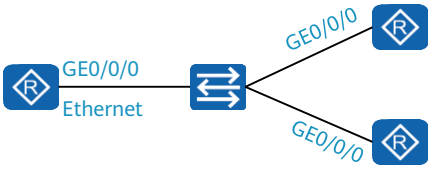
- 一般情况下，链路两端的OSPF接口网络类型必须一致，否则双方无法建立邻居关系。
- OSPF网络类型可以在接口下通过命令手动修改以适应不同网络场景，例如可以将BMA网络类型修改为P2P。

P2P（Point-to-Point，点对点）



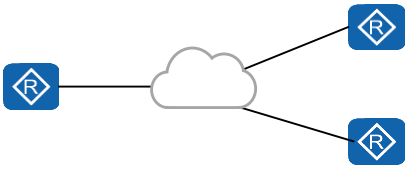
- P2P指的是在一段链路上只能连接两台网络设备的环境。
- 典型的例子是PPP链路。当接口采用PPP封装时，OSPF在该接口上采用的缺省网络类型为P2P 。

BMA（Broadcast Multiple Access，广播式多路访问）



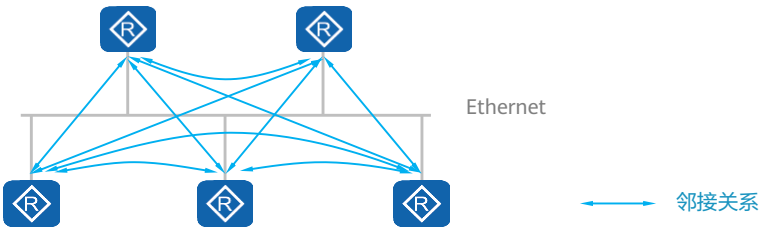
- BMA也被称为Broadcast，指的是一个允许多台设备接入的、支持广播的环境。
- 典型的例子是Ethernet（以太网）。当接口采用Ethernet封装时，OSPF在该接口上采用的缺省网络类型为BMA。

OSPF网络类型（2）

NBMA（Non-Broadcast Multiple Access，非广播式多路访问）	P2MP（Point to Multi-Point，点到多点）
	
• NBMA指的是一个允许多台网络设备接入且不支持广播的环境。 • 典型的例子是帧中继（Frame-Relay）网络。	• P2MP相当于将多条P2P链路的一端进行捆绑得到的网络。 • 没有一种链路层协议会被缺省的认为是P2MP网络类型。该类型必须由其他网络类型手动更改。 • 常用做法是将非全连通的NBMA改为点到多点的网络。

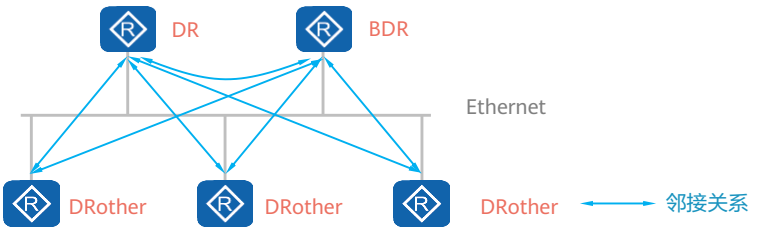
DR与BDR的背景

- MA（Multi-Access）多路访问网络有两种类型：广播型多路访问网络（BMA）及非广播型多路访问网络（NBMA）。以太网（Ethernet）是一种典型的广播型多路访问网络。
- 在MA网络中，如果每台OSPF路由器都与其他的所有路由器建立OSPF邻接关系，便会导致网络中存在过多的OSPF邻接关系，增加设备负担，也增加了网络中泛洪的OSPF报文数量。
- 当拓扑出现变更，网络中的LSA泛洪可能会造成带宽的浪费和设备资源的损耗。



DR与BDR

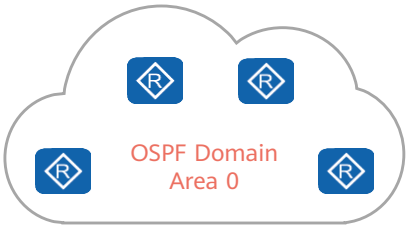
- 为优化MA网络中OSPF邻接关系，OSPF指定了三种OSPF路由器身份，DR（Designated Router，指定路由器）、BDR（Backup Designated Router，备用指定路由器）和DRother路由器。
- 只允许DR、BDR与其他OSPF路由器建立邻接关系。DRother之间不会建立全毗邻的OSPF邻接关系，双方停滞在2-way状态。
- BDR会监控DR的状态，并在当前DR发生故障时接替其角色。



HUAWEI

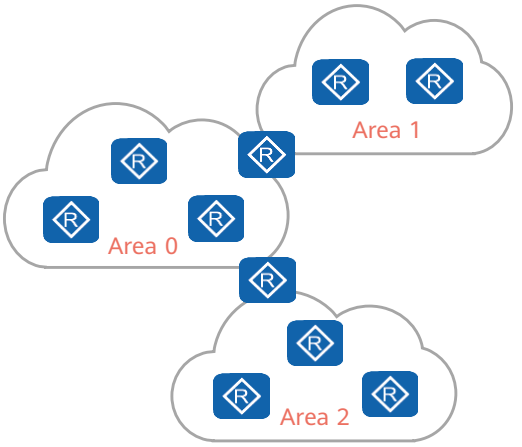
- 选举规则：OSPF DR优先级更高的接口成为该MA的DR，如果优先级相等（默认为1），则具有更高的OSPF Router-ID的路由器（的接口）被选举成DR，并且DR具有非抢占性。

OSPF域与单区域



- OSPF域（Domain）：一系列使用相同策略的连续OSPF网络设备所构成的网络。
- OSPF路由器在同一个区域（Area）内网络中泛洪LSA。为了确保每台路由器都拥有对网络拓扑的一致认知，LSDB需要在区域内进行同步。
- 如果OSPF域仅有一个区域，随着网络规模越来越大，OSPF路由器的数量越来越多，这将导致诸多问题：
  - LSDB越来越庞大，同时导致OSPF路由表规模增加。路由器资源消耗多，设备性能下降，影响数据转发。
  - 基于庞大的LSDB进行路由计算变得困难。
  - 当网络拓扑变更时，LSA全域泛洪和全网SPF重计算带来巨大负担。

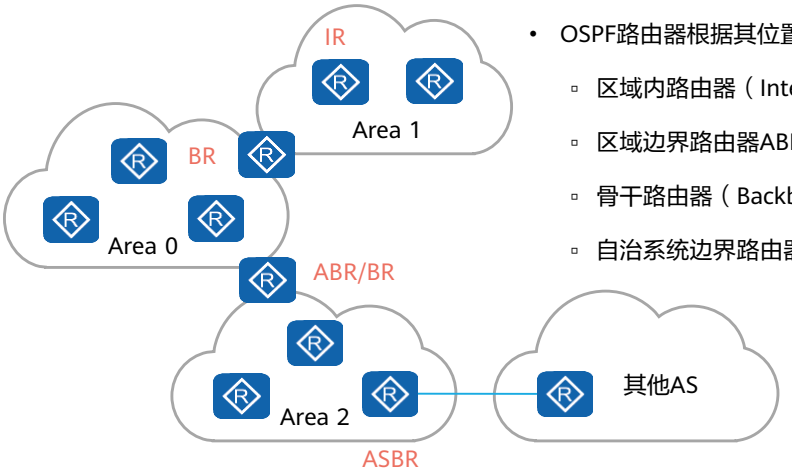
OSPF多区域



- OSPF引入区域（Area）的概念，将一个OSPF域划分成多个区域，可以使OSPF支撑更大规模组网。
- OSPF多区域的设计减小了LSA泛洪的范围，有效的把拓扑变化的影响控制在区域内，达到网络优化的目的。
- 在区域边界可以做路由汇总，减小了路由表规模。
- 多区域提高了网络扩展性，有利于组建大规模的网络。

- 区域的分类：区域可以分为骨干区域与非骨干区域。骨干区域即Area0，除Area0以外其他区域都称为非骨干区域。
- 多区域互联原则：基于防止区域间环路的考虑，非骨干区域与非骨干区域不能直接相连，所有非骨干区域必须与骨干区域相连。

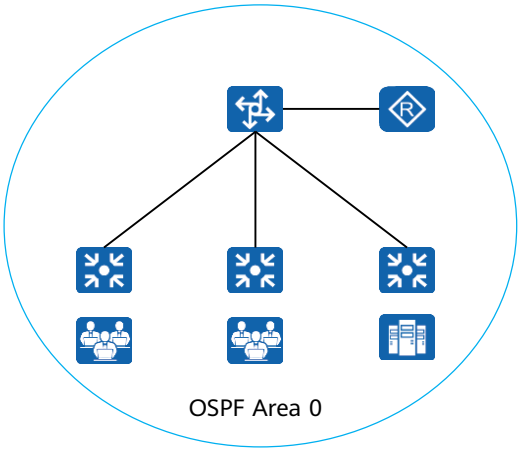
OSPF路由器类型



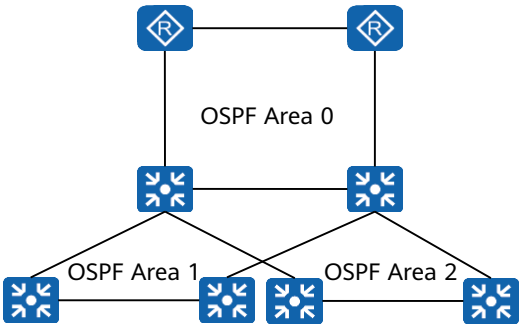
- OSPF路由器根据其位置或功能不同，有这样几种类型：
  - 区域内路由器（Internal Router）
  - 区域边界路由器ABR（Area Border Router）
  - 骨干路由器（Backbone Router）
  - 自治系统边界路由器ASBR（AS Boundary Router）

- 区域内路由器（Internal Router）：该类路由器的所有接口都属于同一个OSPF区域。
- 区域边界路由器ABR（Area Border Router）：该类路由器的接口同时属于两个以上的区域，但至少有一个接口属于骨干区域。
- 骨干路由器（Backbone Router）：该类路由器至少有一个接口属于骨干区域。
- 自治系统边界路由器ASBR（AS Boundary Router）：该类路由器与其他AS交换路由信息。只要一台OSPF路由器引入了外部路由的信息，它就成为ASBR。

OSPF单区域&多区域典型组网



中小型企业网（单区域）



大型企业网（多区域）

- 中小型企业网络规模不大，路由设备数量有限，可以考虑将所有设备都放在同一个OSPF区域。
- 大型企业网络规模大，路由设备数量很多，网络层次分明，建议采用OSPF多区域的方式部署。

目录

OSPF基础配置命令 (1)

1. （系统视图）创建并运行OSPF进程

```
[Huawei] ospf [ process-id | router-id router-id ]
```

*porcess-id*用于标识OSPF进程，默认进程号为1。OSPF支持多进程，在同一台设备上可以运行多个不同的OSPF进程，它们之间互不影响，彼此独立。**router-id**用于手工指定设备的ID号。如果没有通过命令指定ID号，系统会从当前接口的IP地址中自动选取一个作为设备的ID号。

2. （OSPF视图）创建并进入OSPF区域

```
[Huawei-ospf-1] area area-id
```

area命令用来创建OSPF区域，并进入OSPF区域视图。
*area-id*可以是十进制整数或点分十进制格式。采取整数形式时，取值范围是0～4294967295。

3. （OSPF区域视图）指定运行OSPF的接口

```
[Huawei-ospf-1-area-0.0.0.0] network network-address wildcard-mask
```

network命令用来指定运行OSPF协议的接口和接口所属的区域。*network-address*为接口所在的网段地址。
*wildcard-mask*为IP地址的反码，相当于将IP地址的掩码反转（0变1，1变0），例如0.0.0.255表示掩码长度24 bit。

- Router ID的选择顺序是：优先从Loopback地址中选择最大的IP地址作为设备的ID号，如果没有配置Loopback接口，则在接口地址中选取最大的IP地址作为设备的ID号。

OSPF基础配置命令 (2)

4. （接口视图）配置OSPF接口开销

```
[Huawei-GigabitEthernet0/0/0] ospf cost cost
```

ospf cost命令用来配置接口上运行OSPF协议所需的开销。缺省情况下，OSPF会根据该接口的带宽自动计算其开销值cost取值范围是1～65535。

5. （OSPF视图）设置OSPF带宽参考值

```
[Huawei-ospf-1] bandwidth-reference value
```

bandwidth-reference命令用来设置通过公式计算接口开销所依据的带宽参考值。value取值范围是1～2147483648，单位是Mbit/s，缺省值是100Mbit/s。

6. （接口视图）设置接口在选举DR时的优先级

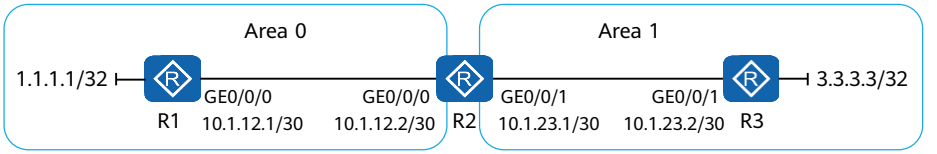
```
[Huawei-GigabitEthernet0/0/0] ospf dr-priority priority
```

ospf dr-priority命令用来设置接口在选举DR时的优先级。*priority*值越大，优先级越高，取值范围是0～255。

OSPF配置案例

案例描述：

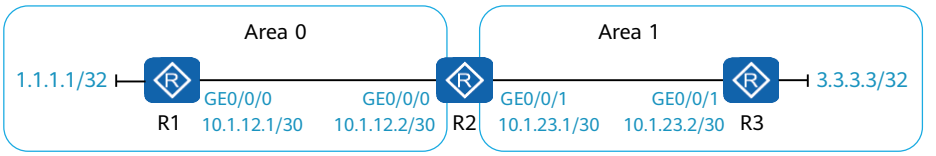
- 有三台路由器R1、R2和R3，其中R1和R3分别连接网络1.1.1.1/32和3.3.3.3/32（LoopBack0模拟），现需要使用OSPF实现这两个网络的互通。具体拓扑如下：



- 配置过程分为三个步骤：配置设备接口、配置OSPF和验证结果。

OSPF配置案例 - 配置接口

配置设备接口 → 配置OSPF → 结果验证



- 根据规划配置R1、R2和R3接口IP地址。

#配置R1的接口
[R1] interface LoopBack 0
[R1-LoopBack0] ip address 1.1.1.1 32
[R1-LoopBack0] interface GigabitEthernet 0/0/0
[R1-GigabitEthernet0/0/0] ip address 10.1.12.1 30

#配置R3的接口
[R3] interface LoopBack 0
[R3-LoopBack0] ip address 3.3.3.3 32
[R3-LoopBack0] interface GigabitEthernet 0/0/1
[R3-GigabitEthernet0/0/1] ip address 10.1.23.2 30

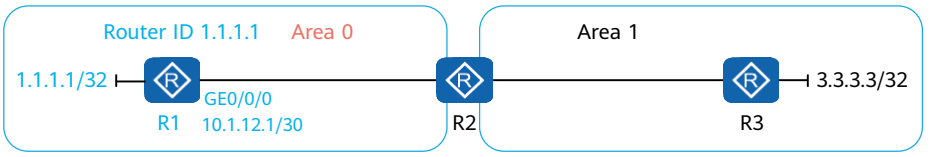
R2配置GE0/0/0和GE0/0/1接口IP地址，详细配置参见备注。

HUAWEI

- 配置R2的接口
  - [R2] interface GigabitEthernet 0/0/0
    [R2-GigabitEthernet0/0/0] ip address 10.1.12.2 30
    [R2-GigabitEthernet0/0/0] interface GigabitEthernet 0/0/1
    [R2-GigabitEthernet0/0/1] ip address 10.1.23.1 30

OSPF配置案例 - 配置OSPF（1）

配置设备接口 → **配置OSPF** → 结果验证



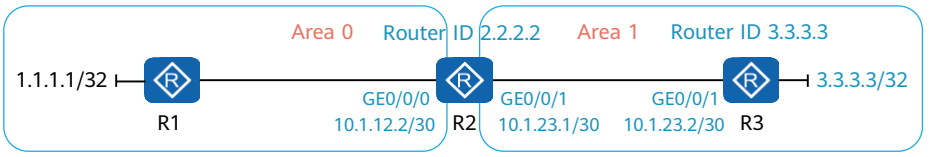
- OSPF参数规划：OSPF进程号为1。R1、R2和R3的Router ID分别为1.1.1.1、2.2.2.2和3.3.3.3。
- 配置步骤：
  - 创建并运行OSPF进程
  - 创建并进入OSPF区域
  - 指定运行OSPF的接口

```
#配置R1 OSPF协议
[R1] ospf 1 router-id 1.1.1.1
[R1-ospf-1] area 0
[R1-ospf-1-area-0.0.0.0] network 1.1.1.1 0.0.0.0
[R1-ospf-1-area-0.0.0.0] network 10.1.12.0 0.0.0.3
```

注意反掩码

OSPF配置案例 - 配置OSPF (2)

配置设备接口 → **配置OSPF** → 结果验证

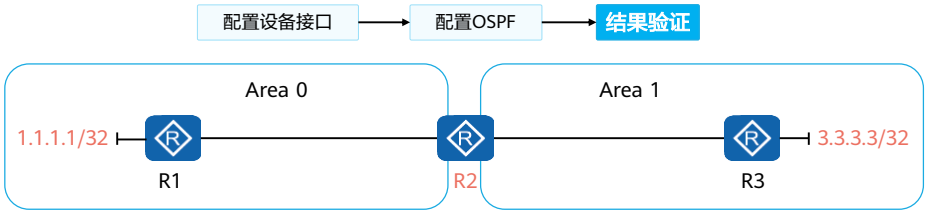


- OSPF多区域的配置请注意在指定区域下通知相应的网段。

#配置R2 OSPF协议
[R2] ospf 1 router-id 2.2.2.2
[R2-ospf-1] area 0
[R2-ospf-1-area-0.0.0.0] network 10.1.12.0 0.0.0.3
[R2-ospf-1-area-0.0.0.0] area 1
[R2-ospf-1-area-0.0.0.1] network 10.1.23.0 0.0.0.3

#配置R3 OSPF协议
[R3] ospf 1 router-id 3.3.3.3
[R3-ospf-1] area 1
[R3-ospf-1-area-0.0.0.1] network 3.3.3.3 0.0.0.0
[R3-ospf-1-area-0.0.0.1] network 10.1.23.0 0.0.0.3

OSPF配置案例 – 结果验证 (1)



- 在路由器R2上查看OSPF邻居表：

```
<R2> display ospf peer brief

        OSPF Process 1 with Router ID 2.2.2.2
                Peer Statistic Information
----------------------------------------------------------------------------
 Area Id          Interface                     Neighbor id      State
 0.0.0.0          GigabitEthernet0/0/0          1.1.1.1          Full
 0.0.0.1          GigabitEthernet0/0/1          3.3.3.3          Full
----------------------------------------------------------------------------
```

邻居的区域ID

邻居的状态
结果验证邻居状态为Full，即成功建立邻接关系。

OSPF配置案例 – 结果验证 (2)

- 在路由器R1上查看路由表，并执行从源1.1.1.1 ping 3.3.3.3。

从OSPF学习到3.3.3.3/32路由

指定源地址为1.1.1.1 ping 3.3.3.3

```
<R1>display ip routing-table
Route Flags: R - relay, D - download to fib
------------------------------------------------------------------------------
Routing Tables: Public
         Destinations : 10        Routes : 10
Destination/Mask    Proto   Pre  Cost      Flags NextHop         Interface
        1.1.1.1/32  Direct  0    0           D   127.0.0.1       LoopBack0
        3.3.3.3/32  OSPF    10   2           D   10.1.12.2       GigabitEthernet 0/0/0
       10.1.12.0/30 Direct  0    0           D   10.1.12.1       GigabitEthernet 0/0/0
...
<R1>ping -a 1.1.1.1 3.3.3.3
  PING 3.3.3.3: 56  data bytes, press CTRL_C to break
    Reply from 3.3.3.3: bytes=56 Sequence=1 ttl=254 time=50 ms
...
```

思考题

1. （多选）在建立OSPF邻居和邻接关系的过程中，稳定的状态是（　　）

   A. Exstart

   B. Two-way

   C. Exchange

   D. Full

2. （多选）以下哪种情况下路由器之间会建立邻接关系（　　）

   A. 点到点链路上的两台路由器

   B. 广播型网络中的DR和BDR

   C. NBMA网络中的DRother和DRother

   D. 广播型网络中的BDR和DRother

1. BD
2. ABD

本章总结

- OSPF是现网中使用广泛的路由协议之一，本章节帮助您初步了解OSPF的基本概念、应用场景和基础配置。
- Router ID、区域、OSPF邻居表、LSDB表和OSPF路由表是OSPF的基本概念。能够阐述OSPF的邻居和邻接关系建立过程，可以帮助您更好的理解链路状态路由协议。
- OSPF有更多有趣的细节，例如LSA的类型、SPF的计算过程和OSPF的特殊区域等。如果您对更多的OSPF知识感兴趣，请继续学习华为HCIP-DataCom认证。


谢 谢
www.huawei.com

以太网交换基础

前言

- 在网络中传输数据时需要遵循一些标准，以太网协议定义了数据帧在以太网上的传输标准，了解以太网协议是充分理解数据链路层通信的基础。以太网交换机是实现数据链路层通信的主要设备，了解以太网交换机的工作原理也是十分必要的。
- 在本课程中，将介绍以太网协议的相关概念、MAC地址的类型、二层交换机的工作流程以及二层交换机的工作原理。

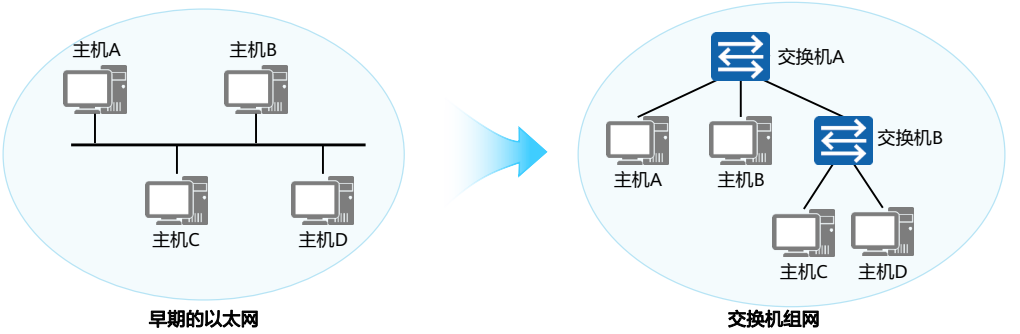
目标

- 学完本课程后，您将能够：
  - 描述以太网的基本概念
  - 区分MAC地址的类型
  - 描述二层交换机的工作流程
  - 描述MAC地址表的构成与形成过程

目录

以太网协议

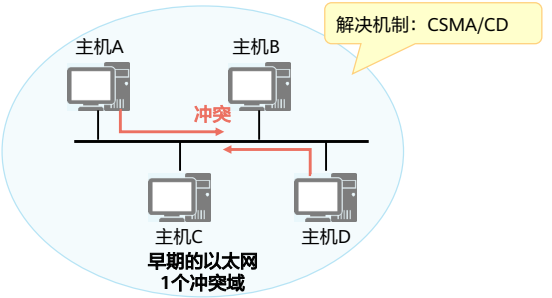
- 以太网是当今现有局域网（Local Area Network, LAN）采用的最通用的通信协议标准，该标准定义了在局域网中采用的电缆类型和信号处理方法。
- 以太网是建立在CSMA/CD (Carrier Sense Multiple Access/Collision Detection，载波监听多路访问/冲突检测)机制上的广播型网络。

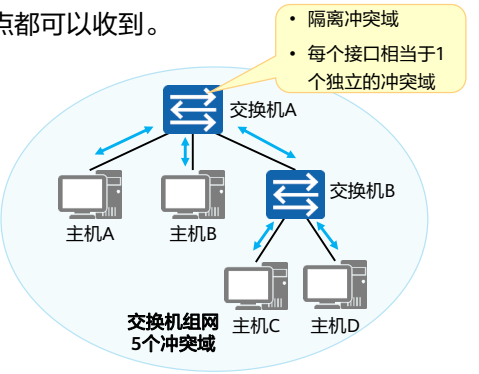


- 早期的以太网：
  - 以太网是建立在CSMA/CD机制上的广播型网络。冲突的产生是限制以太网性能的重要因素，早期的以太网设备如集线器HUB是物理层设备，不能隔绝冲突扩散，限制了网络性能的提高。
- 交换机组网：
  - 交换机做为一种能隔绝冲突的二层网络设备，极大的提高了以太网的性能，并替代HUB成为主流的以太网设备。但是交换机对网络中的广播数据流量不做任何限制，这也影响了网络的性能。

冲突域

- 冲突域是指连接在同一共享介质上的所有节点的集合，冲突域内所有节点竞争同一带宽，一个节点发出的报文（无论是单播、组播、广播），其余节点都可以收到。





- 在传统的以太网中，同一介质上的多个节点共享链路带宽，争用链路的使用权，这样就会发生冲突。
- 同一介质上的节点越多，冲突发生的概率越大。

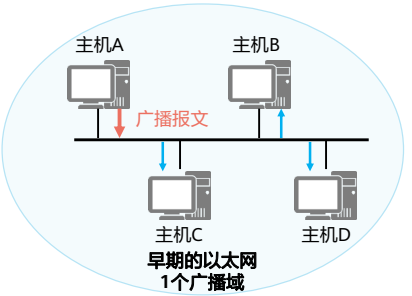
交换机不同的接口发送和接收数据独立，各接口属于不同的冲突域，因此有效地隔离了网络中物理层冲突域，使得通过它互连的主机（或网络）之间不必再担心流量大小对于数据发送冲突的影响。

HUAWEI

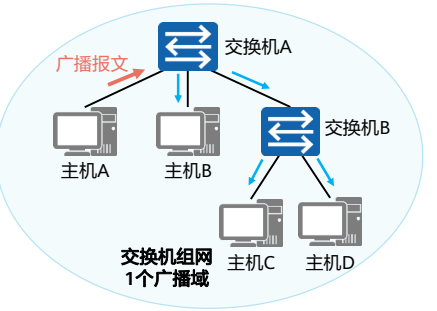
- 在共享网络，以太网络使用CSMA/CD技术，避免冲突问题。CSMA/CD的基本工作过程如下：
  - 终端设备不停的检测共享线路的状态。
    - 如果线路空闲则发送数据。
    - 如果线路不空闲则一直等待。
  - 如果有另外一个设备同时发送数据，两个设备发送的数据必然产生冲突，导致线路上的信号不稳定。
  - 终端设备检测到这种不稳定之后，马上停止发送自己的数据。
  - 终端设备发送一连串干扰脉冲，然后等待一段时间之后再进行发送数据。发送干扰脉冲的目的是为了通知其他设备，特别是跟自己在同一个时刻发送数据的设备，线路上已经产生了冲突。
- CSMA/CD的工作原理可简单总结为：先听后发，边发边听，冲突停发，随机延迟后重发。

广播域

- 广播报文所能到达的整个访问范围称为二层广播域，简称广播域，同一广播域内的主机都能收到广播报文。



在传统的以太网中，同一介质上的多个节点共享链路，一台设备发出的广播报文，所有设备均会收到。



交换机对广播报文会向所有的接口都转发，所以交换机的所有接口连接的节点属于一个广播域。

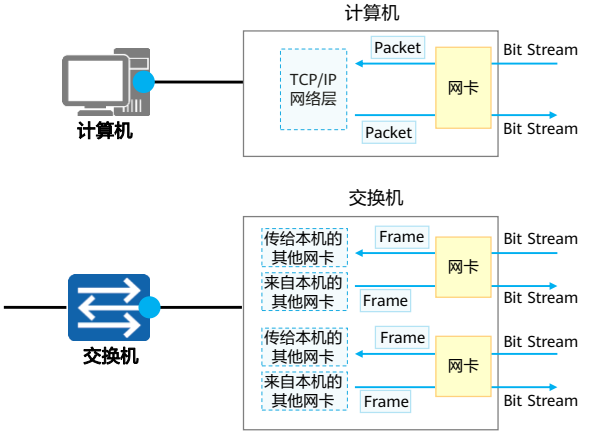
HUAWEI

- 全1MAC地址FF-FF-FF-FF-FF-FF为广播地址，所有节点都会处理目的地址为广播地址的数据帧，该数据帧所能到达的整个访问范围称为二层广播域，简称广播域。

- 注：MAC (Media Access Control)地址，在网络中唯一标识一个网卡，每个网卡都需要并会有唯一的一个MAC地址，后面课程内容会具体讲。

以太网卡

- 网络接口卡（Network Interface Card，NIC)也称为“网卡”，是计算机、交换机、路由器等网络设备与外部网络世界相连的关键部件。



- **网络接口**
  - 简称“网口”或“接口”或“端口”。
- **网卡**
  - 每个网口都有一块网卡与之对应。
  - 计算机或交换机通过网卡来转发数据。

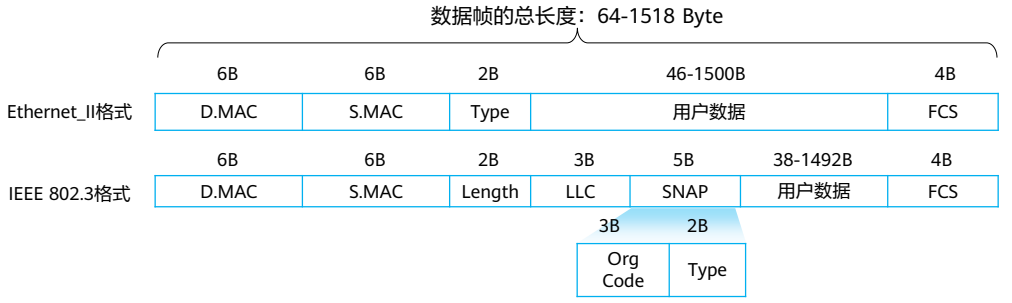
HUAWEI

- 网卡有很多类型，本章节我们提及的都指以太网接口卡，简称以太网卡或以太卡。
- 我们所说的交换机也均为以太网交换机，则交换机上每个转发数据的网口所使用的网卡都是以太网卡。

目录

以太网帧格式

- 以太网技术所使用的帧称为以太网帧 (Ethernet Frame)，或简称以太帧。
- 以太帧的格式有两个标准：Ethernet_II格式和IEEE 802.3格式。



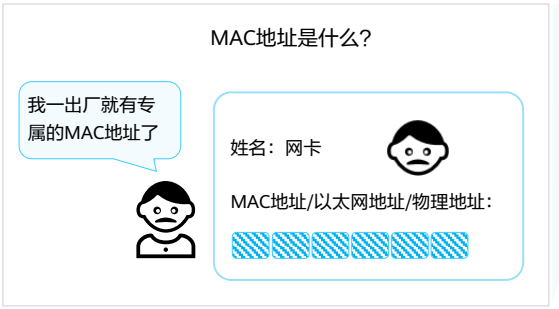
  HUAWEI

- 在以太网中，数据通信的基本单位是以太网帧（Frame）。以太帧的格式有两个标准：Ethernet_II格式和IEEE 802.3格式，因此协议规定以太网帧的数据格式如图所示。
- Ethernet II以太帧：
  - DMAC：6字节，目的MAC地址，6字节，该字段标识帧的接收者。
  - SMAC：6字节，源MAC地址，6字节，该字段标识帧的发送者。
  - Type：2字节，协议类型。常见值：
    - 0x0800：Internet Protocol Version 4 (IPv4)；
    - 0x0806：Address Resolution Protocol (ARP)。
- IEEE 802.3 LLC以太帧：
  - 逻辑链路控制LLC（Logical Link Control）由目的服务访问点DSAP（Destination Service Access Point）、源服务访问点SSAP（Source Service Access Point）和Control字段组成。
    - DSAP：1字节，目的服务访问点，若后面类型为IP值设为0x06。服务访问点的功能类似于Ethernet II帧中的Type字段或TCP/UDP传输协议中的端口号。
    - SSAP：1字节，源服务访问点，若后面类型为IP值设为0x06。
    - Ctrl：1字节，该字段值通常设为0x03，表示无连接服务的IEEE 802.2无编号数据格式。

    - SNAP（Sub-network Access Protocol）由机构代码（Org Code）和类型（Type）字段组成。
        - Org Code三个字节都为0。
        - Type字段的含义与Ethernet_II帧中的Type字段相同。
- 数据帧的总长度为64-1518字节，这样设计的原因是什么？（另外，以太网口的最大传输单元是1500字节，即MTU=1500B。）
    - 以太网中，最小帧长为64字节，这是由最大传输距离和CSMA/CD机制共同决定的。
        - 规定最小帧长是为了避免这种情况发生：A站点已经将一个数据包的最后一个Bit发送完毕，但这个报文的第一个Bit还没有传送到距离很远的B站点。B站点认为线路空闲继续发送数据，导致冲突。
        - 高层协议必须保证Data域至少包含46字节，这样加上以太网帧头的14字节和帧尾的4字节校验码正好满足64字节的最小帧长，如果实际数据不足46个字节，则高层协议必须填充一些数据单元。
    - 而出于对传输效率和传输可靠性的折中考虑，使得以太网帧的最大长度为1518字节，对应IP数据包就是1500字节。
        - 较大的帧长度，数据的有效传输效率会更高；但是数据帧过长，传输时会占用共享链路过多的时间，对时延敏感应用造成极大的影响。
        - 因此最终选择了一个折中的长度：1518字节的数据帧长，对应1500字节的IP数据包长度，这就是最大传输单元MTU的由来。

什么是MAC地址

- MAC (Media Access Control)地址在网络中唯一标识一个网卡，每个网卡都需要并拥有唯一的一个MAC地址。
- 一块网卡的MAC地址是具有全球唯一性的。



MAC地址

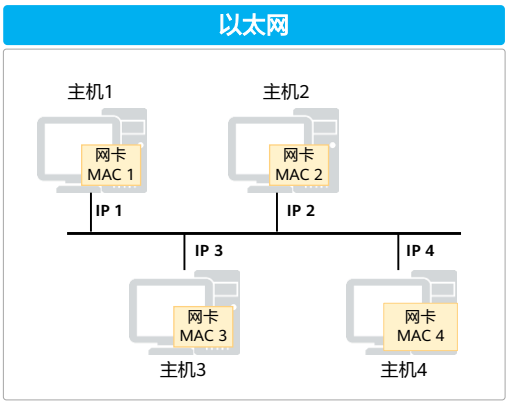
如同每个人都有身份证号码来标识自己一样，每块网卡也拥有一个用来标识自己的号码，即MAC地址。

HUAWEI

- MAC地址是在IEEE 802标准中定义并规范的，凡是符合IEEE 802标准的以太网卡，都必须拥有一个MAC地址，用MAC地址来定义网络设备的位置。不同的网卡，MAC地址也不同。

IP地址 vs MAC地址

- 每个以太网设备在出厂时都有一个唯一的MAC地址，但在设备接入网络时，会同时为每台主机再分配一个IP地址，这个原因是什么?



IP地址的特点:

- IP地址是唯一的
- IP地址可变
- 基于网络拓扑进行IP地址分配

MAC地址的特点:

- MAC地址是唯一的
- MAC地址不可变
- 基于制造商进行MAC地址分配

- 每个以太网设备在出厂时都有一个唯一的MAC地址，那为什么还需要为每台主机再分配一个IP地址呢？或者说每台主机都分配唯一的IP地址了，为什么还要在网络设备 (如：网卡) 生产时内嵌一个唯一的MAC地址呢?
- 主要原因有:
  - IP地址是根据网络的拓朴结构分配的，MAC地址是根据制造商分配的，若路由选择建立在设备制造商的基础上，这种方案是不可行的。
  - 当存在两层地址寻址时，设备更灵活，易于移动和维修。
    - 例如，如果一个以太网卡坏了，可以被更换，而无须更换一个新的IP地址；如果一个IP主机从一个网络移到另一个网络，可以给它一个新的IP地址，而无须换一个新的网卡。
- 总结:
  - IP地址的作用是唯一标识网络中的一个节点，可以通过IP地址进行不同网段的数据访问。
  - MAC地址的作用是唯一标识一个网卡，可以通过MAC地址进行同网段的数据访问。

MAC地址表示

- 一个MAC地址有48 bit，6 Byte。
- MAC地址通常采用“十六进制”+“-”表示。

如: 00-1E-10-DD-DD-02，或 001E-10DD-DD02

十六进制	00	1E	10	DD	DD	02	6 Byte
二进制	0000 0000	0001 1110	0001 0000	1101 1101	1101 1101	0000 0010	48 bit

十六进制与二进制的转换

幂	2^3	2^2	2^1	2^0	2^3	2^2	2^1	2^0
	8	4	2	1	8	4	2	1
位	0	0	0	1	1	1	1	0
	= 1				= 8+4+2=14=E			

HUAWEI

- MAC地址由48比特（6个字节）长，12位的16进制数字组成。

MAC地址构成及分类

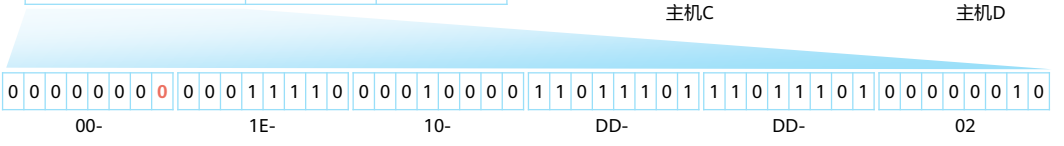
- OUI（Organizationally Unique Identifier）：厂商代码，由IEEE分配，3 Byte，24 bit。
- 制造商分配：3 Byte，24 bit

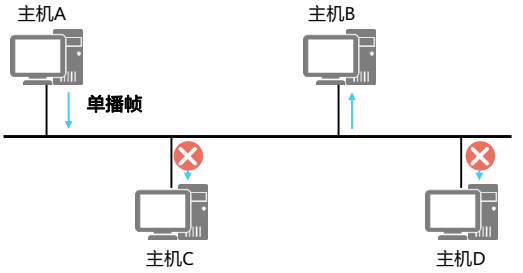
- MAC地址分类：

- 一个制造商在生产制造网卡之前，必须先向IEEE注册，以获取一个长度为24bit (3字节)的厂商代码，也称为OUI。
- 后24bit由厂商自行分派，是各个厂商制造的所有网卡的唯一编号。
- MAC地址可以分为3种类型：
  - 单播MAC地址：也称物理MAC地址，这种类型的MAC地址唯一的标识了以太网上的一个终端，该地址为全球唯一的硬件地址。
    - 单播MAC地址用于标识链路上的一个单一节点。
    - 目的MAC地址为单播MAC地址的帧发往一个单一的节点。
    - 单播MAC地址可以作为源或目的地址。
    - 注意：单播MAC地址具有全球唯一性，当一个二层网络中接入了两台具有相同MAC地址的终端时（例如误操作等），将会引发通信故障（例如这两台终端无法相互通信），且其他设备与它们之间的通信也会存在问题。
  - 广播MAC地址：全1的MAC地址（FF-FF-FF-FF-FF-FF），用来表示局域网上的所有终端设备。
    - 广播MAC地址可以理解为一种特殊的组播MAC地址。
    - 其具体格式为：FFFF-FFFF-FFFF。
    - 目的MAC地址为广播MAC地址的帧发往链路上的所有节点。
  - 组播MAC地址：除广播地址外，第7bit为1的MAC地址为组播MAC地址（例如01-00-00-00-00-00），用来代表局域网上的一组终端。
    - 组播MAC地址用于标识链路上的一组节点。
    - 目的MAC地址为组播MAC地址的帧发往一组节点。
    - 组播MAC地址不能作为源地址，只能作为目的地址。

单播以太帧

- 简称：单播帧
- 目的MAC地址为单播MAC地址的帧

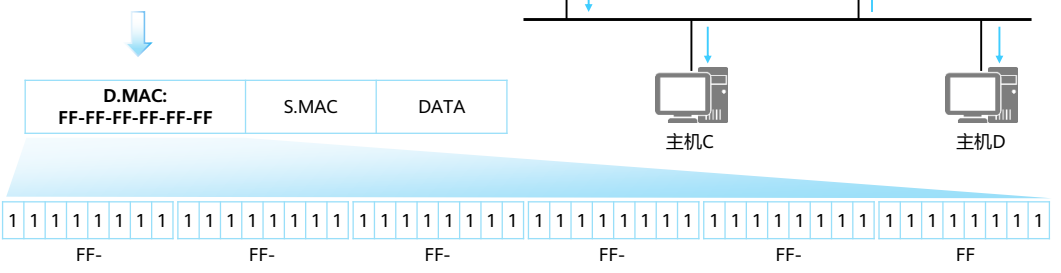




- 局域网上的帧可以通过三种方式发送。
- 第一种是单播，指从单一的源端发送到单一的目的端。
  - 每个主机接口由一个MAC地址唯一标识，MAC地址的OUI中，第一字节第8个比特表示地址类型。对于主机MAC地址，这个比特固定为0，表示目的MAC地址为此MAC地址的帧都是发送到某个唯一的目的端。

广播以太帧

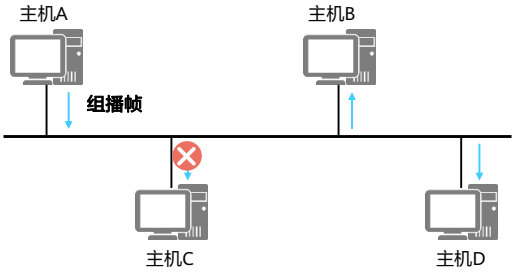
- 简称：广播帧
- 目的MAC地址为广播MAC地址的帧

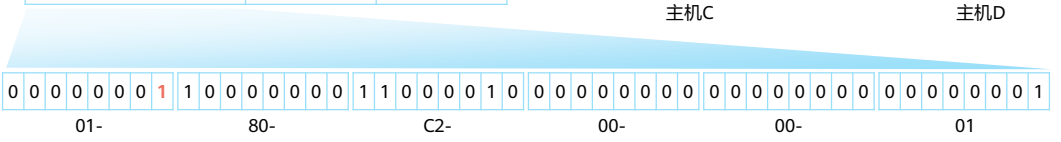


- 第二种发送方式是广播，表示帧从单一的源发送到共享以太网上的所有主机。
  - 广播帧的目的MAC地址为十六进制的FF-FF-FF-FF-FF-FF，所有收到该广播帧的主机都要接收并处理这个帧。
  - 广播方式会产生大量流量，导致带宽利用率降低，进而影响整个网络的性能。
  - 当需要网络中的所有主机都能接收到相同的信息并进行处理的情况下，通常会使用广播方式。

组播以太帧

- 简称：组播帧
- 目的MAC地址为组播MAC地址的帧

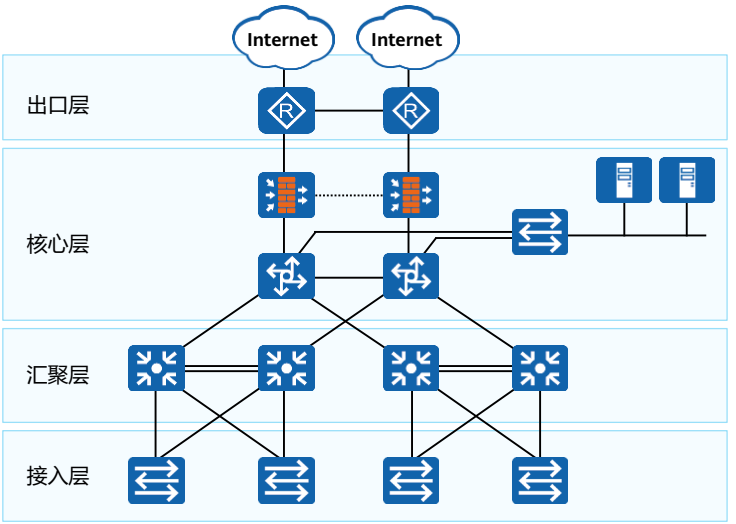




- 第三种发送方式为组播，组播比广播更加高效。
  - 组播转发可以理解为选择性的广播，主机侦听特定组播地址，接收并处理目的MAC地址为该组播MAC地址的帧。
  - 组播MAC地址和单播MAC地址是通过第一字节中的第8个比特区分的。组播MAC地址的第8个比特为1。
  - 当需要网络上的一组主机（而不是全部主机）接收相同信息，并且其他主机不受影响的情况下通常会使用组播方式。

目录

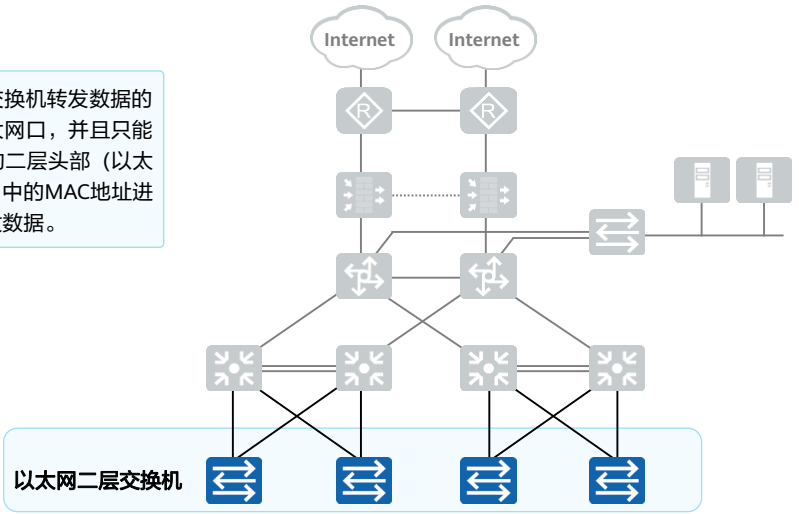
园区网典型架构



- 一个典型的园区数据网络由路由器、交换机、防火墙等设备构成，通常会采用多层架构，包括：接入层、汇聚层、核心层和出口层。

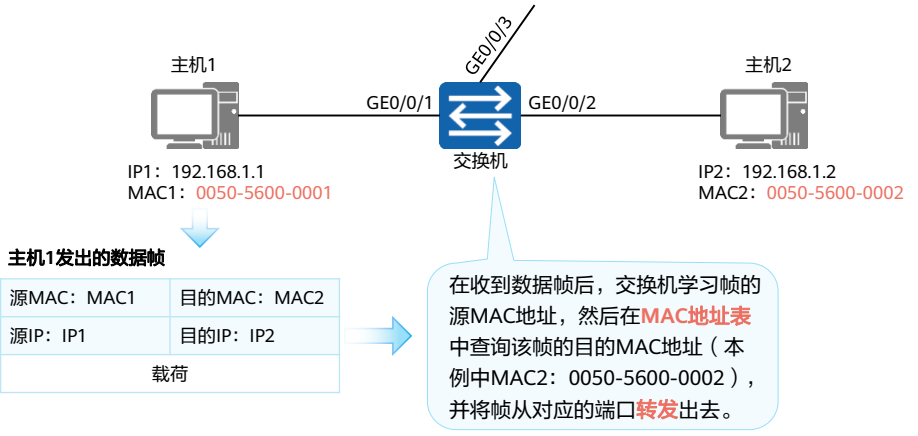
以太网二层交换机

以太网二层交换机转发数据的端口都是以太网口，并且只能够针对数据的二层头部（以太网数据帧头）中的MAC地址进行寻址并转发数据。



- 以太网二层交换机：
  - 在园区网络中，交换机一般来说是距离终端用户最近的设备，用于终端接入园区网，接入层的交换机一般为二层交换机。
  - 二层交换设备工作在TCP/IP对等模型的第二层，即数据链路层，它对数据包的转发是建立在MAC（ Media Access Control ）地址基础之上的。
- 以太网三层交换机：
  - 不同局域网之间的网络互通需要由路由器来完成。随着数据通信网络范围的不断扩大，网络业务的不断丰富，网络间互访的需求越来越大，而路由器由于自身成本高、转发性能低、接口数量少等特点无法很好的满足网络发展的需求。因此出现了三层交换机这样一种能实现高速三层转发的设备。
- 注意：本课程中所涉及的交换机，均指以太网二层交换机。

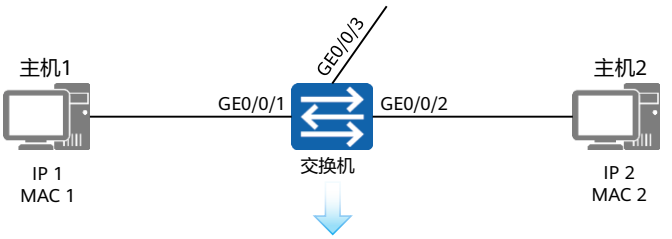
交换机的工作原理



- 二层交换机工作在数据链路层，它对数据帧的转发是建立在MAC地址基础之上的。交换机不同的接口发送和接收数据是独立的，各接口属于不同的冲突域，因此有效地隔离了网络中的冲突域。
- 二层交换设备通过学习以太网数据帧的源MAC地址来维护MAC地址与接口的对应关系（保存MAC与接口对应关系的表称为MAC地址表），通过其目的MAC地址来查找MAC地址表决定向哪个接口转发。

MAC地址表

- 每台交换机中都有一个MAC地址表，存放了MAC地址与交换机端口编号之间的映射关系。



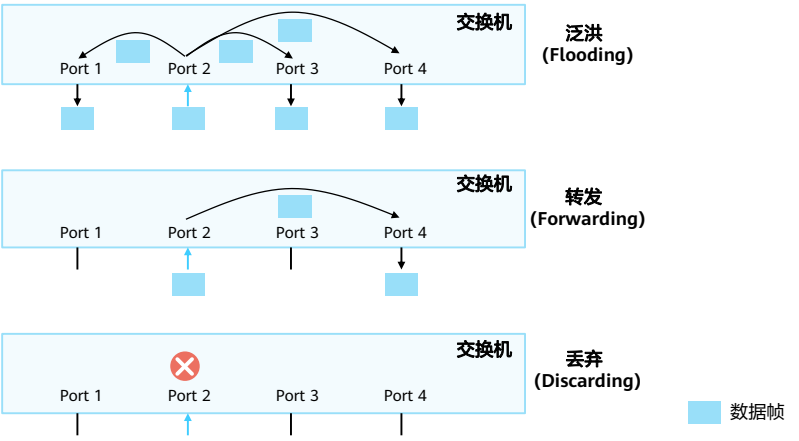
MAC地址	Port
MAC 1	GE0/0/1
MAC 2	GE0/0/2
.....	.....

HUAWEI

- MAC地址表记录了交换机学习到的其他设备的MAC地址与接口的对应关系。交换机在转发数据帧时，根据数据帧的目的MAC地址查询MAC地址表。如果MAC地址表中包含与该帧目的MAC地址对应的表项，则直接通过该表项中的出接口转发该报文；如果MAC地址表中没有包含该帧目的MAC地址对应的表项时，交换机将采取泛洪方式在除接收接口外的所有接口发送该报文。

交换机的3种数据帧处理行为

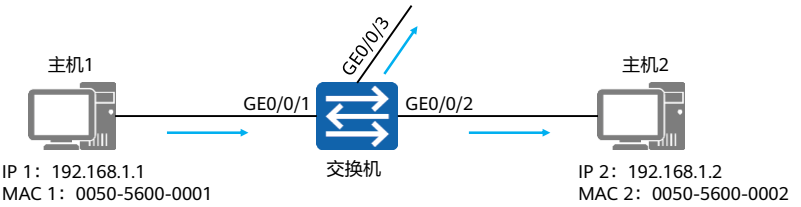
- 交换机对于从传输介质进入某一端口的帧的处理行为一共有3种：



HUAWEI

- 交换机会通过传输介质进入其端口的每一个帧都进行转发操作，交换机的基本作用就是用来转发数据帧。
- 交换机对帧的处理行为一共有三种：泛洪（Flooding），转发（Forwarding），丢弃（Discarding）。
  - 泛洪：交换机把从某一端口进来的帧通过所有其它的端口转发出去（注意，“所有其它的端口”是指除了这个帧进入交换机的那个端口以外的所有端口）。
  - 转发：交换机把从某一端口进来的帧通过另一个端口转发出去（注意，“另一个端口”不能是这个帧进入交换机的那个端口）。
  - 丢弃：交换机把从某一端口进来的帧直接丢弃。

泛洪



1 主机1发出的数据帧

源MAC：MAC1

目的MAC：MAC2

或

源MAC：MAC1

目的MAC：FF-FF-FF-FF-FF-FF

2 交换机查MAC地址表

MAC地址	Port
MAC 1	GE0/0/1

3 交换机处理数据帧行为

未知单播帧

- **如果接收的是单播帧：**

  交换机在MAC地址表中查不到这个帧的目的MAC地址，则交换机对该单播帧执行泛洪操作。

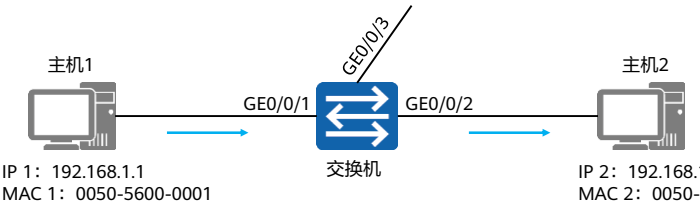
- **如果接收的是广播帧：**

  交换机不会去查MAC地址表，直接对该广播帧执行泛洪操作。

HUAWEI

- 如果从传输介质进入交换机的某个端口的帧是一个单播帧，交换机会去MAC表查这个帧的目的MAC地址。如果查不到这个MAC地址，则交换机将对该单播帧执行泛洪操作。
- 如果从传输介质进入交换机的某个端口的帧是一个广播帧，交换机不会去查MAC地址表，而是直接对该广播帧执行泛洪操作。
- 如图所示：
  - 场景一：主机1想要访问主机2，发送单播数据帧，交换机收到后，若MAC地址表中查不到对应的表项，则会泛洪该数据帧。
  - 场景二：主机1想要访问主机2，但不知道对应的MAC地址，则会发送ARP请求报文，该报文为广播数据帧，交换机收到后，则会泛洪该数据帧。

转发



1 主机1发出的数据帧

源MAC：MAC1

目的MAC：MAC2

2 交换机查MAC地址表

MAC地址	Port
MAC 1	GE0/0/1
MAC 2	GE0/0/2

3 交换机处理数据帧行为

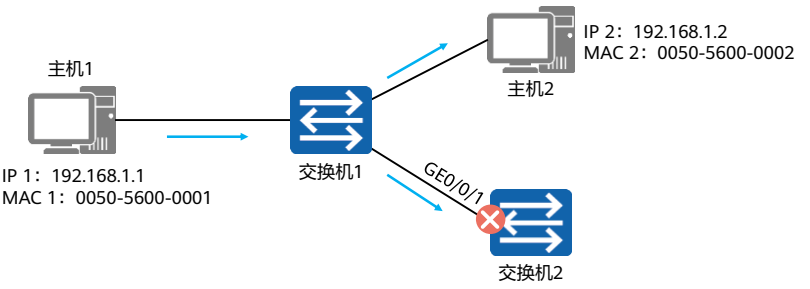
- **如果接收的是单播帧：**

  交换机在MAC地址表中查到了这个帧的目的MAC地址，并且表中对应的端口编号不是这个帧从传输介质进入交换机的那个端口编号，则交换机对该单播帧执行转发操作。

HUAWEI

- 如果从传输介质进入交换机的某个端口的帧是一个单播帧，则交换机会去MAC表查这个帧的目的MAC地址。如果查到了这个MAC地址表，则比较这个MAC地址在MAC地址表中对应的端口编号是不是这个帧从传输介质进入交换机的那个端口的端口编号。如果不是，则交换机执行转发操作（将该帧送至该帧目的MAC地址在MAC地址表中对应的那个端口，并从那个端口发送出去）。
- 如图所示：
  - 主机1想要访问主机2，发送单播数据帧，交换机收到后，在MAC地址表中查到了对应的表项，则会点对点转发该数据帧。

丢弃



1 主机1发出的数据帧

源MAC：MAC1

目的MAC：MAC2

2 交换机2查MAC地址表

MAC地址	Port
MAC 2	GE0/0/1

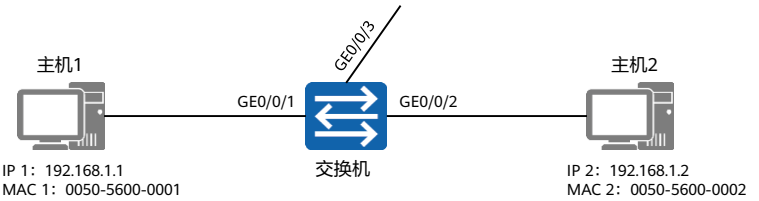
3 交换机处理数据帧行为

- **如果接收的是单播帧：**
- 交换机在MAC地址表中查到了这个帧的目的MAC地址，但是表中对应的端口编号是该帧从传输介质进入交换机的那个端口编号，则交换机对该单播帧执行丢弃操作。

HUAWEI

- 如果从传输介质进入交换机的某个端口的帧是一个单播帧，则交换机会去MAC表查这个帧的目的MAC地址。如果查到了这个MAC地址表，则比较这个MAC地址在MAC地址表中对应的端口编号是不是这个帧从传输介质进入交换机的那个端口的端口编号。如果是，则交换机将对该帧执行丢弃操作。
- 如图所示：
  - 主机1想要访问主机2，发送单播数据帧，交换机1收到后，若MAC地址表中查不到对应的表项，则会泛洪该数据帧。
  - 交换机2收到该数据帧后，发现目的MAC地址对应的端口就是接收数据帧的端口，则会丢弃该数据帧。

交换机的MAC地址学习（1）



交换机的MAC地址表

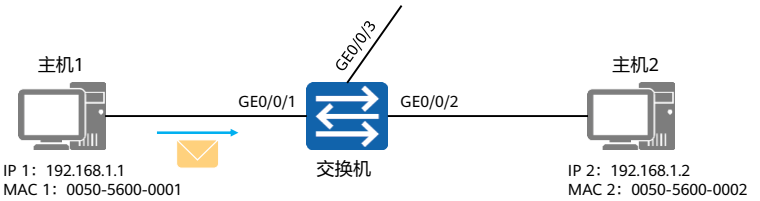
MAC地址	Port

1

初始情况，交换机的MAC地址表是空的。

- 初始状态下，交换机并不知道所连接主机的MAC地址，所以MAC地址表为空。

交换机的MAC地址学习 (2)



主机1发出的数据帧

源MAC：MAC1

目的MAC：MAC2

（注：假设主机1已知主机2的MAC地址）

交换机查MAC地址表

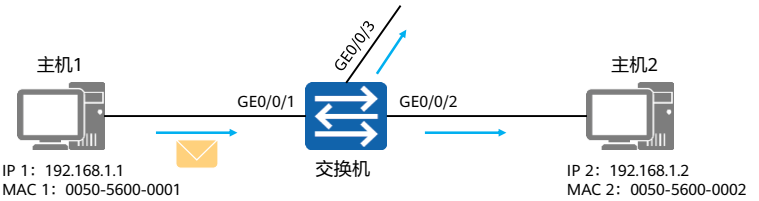
MAC地址	Port

2

- 主机1发送数据帧给主机2。
- 交换机GE0/0/1口接收到数据帧后，在MAC地址表中查询该帧的目的MAC地址，发现没有对应表项，则收到的数据帧是“未知单播帧”。

- 主机1想要发送数据给主机2（假设已知对端的IP地址和MAC地址），会封装数据帧，包含自己的源IP地址和源MAC地址。
- 交换机收到后会查自己的MAC地址表，发现没有对应表项，则收到的数据帧是“未知单播帧”。

交换机的MAC地址学习 (3)



主机1发出的数据帧

源MAC：MAC1

目的MAC：MAC2

交换机查MAC地址表

MAC地址	Port
MAC 1	GE0/0/1

3

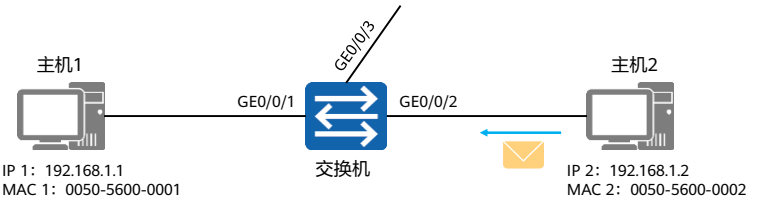
- 交换机在MAC地址表中没有查到对应表项，则交换机对该单播帧执行泛洪操作。
- 同时，交换机学习该数据帧的源MAC地址，并创建对应的MAC地址表项，与接收口GE0/0/1关联。

HUAWEI

- 由于收到的数据帧是“未知单播帧”，因此交换机会泛洪该数据帧。
- 同时，交换机将收到的数据帧的源MAC地址和对应端口编号记录到MAC地址表中。

- 注意：MAC地址表中动态学习的表项并非永远有效，每一条表项都有一个生存周期，到达生存周期仍得不到更新的表项将被删除，这个生存周期被称作老化时间。例如华为S系列交换机的老化时间缺省值是300秒。

交换机的MAC地址学习 (4)



4

- 交换机其他端口连接的主机，也会收到该数据帧，但是会丢弃。
- 主机2收到并处理该数据帧，向主机1回复，将数据帧发往交换机。

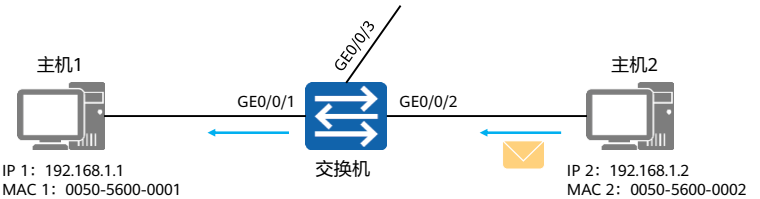
主机2发出的数据帧

源MAC：MAC2

目的MAC：MAC1

- 广播网络中的所有主机均会收到该数据帧，但是只有主机2会处理（因为目的MAC地址是主机2）。
- 主机2会回复数据帧给主机1，也是单播数据帧。

交换机的MAC地址学习 (5)



交换机查MAC地址表

MAC地址	Port
MAC 1	GE0/0/1
MAC 2	GE0/0/2

主机2发出的数据帧

源MAC：MAC2

目的MAC：MAC1

5

- 交换机在MAC地址表中查到了对应表项，则交换机对该单播帧执行转发操作，将数据帧从GE0/0/1口转发出去。
- 同时，交换机学习该数据帧的源MAC地址，并创建对应的MAC地址表项，与接收口GE0/0/2关联。

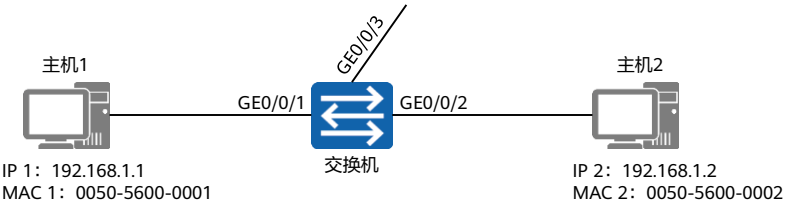
HUAWEI

- 交换机收到该单播数据帧后，会查看自己的MAC地址表，发现有对应的表项，则将数据从对应的端口转发出去。
- 同时，交换机将收到的数据帧的源MAC地址和对应端口编号记录到MAC地址表中。

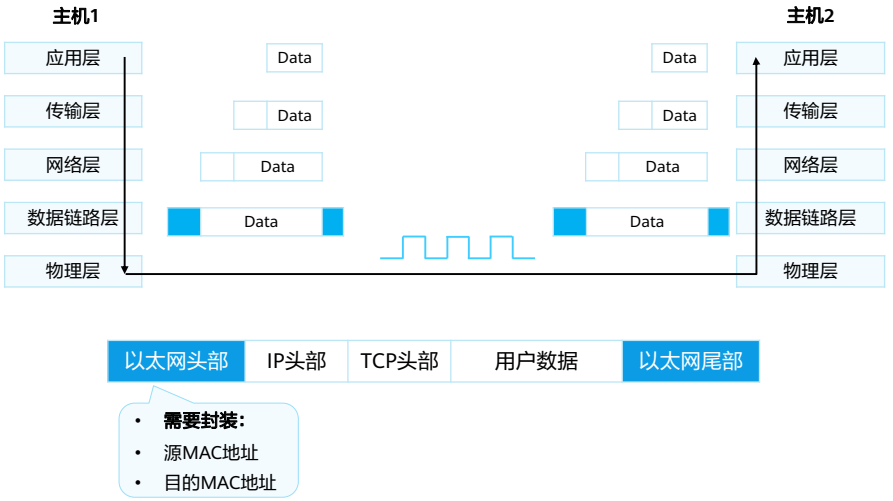
目录

同网段数据通信全过程

- 场景描述：
  - 任务：主机1想要访问主机2
  - 主机：初始化状态，仅知道本机IP地址和MAC地址（假设已获取对端IP地址）
  - 交换机：刚上电，初始化状态

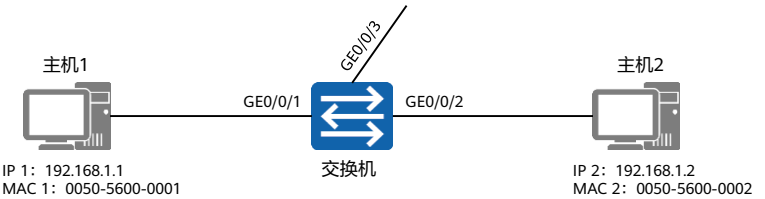


数据封装过程



- 主机1在发送数据报文前，需要先进行报文封装，包括源目IP地址、源目MAC地址等。

初始状态



主机1的ARP缓存表

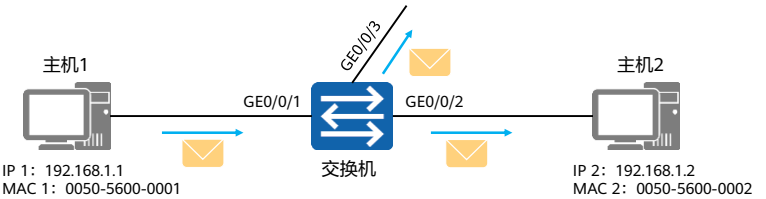
```
Host 1>arp -a
Internet Address    Physical Address    Type
```

交换机的MAC地址表

```
[Switch]display mac-address verbose
MAC address table of slot 0:
------------------------------------------------------------
MAC Address        Port        Type
------------------------------------------------------------


------------------------------------------------------------
```


- 主机1为了进行报文封装，会查本地的ARP缓存表。初始状态下，主机1的ARP缓存是空的。
- 而刚上电的交换机，初始状态下，交换机的MAC地址表也是空的。

泛洪数据帧



主机1发出的ARP Request

源MAC：MAC 1	目的MAC：FF-FF-FF-FF-FF-FF
源IP：IP 1	目的IP：IP 2
操作类型：ARP Request 发送端MAC： MAC 1 发送端IP：IP 1 目的端MAC：00-00-00-00-00-00 目的端IP：IP 2	

交换机的MAC地址表

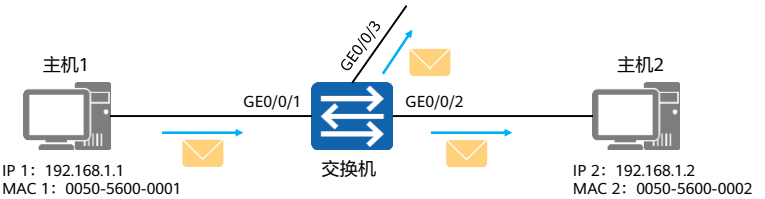
[Switch]display mac-address verbose
MAC address table of slot 0:

MAC Address Port Type

HUAWEI

- 主机1发送ARP请求报文，请求目的MAC地址。
- 交换机收到数据帧后，直接向所有非接收端口泛洪该数据帧。

学习MAC地址



主机1发出的ARP Request

源MAC：MAC 1	目的MAC：FF-FF-FF-FF-FF-FF
源IP：IP 1	目的IP：IP 2

操作类型：ARP Request
发送端MAC： MAC 1
发送端IP：IP 1
目的端MAC：00-00-00-00-00-00
目的端IP：IP 2

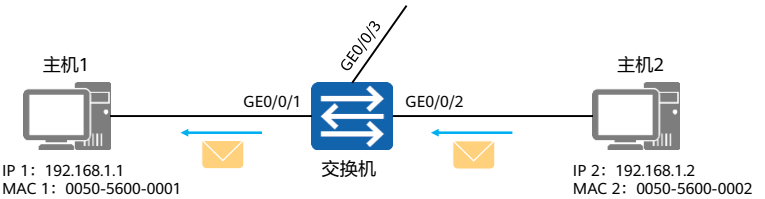
交换机的MAC地址表

```
[Switch]display mac-address verbose
MAC address table of slot 0:
-------------------------------------------------
MAC Address      Port        Type
-------------------------------------------------
0050-5600-0001   GE0/0/1     dynamic

-------------------------------------------------
```


HUAWEI

- 交换机将收到的数据帧的源MAC地址和对应端口记录到MAC地址表中。

目标主机回复



交换机的MAC地址表

[Switch]display mac-address verbose
MAC address table of slot 0:

MAC Address	Port	Type
0050-5600-0001	GE0/0/1	dynamic
0050-5600-0002	GE0/0/2	dynamic

主机2发出的ARP Reply

源MAC：MAC 2	目的MAC：MAC 1
源IP：IP 2	目的IP：IP 1

操作类型：ARP Reply
发送端MAC： MAC 2
发送端IP：IP 2
目的端MAC：MAC 1
目的端IP：IP 1

HUAWEI

- 主机2收到ARP请求报文后，会进行相应的处理，并发送ARP响应报文，回复主机1。
- 交换机收到的数据帧后查MAC地址表，发现有对应表项，则向对应端口转发该数据帧；并且交换机将收到的数据帧的源MAC地址和对应端口记录到MAC地址表中。
- 最终，主机1收到主机2的ARP响应报文后，就会将对应的IP地址和MAC地址记录到自己的ARP缓存中，并封装自己的报文，访问主机2。

思考题

1. 二层以太网交换机根据端口所接收到报文的（　）生成 MAC 地址表选项。

   A. 源 MAC 地址

   B. 目的 MAC 地址

   C. 源 IP 地址

   D. 目的 IP 地址

2. 一台交换机有8个端口，一个单播帧从某一端口进入了该交换机，但交换机在MAC地址表中查不到关于该帧的目的MAC地址表项，那么交换机对该帧进行的转发操作是？（　）

   A. 丢弃

   B. 泛洪

   C. 点对点转发

1. A
2. B

本章总结

- 在本章节中，介绍了以太网协议的基本概况，并介绍了以太网帧格式和MAC地址，还介绍了二层交换机的工作原理：在收到数据帧后，交换机学习帧的源MAC地址，然后在MAC地址表中查询该帧的目的MAC地址，并将帧从对应的端口转发出去。
- 最后，基于交换机的工作原理，回顾了同网段数据通信全过程。


谢 谢
www.huawei.com

VLAN原理与配置

前言

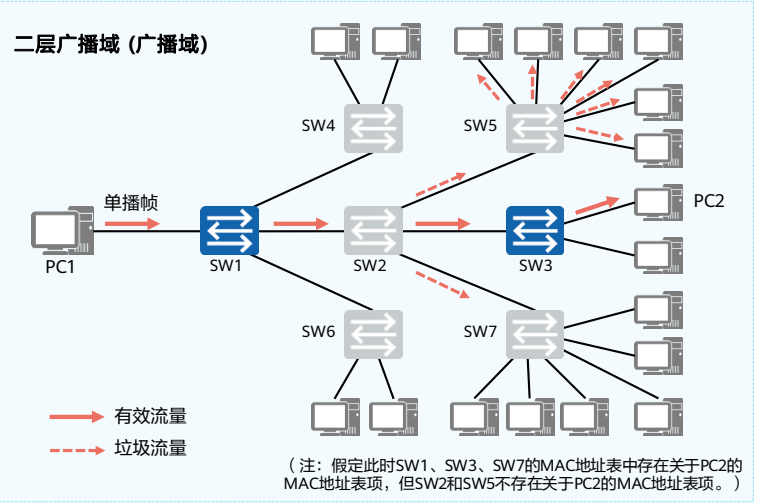
- 以太网是一种基于CSMA/CD的数据网络通信技术，其特征是共享通信介质。当主机数目较多时会导致安全隐患、广播泛滥、性能显著下降甚至造成网络不可用。
- 在这种情况下出现了VLAN (Virtual Local Area Network)技术解决以上问题。
- 在本课程中，将介绍VLAN技术的相关概念，介绍不同二层接口的工作原理，并且会介绍VLAN的应用及其数据转发原理和相关配置。

目标

- 学完本课程后，您将能够：
  - 了解VLAN技术的产生背景
  - 识别数据所属的VLAN
  - 掌握不同的VLAN划分方式
  - 描述网络中VLAN数据的通信过程
  - 掌握VLAN的基本配置

目录

传统以太网的问题

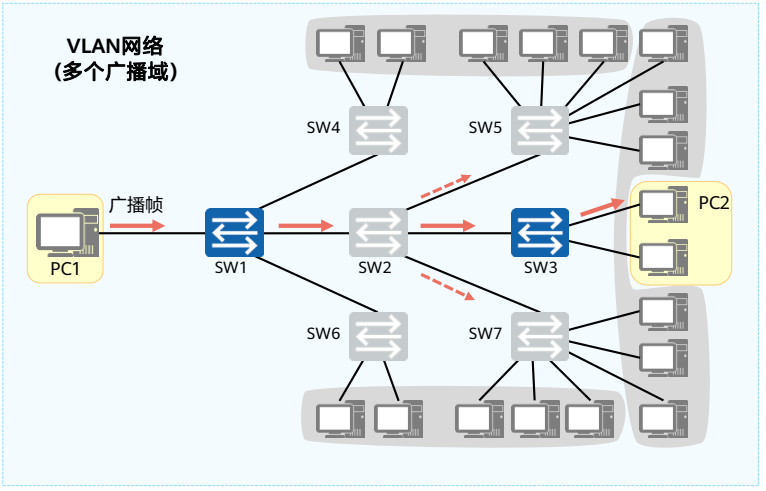


- 在典型交换网络中，当某台主机发送一个广播帧或未知单播帧时，该数据帧会被泛洪，甚至传递到整个广播域。
- 广播域越大，产生的网络安全问题、垃圾流量问题，就越严重。

HUAWEI

- 广播域：
  - 如图是一个典型的交换网络，网络中只有终端计算机和交换机。在这样的网络中，如果某一台计算机发送了一个广播帧，由于交换机对广播帧执行泛洪操作，结果所有其他的计算机都会收到这个广播帧。
  - 把广播帧所能到达的整个访问范围称为二层广播域，简称广播域 (Broadcast Domain)。显然，一个交换网络其实就是一个广播域。
- 网络安全问题和垃圾流量问题：
  - 如图：如果PC1向PC2发送了一个单播帧。此时SW1、SW3、SW7的MAC地址表中存在关于PC2的MAC地址表项，但SW2和SW5不存在关于PC2的MAC地址表项。那么，SW1和SW3将对该单播帧执行点对点的转发操作，SW7将对该单播帧执行丢弃操作，SW2和SW5将对该单播帧执行泛洪操作。最后的结果是，PC2虽然收到了该单播帧，但网络中的很多其他非目的主机，同样收到了不该接收的数据帧。
- 显然，广播域越大，网络安全问题和垃圾流量问题就越严重。

虚拟局域网 (VLAN, Virtual LAN)



- 虚拟局域网VLAN可以隔离广播域。
- 特点：
  - 不受地域限制。
  - 同一VLAN内的设备才能直接进行二层通信。

HUAWEI

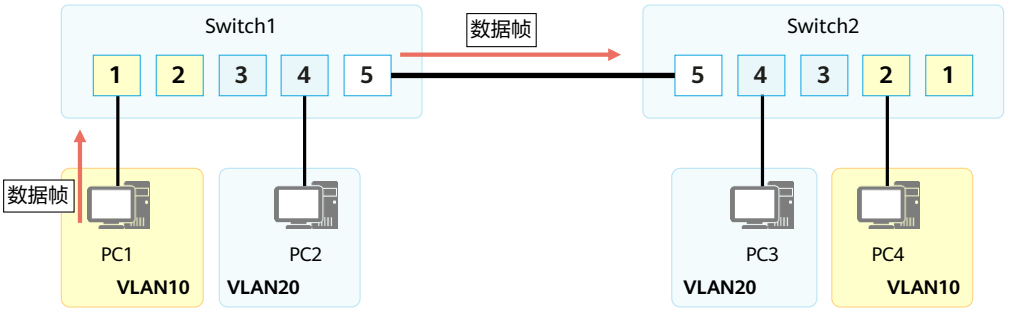
- 为了解决广播域带来的问题，人们引入了VLAN (Virtual Local Area Network)，即虚拟局域网技术：
  - 通过在交换机上部署VLAN，可以将一个规模较大的广播域在逻辑上划分成若干个不同的、规模较小的广播域，由此可以有效地提升网络的安全性，同时减少垃圾流量，节约网络资源。
- VLAN的特点：
  - 一个VLAN就是一个广播域，所以在同一个VLAN内部，计算机可以直接进行二层通信；而不同VLAN内的计算机，无法直接进行二层通信，只能进行三层通信来传递信息，即广播报文被限制在一个VLAN内。
  - VLAN的划分不受地域的限制。
- VLAN的好处：
  - 灵活构建虚拟工作组：用VLAN可以划分不同的用户到不同的工作组，同一工作组的用户也不必局限于某一固定的物理范围，网络构建和维护更方便灵活。
  - 限制广播域：广播域被限制在一个VLAN内，节省了带宽，提高了网络处理能力。
  - 增强局域网的安全性：不同VLAN内的报文在传输时是相互隔离的，即一个VLAN内的用户不能和其它VLAN内的用户直接通信。
  - 提高了网络的健壮性：故障被限制在一个VLAN内，本VLAN内的故障不会影响其他VLAN的正常工作。
- 注：二层，即数据链路层。

目录

  HUAWEI

- 此部分内容，将从：如何识别VLAN，网络中VLAN的划分方式，以及交换机如何进行VLAN的数据交互三部分来讲解VLAN的基本原理。

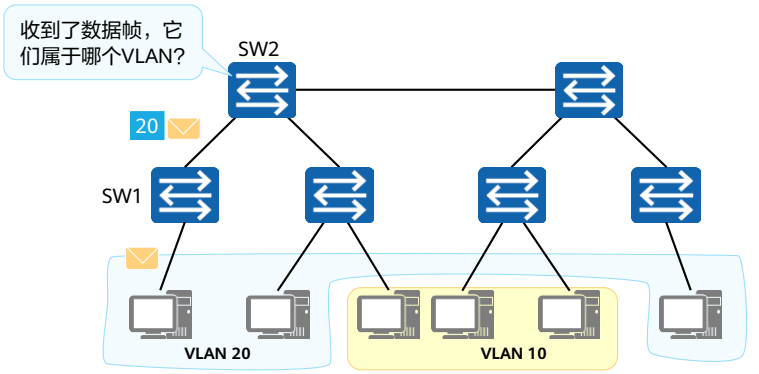
如何实现VLAN



- Switch1与Switch2同属一个企业，该企业统一规划了网络中的VLAN。其中VLAN10用于A部门，VLAN20用于B部门。A、B部门的员工在Switch1和Switch2上都有接入。
- PC1发出的数据经过Switch1和Switch2之间的链路到达了Switch2。如果不加处理，后者无法判断该数据所属的VLAN，也不知道应该将这个数据输出到本地哪个VLAN中。

VLAN标签 (VLAN Tag)

- 交换机如何识别接收到的数据帧属于哪个VLAN?



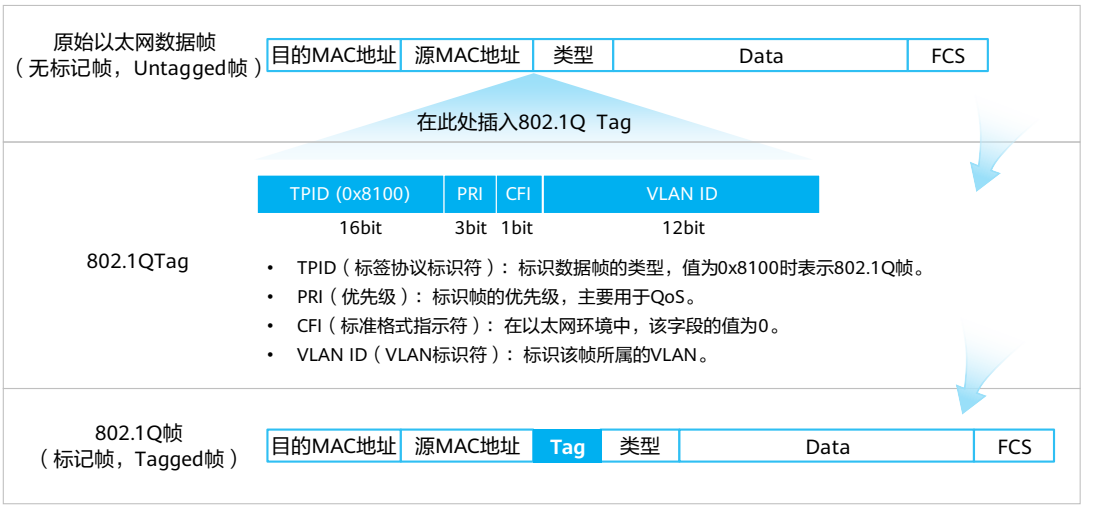
VLAN标签

- 要使交换机能够分辨不同VLAN的报文，需要在报文中添加标识VLAN信息的字段。
- IEEE 802.1Q协议规定，在以太网数据帧中加入4个字节的VLAN标签，又称VLAN Tag，简称Tag。

- 如图所示，SW1识别出某个帧是属于哪个VLAN后，会在这个帧的特定位置上添加一个标签。这个标签明确地标明了这个帧是属于哪个VLAN的。其他交换机（如SW2）收到这个带标签的数据帧后，就能轻而易举地直接根据标签信息识别出这个帧属于哪个VLAN。

- IEEE 802.1Q定义了这种带标签的数据帧的格式。满足这种格式的数据帧称为IEEE 802.1Q数据帧，也称VLAN数据帧。

VLAN数据帧

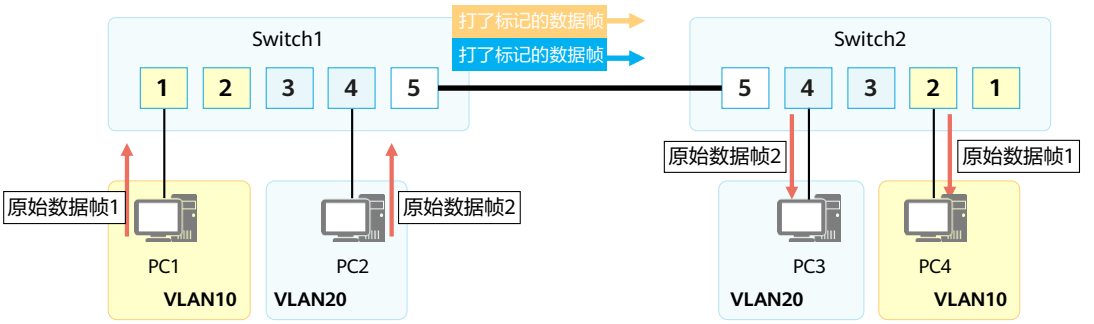


HUAWEI

- 在一个VLAN交换网络中，以太网帧主要有以下两种形式：
  - 有标记帧（Tagged帧）：IEEE 802.1Q协议规定，在以太网数据帧的目的MAC地址和源MAC地址字段之后、协议类型字段之前加入4个字节的VLAN标签（又称VLAN Tag，简称Tag）的数据帧。
  - 无标记帧（Untagged帧）：原始的、未加入4字节VLAN标签的数据帧。
- VLAN数据帧中的主要字段：
  - TPID：2字节，Tag Protocol Identifier（标签协议标识符），表示数据帧类型。
    - 取值为0x8100时表示IEEE 802.1Q的VLAN数据帧。如果不支持802.1Q的设备收到这样的帧，会将其丢弃。
    - 各设备厂商可以自定义该字段的值。当邻居设备将TPID值配置为非0x8100时，为了能够识别这样的报文，实现互通，必须在本设备上修改TPID值，确保和邻居设备的TPID值配置一致。
  - PRI：3 bit，Priority，表示数据帧的优先级，用于QoS。
    - 取值范围为0～7，值越大优先级越高。当网络阻塞时，交换机优先发送优先级高的数据帧。

    - CFI：1 bit，Canonical Format Indicator（标准格式指示位），表示MAC地址在不同的传输介质中是否以标准格式进行封装，用于兼容以太网和令牌环网
        - CFI取值为0表示MAC地址以标准格式进行封装，为1表示以非标准格式封装。
        - 在以太网中，CFI的值为0。
    - VID：12 bit，VLAN ID，表示该数据帧所属VLAN的编号。
        - VLAN ID取值范围是0～4095。由于0和4095为协议保留取值，所以VLAN ID的有效取值范围是1～4094。
        - 交换机利用VLAN标签中的VID来识别数据帧所属的VLAN，广播帧只在同一VLAN内转发，这就将广播域限制在一个VLAN内。
- 如何识别带VLAN标签的数据帧：
    - 数据帧的Length/Type = 0x8100。
- 注意：计算机无法识别Tagged数据帧，因此计算机处理和发出的都是Untagged数据帧；为了提高处理效率，交换机内部处理的数据帧一律都是Tagged帧。

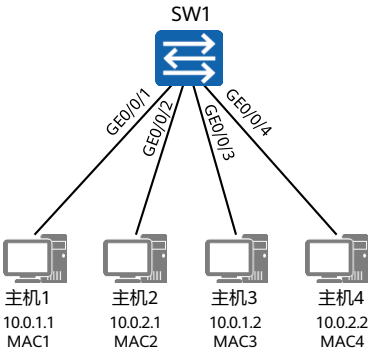
VLAN的实现



- Switch1和Switch2之间的链路要承载多个VLAN的数据，需要一种基于VLAN的数据“标记”手段，以便对不同VLAN的数据帧进行区分。
- IEEE 802.1Q标准（也被称为Dot1Q）定义了该“标记”方法。该标准对传统的以太网数据帧进行修改，在帧头中插入802.1Q Tag，而在该Tag中，便可以写入VLAN信息。

VLAN的划分方式

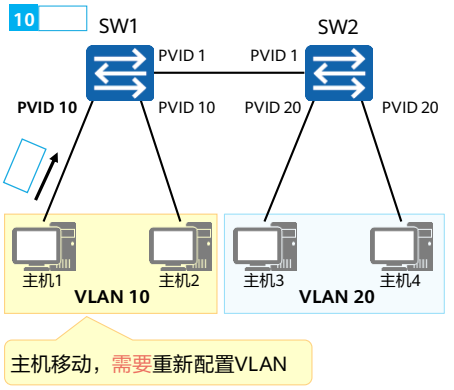
整个网络是如何划分VLAN的?



VLAN划分方式	VLAN 10	VLAN 20
基于接口	GE0/0/1，GE0/0/3	GE0/0/2，GE0/0/4
基于MAC地址	MAC 1，MAC 3	MAC 2，MAC 4
基于IP子网划分	10.0.1.*	10.0.2.*
基于协议划分	IP	IPv6
基于策略	10.0.1.* + GE0/0/1+ MAC 1	10.0.2.* + GE0/0/2 + MAC 2

- 计算机发出的数据帧不带任何标签。对已支持VLAN特性的交换机来说，当计算机发出的Untagged帧一旦进入交换机后，交换机必须通过某种划分原则把这个帧划分到某个特定的VLAN中去。
- VLAN的划分包括如下5种方法：
  - 基于接口划分：根据交换机的接口来划分VLAN。
    - 网络管理员预先给交换机的每个接口配置不同的PVID，当一个数据帧进入交换机时，如果没有带VLAN标签，该数据帧就会被打上接口指定PVID的标签，然后数据帧将在指定VLAN中传输。
  - 基于MAC地址划分：根据数据帧的源MAC地址来划分VLAN。
    - 网络管理员预先配置MAC地址和VLAN ID映射关系表，当交换机收到的是Untagged帧时，就依据该表给数据帧添加指定VLAN的标签，然后数据帧将在指定VLAN中传输。
  - 基于IP子网划分：根据数据帧中的源IP地址和子网掩码来划分VLAN。
    - 网络管理员预先配置IP地址和VLAN ID映射关系表，当交换机收到的是Untagged帧，就依据该表给数据帧添加指定VLAN的标签，然后数据帧将在指定VLAN中传输。
  - 基于协议划分：根据数据帧所属的协议（族）类型及封装格式来划分VLAN。
    - 网络管理员预先配置以太网帧中的协议域和VLAN ID的映射关系表，如果收到的是Untagged帧，就依据该表给数据帧添加指定VLAN的标签，然后数据帧将在指定VLAN中传输。
  - 基于策略划分：根据配置的策略划分VLAN，能实现多种组合的划分方式，包括接口、MAC地址、IP地址等。
    - 网络管理员预先配置策略，如果收到的是Untagged帧，且匹配配置的策略时，给数据帧添加指定VLAN的标签，然后数据帧将在指定VLAN中传输。

基于接口的VLAN划分



基于接口的VLAN划分

- **原理**
  - 根据交换机的接口来划分VLAN。
  - 网络管理员预先给交换机的每个接口配置不同的**PVID**，将该接口划入PVID对应的VLAN。
  - 当一个数据帧进入交换机时，如果没有带VLAN标签，该数据帧就会被打上接口指定PVID的Tag，然后数据帧将在指定PVID中传输。
- **缺省VLAN，PVID**
  - Port VLAN ID，是接口上的缺省VLAN。
  - 取值：1~4094。

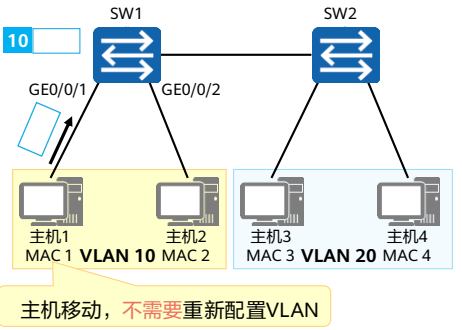
HUAWEI

- 划分原则：
  - 将VLAN ID配置到交换机的物理接口上，从某一个物理接口进入交换机的、由终端计算机发送的Untagged数据帧都被划分到该接口的VLAN ID所表明的那个VLAN。
- 特点：
  - 这种划分原则简单而直观，实现容易，是目前实际的网络应用中最为广泛的划分VLAN的方式。
  - 当计算机接入交换机的端口发生了变化时，该计算机发送的帧的VLAN归属可能会发生变化。
- 缺省VLAN，PVID (Port VLAN ID)
  - 每个交换机的接口都应该配置一个PVID，到达这个端口的Untagged帧将一律被交换机划分到PVID所指代的VLAN。
  - 默认情况下，PVID的值为1。

基于MAC地址的VLAN划分

SW1的MAC地址与VLAN表

MAC地址	VLAN ID
MAC 1	10
MAC 2	10
.....	.....

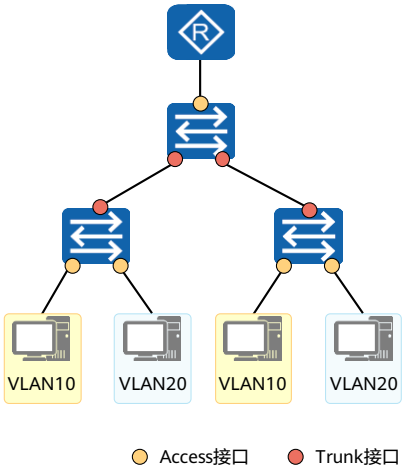


基于MAC地址的VLAN划分

- **原理**
  - 根据数据帧的源MAC地址来划分VLAN。
  - 网络管理员预先配置MAC地址和VLAN ID映射关系表。
  - 当交换机收到的是Untagged帧时，就依据该表给数据帧添加指定VLAN的Tag，然后数据帧将在指定VLAN中传输。
- **映射表**
  - 记录了MAC地址和VLAN ID的关联情况。

- 划分原则：
  - 交换机内部建立并维护了一个MAC地址与VLAN ID的对应表。当交换机接收到计算机发送的Untagged帧时，交换机将分析帧中的源MAC地址，然后查询MAC地址与VLAN ID的对应表，并根据对应关系把这个帧划分到相应的VLAN中。
- 特点：
  - 这种划分实现稍微复杂，但灵活性得到了提高。
  - 当计算机接入交换机的端口发生了变化时，该计算机发送的帧的VLAN归属不会发生变化（因为计算机的MAC地址没有变）。
  - 但这种类型的VLAN划分安全性不是很高，因为恶意计算机很容易伪造MAC地址。

以太网二层接口类型



接口类型

- **Access接口**

  交换机上常用来连接用户PC、服务器等终端设备的接口。Access接口所连接的这些设备的网卡往往只收发无标记帧。Access接口只能加入一个VLAN。

- **Trunk接口**

  Trunk接口允许多个VLAN的数据帧通过，这些数据帧通过802.1Q Tag实现区分。Trunk接口常用于交换机之间的互联，也用于连接路由器、防火墙等设备的子接口。

- **Hybrid接口**

  Hybrid接口与Trunk接口类似，也允许多个VLAN的数据帧通过，这些数据帧通过802.1Q Tag实现区分。用户可以灵活指定Hybrid接口在发送某个（或某些）VLAN的数据帧时是否携带Tag。

- 基于接口的VLAN划分依赖于交换机的接口类型。
- Access接口

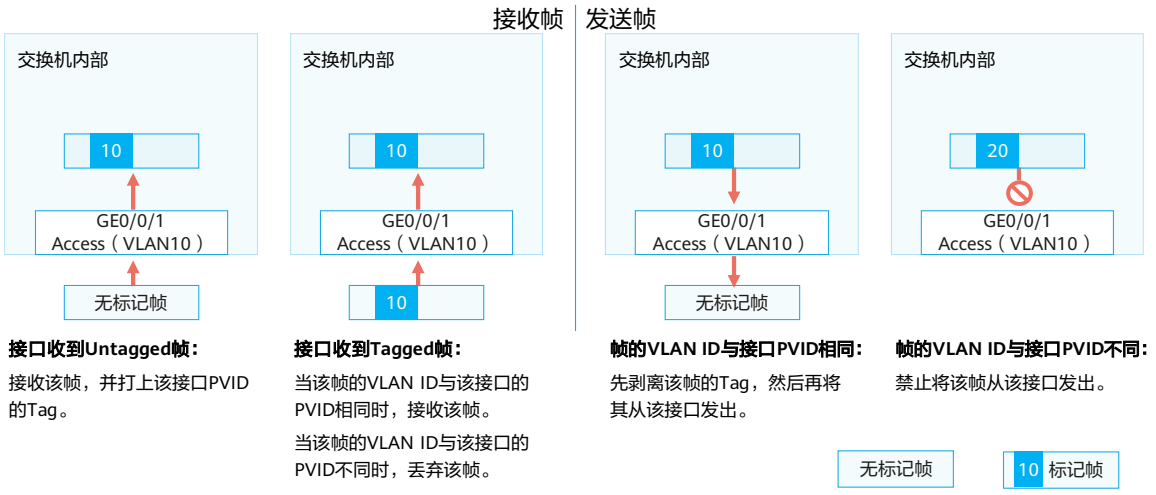
  Access接口一般用于和不能识别Tag的用户终端（如用户主机、服务器等）相连，或者不需要区分不同VLAN成员时使用。

- Trunk接口

  Trunk接口一般用于连接交换机、路由器、AP以及可同时收发Tagged帧和Untagged帧的语音终端。

- Hybrid接口
  - Hybrid接口既可以用于连接不能识别Tag的用户终端（如用户主机、服务器等），也可以用于连接交换机、路由器以及可同时收发Tagged帧和Untagged帧的语音终端、AP。
  - 华为设备默认的接口类型是Hybrid。

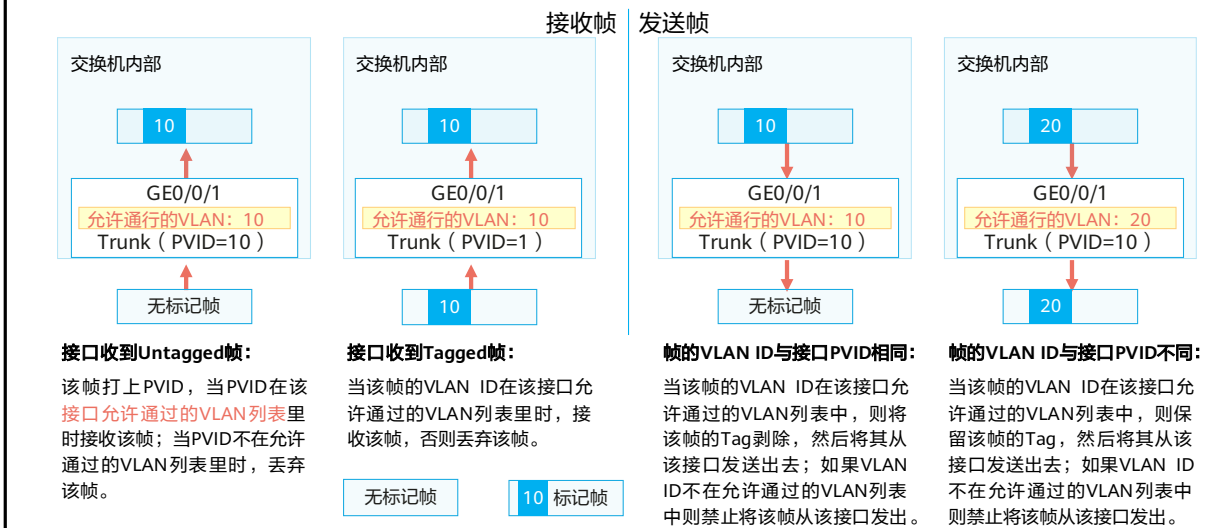
Access接口



  HUAWEI

- 上文已经介绍了交换机如何识别数据帧属于哪个VLAN以及VLAN的划分方式，那交换机对于Untagged帧和Tagged帧又是如何处理的呢?
- Access接口特点：
  - 仅允许VLAN ID与接口PVID相同的数据帧通过。
- Access接口接收数据帧：
  - 当Access接口从链路上收到一个Untagged帧，交换机会在这个帧中添加上VID为PVID的Tag，然后对得到的Tagged帧进行转发操作（泛洪、转发、丢弃）。
  - 当Access接口从链路上收到一个Tagged帧，交换机会检查这个帧的Tag中的VID是否与PVID相同。如果相同，则对这个Tagged帧进行转发操作；如果不同，则直接丢弃这个Tagged帧。
- Access接口发送数据帧：
  - 当一个Tagged帧从本交换机的其他接口到达一个Access接口后，交换机会检查这个帧的Tag中的VID是否与PVID相同：
    - 如果相同，则将这个Tagged帧的Tag进行剥离，然后将得到的Untagged帧从链路上发送出去；
    - 如果不同，则直接丢弃这个Tagged帧。

Trunk接口



接口收到Untagged帧：

该帧打上PVID，当PVID在该接口允许通过的VLAN列表里时接收该帧；当PVID不在允许通过的VLAN列表里时，丢弃该帧。

接口收到Tagged帧：

当该帧的VLAN ID在该接口允许通过的VLAN列表里时，接收该帧，否则丢弃该帧。

无标记帧　10 标记帧

帧的VLAN ID与接口PVID相同：

当该帧的VLAN ID在该接口允许通过的VLAN列表中，则将该帧的Tag剥除，然后将其从该接口发送出去；如果VLAN ID不在允许通过的VLAN列表中则禁止将该帧从该接口发出。

帧的VLAN ID与接口PVID不同：

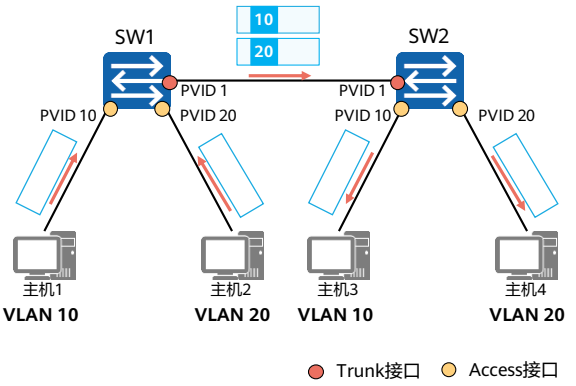
当该帧的VLAN ID在该接口允许通过的VLAN列表中，则保留该帧的Tag，然后将其从该接口发送出去；如果VLAN ID不在允许通过的VLAN列表中则禁止将该帧从该接口发出。

HUAWEI

- 对于Trunk接口，除了要配置PVID外，还必须配置允许通过的VLAN ID列表，其中VLAN 1是默认存在的。
- Trunk接口特点：
  - Trunk接口仅允许VLAN ID在允许通过列表中的数据帧通过。
  - Trunk接口可以允许多个VLAN的帧带Tag通过，但只允许一个VLAN的帧从该类接口上发出时不带Tag（即剥除Tag）。
- Trunk接口接收数据帧：
  - 当Trunk接口从链路上收到一个Untagged帧，交换机会在这个帧中添加上VID为PVID的Tag，然后查看PVID是否在允许通过的VLAN ID列表中。如果在，则对得到的Tagged帧进行转发操作；如果不在，则直接丢弃得到的Tagged帧。
  - 当Trunk接口从链路上收到一个Tagged帧，交换机会检查这个帧的Tag中的VID是否在允许通过的VLAN ID列表中。如果在，则对这个Tagged帧进行转发操作；如果不在，则直接丢弃这个Tagged帧。
- Trunk接口发送数据帧：
  - 当一个Tagged帧从本交换机的其他接口到达一个Trunk接口后，如果这个帧的Tag中的VID不在允许通过的VLAN ID列表中，则该Tagged帧会被直接丢弃。
  - 当一个Tagged帧从本交换机的其他接口到达一个Trunk接口后，如果这个帧的Tag中的VID在允许通过的VLAN ID列表中，则会比较该Tag中的VID是否与接口的PVID相同：
    - 如果相同，则交换机会对这个Tagged帧的Tag进行剥离，然后将得到的Untagged帧从链路上发送出去；
    - 如果不同，则交换机不会对这个Tagged帧的Tag进行剥离，而是直接将它从链路上发送出去。

Access接口与Trunk接口举例

- 请描述主机之间数据访问的全流程。

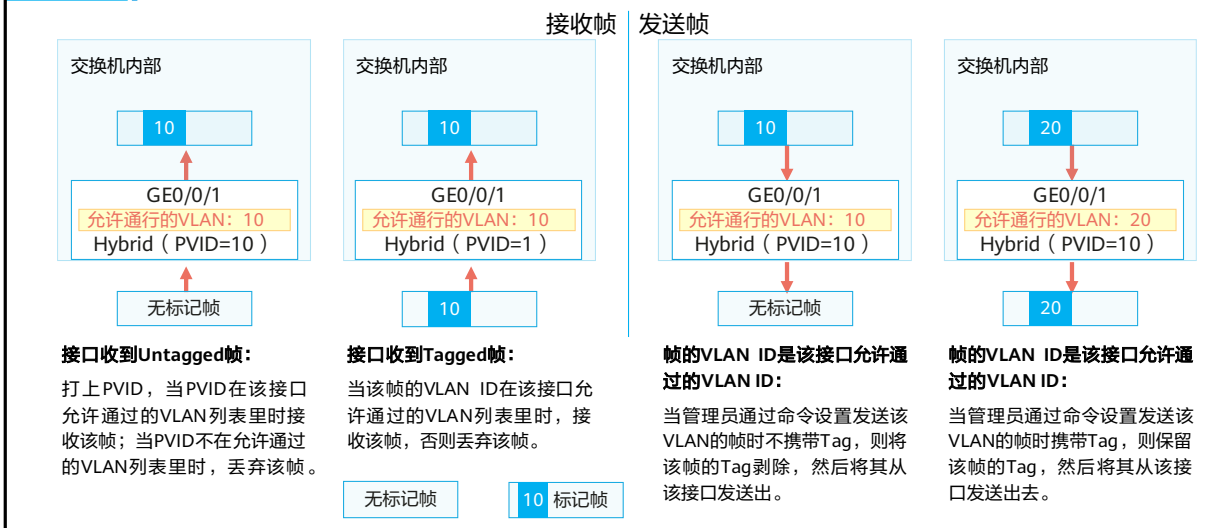


SW1与SW2的Trunk接口

允许通过列表	
VLAN ID	1
	10
	20

- 在本例中，SW1和SW2连接主机的接口为Access接口，PVID如图所示。SW1和SW2互连的接口为Trunk接口，PVID都为1，此Trunk接口的允许通过的VLAN ID列表也如图所示。
- 请描述主机之间数据互访的全流程。

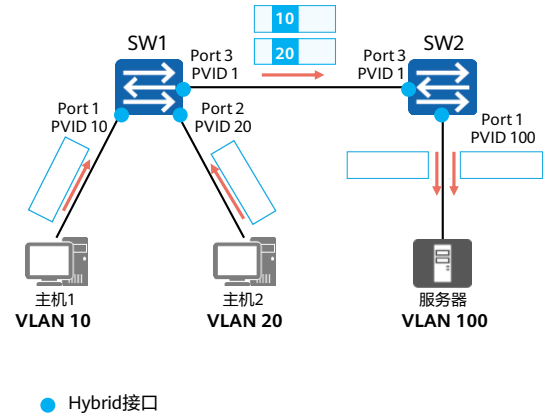
Hybrid接口



- 对于Hybrid接口，除了要配置PVID外，还存在两个允许通过的VLAN ID列表，一个是Untagged VLAN ID列表，另一个是Tagged VLAN ID列表，其中VLAN 1默认在Untagged VLAN列表中。这两个允许通过列表中的所有VLAN的帧都是允许通过这个Hybrid接口的。
- Hybrid接口特点：
  - Hybrid接口仅允许VLAN ID在允许通过列表中的数据帧通过。
  - Hybrid接口可以允许多个VLAN的帧带Tag通过，且允许从该类接口发出的帧根据需要配置某些VLAN的帧带Tag、某些VLAN的帧不带Tag。
  - 与Trunk最主要的区别就是，能够支持多个VLAN的数据帧，不带标签通过。
- Hybrid接口接收数据帧：
  - 当Hybrid接口从链路上收到一个Untagged帧，交换机会在这个帧中添加上VID为PVID的Tag，然后查看PVID是否在Untagged或Tagged VLAN ID列表中。如果在，则对得到的Tagged帧进行转发操作；如果不在，则直接丢弃得到的Tagged帧。
  - 当Hybrid接口从链路上收到一个Tagged帧，交换机会检查这个帧的Tag中的VID是否在Untagged或Tagged VLAN ID列表中。如果在，则对这个Tagged帧进行转发操作；如果不在，则直接丢弃这个Tagged帧。
- Hybrid接口发送数据帧：
  - 当一个Tagged帧从本交换机的其他接口到达一个Hybrid接口后，如果这个帧的Tag中的VID既不在Untagged VLAN ID列表中，也不在Tagged VLAN ID列表中，则该Tagged帧会被直接丢弃。
  - 当一个Tagged帧从本交换机的其他接口到达一个Hybrid接口后，如果这个帧的Tag中的VID在Untagged VLAN ID列表中，则交换机会对这个Tagged帧的Tag进行剥离，然后将得到的Untagged帧从链路上发送出去。
  - 当一个Tagged帧从本交换机的其他接口到达一个Hybrid接口后，如果这个帧的Tag中的VID在Tagged VLAN ID列表中，则交换机不会对这个Tagged帧的Tag进行剥离，而是直接将它从链路上发送出去。

Hybrid接口举例

- 请描述主机访问服务器的全流程。



交换机1的允许通过列表

	Port1	Port2	Port3
	Untagged	Untagged	Tagged
VLAN ID	1	1	10
	10	20	20
	100	100	100

交换机2的允许通过列表

	Port1	Port3
	Untagged	Tagged
VLAN ID	1	10
	10	20
	20	100
	100	

- 在本例中，SW1和SW2连接主机的接口以及互连的接口均为Hybrid接口，PVID如图所示，Hybrid接口的允许通过的VLAN ID列表也如图所示。
- 请描述两个主机互访服务器的全流程。

小结

Access接口	Trunk接口	Hybrid接口
接收数据帧 ▫ Untagged数据帧，打上PVID，接收。 ▫ Tagged数据帧，与PVID比较，相同则接收；不同则丢弃。	接收数据帧 ▫ Untagged数据帧，打上PVID，且VID在允许列表中，则接收；VID不在允许列表，则丢弃。 ▫ Tagged数据帧，查看VID是否在允许列表中，在允许列表中，则接收；VID不在允许列表，则丢弃。	接收数据帧 ▫ Untagged数据帧，打上PVID，且VID在允许列表中，则接收；VID不在允许列表中，则丢弃。 ▫ Tagged数据帧，查看VID是否在允许列表中，在允许列表中，则接收；VID不在允许列表，则丢弃。
发送数据帧 ▫ VID与PVID比较，相同则剥离标签发送；不同则丢弃。	发送数据帧 ▫ VID在允许列表中，且VID与PVID一致，则剥离标签发送。 ▫ VID在允许列表，但VID与PVID不一致，则直接带标签发送。 ▫ 不在允许列表中，则直接丢弃。	发送数据帧 ▫ VID不在允许列表中，直接丢弃。 ▫ VID在Untagged列表中，剥离标签发送。 ▫ VID在Tagged列表中，带标签直接发送。

- 各类接口添加或剥除VLAN标签的处理过程总结如下：
  - 当接收数据帧时：
    - 当接收到不带VLAN标签的数据帧时，Access接口、Trunk接口、Hybrid接口都会给数据帧打上VLAN标签，但Trunk接口、Hybrid接口会根据数据帧的VID是否为其允许通过的VLAN来判断是否接收，而Access接口则无条件接收。
    - 当接收到带VLAN标签的数据帧时，Access接口、Trunk接口、Hybrid接口都会根据数据帧的VID是否为其允许通过的VLAN（Access接口允许通过的VLAN就是缺省VLAN）来判断是否接收。
  - 当发送数据帧时：
    - Access接口直接剥离数据帧中的VLAN标签。
    - Trunk接口只有在数据帧中的VID与接口的PVID相等时才会剥离数据帧中的VLAN标签。
    - Hybrid接口会根据接口上的配置判断是否剥离数据帧中的VLAN标签。
- 因此，Access接口发出的数据帧肯定不带Tag；Trunk接口发出的数据帧只有一个VLAN的数据帧不带Tag，其他都带VLAN标签；Hybrid接口发出的数据帧可根据需要设置某些VLAN的数据帧带Tag，某些VLAN的数据帧不带Tag。

目录

VLAN的规划

- **VLAN分配原则**
  - 按业务规划：可分为语音、视频和数据。
  - 按部门规划：可分为工程部、市场部、财经部等。
  - 按应用规划：可分为服务器、办公、教室等。
- **VLAN分配技巧**

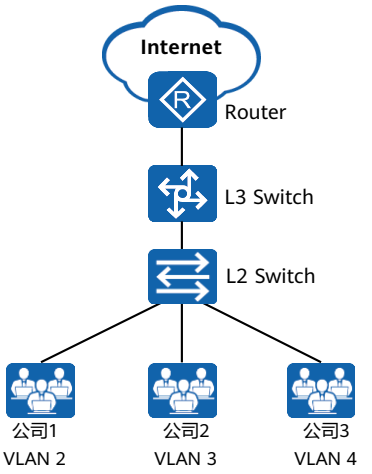
  VLAN ID的分配在有效范围内，可以随意分配和选取，但是为了提高VLAN ID的连续性，可以采用VLAN ID和子网关联的方式进行分配。
- **VLAN规划示例**
  - 假设某园区有三栋楼，分别为行政楼、教学楼、办公楼；每栋楼各有1台接入交换机，核心交换机在行政楼；行政楼内有办公室、财务部和教室；办公楼内有办公室和财务部；教学楼内有办公室和教室。
  - VLAN规划如下：

VLAN	IP地址段	描述
1	X.16.10.0/24	办公室用户所属的VLAN
2	X.16.20.0/24	财务部用户所属的VLAN
3	X.16.30.0/24	教室用户所属的VLAN
100	Y.16.100.0/24	设备管理所属的VLAN

- VLAN编号建议连续分配，以保证VLAN资源合理利用。最常用的划分方式是基于接口的方式。

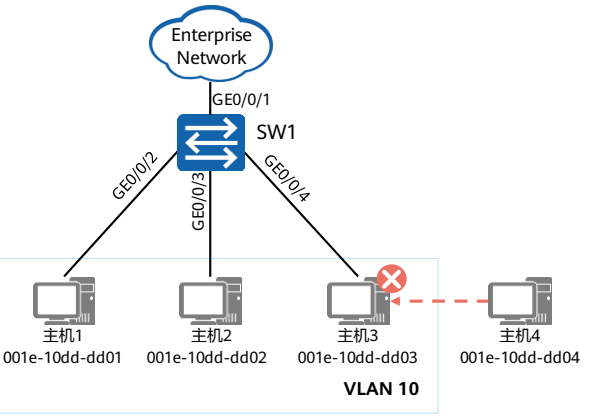
应用场景 - 基于接口的VLAN划分

- 应用场景：
  - 某商务楼内有多家公司，为了降低成本，多家公司共用网络资源，各公司分别连接到一台二层交换机的不同接口，并通过统一的出口访问Internet。
- VLAN划分：
  - 为了保证各公司业务的独立和安全，可将每个公司所连接的接口划分到不同的VLAN，实现公司间业务数据的完全隔离。可以认为每个公司拥有独立的网络，每个VLAN就是一个“虚拟工作组”。



应用场景 - 基于MAC的VLAN划分

- 应用场景：
  - 某个公司的网络中，网络管理者将同一部门的员工划分到同一VLAN。为了提高部门内的信息安全，要求只有本部门员工的主机才可以访问特定网络资源。
- VLAN划分：
  - 为了保证非本部门员工不能访问网络资源，可在SW1上配置基于MAC地址划分VLAN。这样，新的主机接入网络，就无法访问公司的网络资源。



目录

VLAN的基础配置命令

1. 创建VLAN

[Huawei] **vlan** *vlan-id*

通过此命令创建VLAN并进入VLAN视图，如果VLAN已存在，直接进入该VLAN的视图。

- *vlan-id*是整数形式，取值范围是1～4094。

[Huawei] **vlan batch** { *vlan-id1* [**to** *vlan-id2*] }

通过此命令批量创建VLAN。其中：

- batch：指定批量创建的VLAN ID。
- *vlan-id1*：表示第一个VLAN的编号。
- *vlan-id2*：表示最后一个VLAN的编号。

- **vlan**命令用来创建VLAN并进入VLAN视图，如果VLAN已存在，直接进入该VLAN的视图。
- **undo vlan**用来删除指定VLAN。
- 缺省情况下，将所有接口都加入到一个缺省的VLAN中，该VLAN标识为1。
  - 命令：
  - **vlan** *vlan-id*
    - *vlan-id*：指定VLAN ID。整数形式，取值范围是1～4094。
  - **vlan batch** { *vlan-id1* [**to** v*lan-id2*] }
    - batch：指定批量创建VLAN。
    - *vlan-id1* to *vlan-id2*：指定批量创建的VLAN ID，其中：
      - *vlan-id1*表示第一个VLAN的编号。
      - *vlan-id2*表示最后一个VLAN的编号。*vlan-id2*的取值必须大于等于*vlan-id1*，它与*vlan-id1*共同确定一个VLAN范围。
    - 如果不指定to *vlan-id2*参数，则只创建*vlan-id1*所指定的VLAN。
    - *vlan-id1*和*vlan-id2*是整数形式，取值范围是1～4094。

Access接口的基础配置命令

1. 配置接口类型

```
[Huawei-GigabitEthernet0/0/1] port link-type access
```

在接口视图下，配置接口的链路类型为Access。

2. 配置Access接口的缺省VLAN

```
[Huawei-GigabitEthernet0/0/1] port default vlan vlan-id
```

在接口视图下，配置接口的缺省VLAN并同时加入这个VLAN。

- *vlan-id*：配置缺省VLAN的编号。整数形式，取值范围是1～4094。

Trunk接口的基础配置命令

1. 配置接口类型

```
[Huawei-GigabitEthernet0/0/1] port link-type trunk
```

在接口视图下，配置接口的链路类型为Trunk。

2. 配置Trunk接口加入指定VLAN

```
[Huawei-GigabitEthernet0/0/1] port trunk allow-pass vlan { { vlan-id1 [ to vlan-id2 ] } | all }
```

在接口视图下，配置Trunk类型接口加入的VLAN。

3. (可选) 配置Trunk接口的缺省VLAN

```
[Huawei-GigabitEthernet0/0/1] port trunk pvid vlan vlan-id
```

在接口视图下，配置Trunk类型接口的缺省VLAN。

- 命令：**port trunk allow-pass vlan** { { *vlan-id1* [**to** *vlan-id2*] | **all** }
  - *vlan-id1* [to *vlan-id2*]：指定Trunk类型接口加入的VLAN，其中：
    - *vlan-id1*表示第一个VLAN的编号。
    - to *vlan-id2*表示最后一个VLAN的编号。*vlan-id2*的取值必须大于等于*vlan-id1*的取值。
    - *vlan-id1*和*vlan-id 2*为整数形式，取值范围是1～4094。
  - all：指定Trunk接口加入所有VLAN。
- 命令：**port trunk pvid vlan** *vlan-id*，设置Trunk类型接口的缺省VLAN。
  - *vlan-id*：指定Trunk类型接口的缺省VLAN编号。整数形式，取值范围是1～4094。

Hybrid接口的基础配置命令

1. 配置接口类型

```
[Huawei-GigabitEthernet0/0/1] port link-type hybrid
```

在接口视图下，配置接口的链路类型为Hybrid。

2. 配置Hybrid接口加入指定VLAN

```
[Huawei-GigabitEthernet0/0/1] port hybrid untagged vlan { { vlan-id1 [ to vlan-id2 ] } | all }
```

在接口视图下，配置Hybrid类型接口加入的VLAN，这些VLAN的帧以Untagged方式通过接口。

```
[Huawei-GigabitEthernet0/0/1] port hybrid tagged vlan { { vlan-id1 [ to vlan-id2 ] } | all }
```

在接口视图下，配置Hybrid类型接口加入的VLAN，这些VLAN的帧以Tagged方式通过接口。

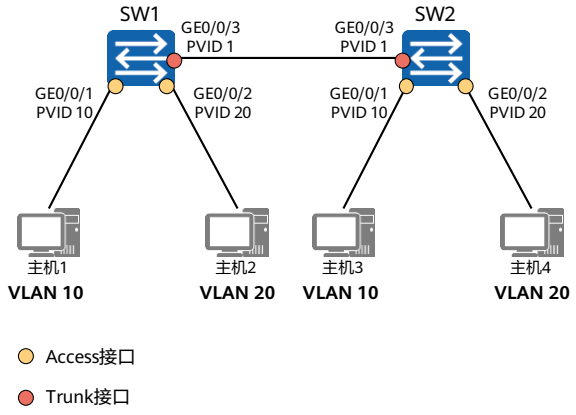
3. (可选) 配置Hybrid接口的缺省VLAN

```
[Huawei-GigabitEthernet0/0/1] port hybrid pvid vlan vlan-id
```

在接口视图下，配置Hybrid类型接口的缺省VLAN。

- 命令：**port hybrid untagged vlan** { { *vlan-id1* [**to** *vlan-id2*] } | **all** }
  - *vlan-id1* [to *vlan-id2*]：指定Hybrid类型接口加入的VLAN，其中：
    - *vlan-id1*表示第一个VLAN的编号。
    - to *vlan-id2*表示最后一个VLAN的编号。*vlan-id2*的取值必须大于等于*vlan-id1*的取值。
    - *vlan-id1*和*vlan-id 2*为整数形式，取值范围是1～4094。
  - all：指定Hybrid接口加入所有VLAN。
- 命令：**port hybrid tagged vlan** { { *vlan-id1* [**to** *vlan-id2*] } | **all** }
  - *vlan-id1* [to *vlan-id2*]：指定Hybrid类型接口加入的VLAN，其中：
    - *vlan-id1*表示第一个VLAN的编号。
    - to *vlan-id2*表示最后一个VLAN的编号。*vlan-id2*的取值必须大于等于*vlan-id1*的取值。
    - *vlan-id1*和*vlan-id 2*为整数形式，取值范围是1～4094。
  - all：指定Hybrid接口加入所有VLAN。
- 命令：**port hybrid pvid vlan** *vlan-id*，设置Hybrid类型接口的缺省VLAN。
  - *vlan-id*：指定Hybrid类型接口的缺省VLAN编号。整数形式，取值范围是1～4094。

案例1：基于接口划分VLAN

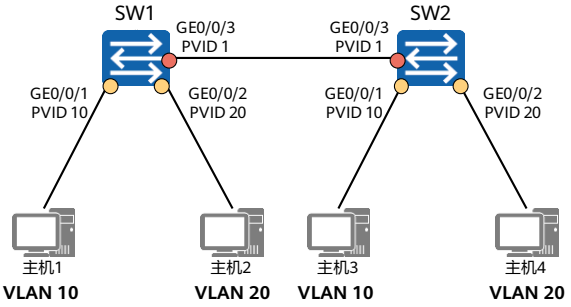


- 组网需求：
  - 某企业的交换机连接有很多用户，且相同业务用户通过不同的设备接入企业网络。为了通信的安全性，企业希望业务相同用户之间可以互相访问，业务不同用户不能直接访问。
  - 可以在交换机上配置基于接口划分VLAN，把业务相同的用户连接的接口划分到同一VLAN。这样属于不同VLAN的用户不能直接进行二层通信，同一VLAN内的用户可以直接互相通信。

  HUAWEI

- 配置思路：
  - 创建VLAN并将连接用户的接口加入VLAN，实现不同业务用户之间的二层流量隔离。
  - 配置SW1和SW2的各接口类型以及通过的VLAN，实现相同业务用户通过SW1和SW2通信。

创建VLAN

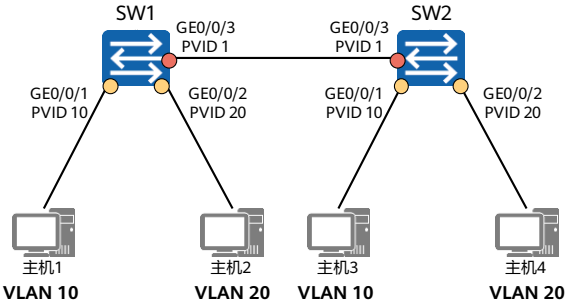


创建VLAN：

```
[SW1] vlan 10
[SW1-vlan10] quit
[SW1] vlan 20
[SW1-vlan20] quit
```

```
[SW2] vlan batch 10 20
```

配置Access接口和Trunk接口



配置Access接口，并加入对应的VLAN：

```
[SW1] interface GigabitEthernet 0/0/1
[SW1-GigabitEthernet0/0/1] port link-type access
[SW1-GigabitEthernet0/0/1] port default vlan 10
```

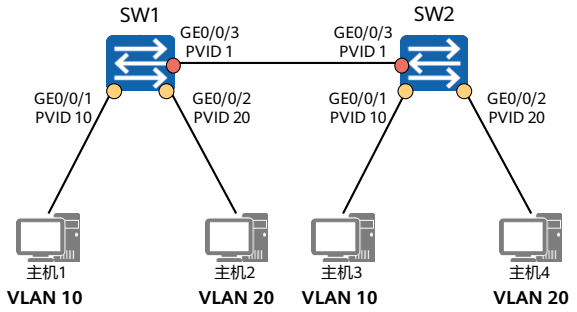
```
[SW1] interface GigabitEthernet 0/0/2
[SW1-GigabitEthernet0/0/2] port link-type access
[SW1] vlan 20
[SW1-vlan20] port GigabitEthernet0/0/2
[SW1-vlan20] quit
```

配置Trunk接口，并创建对应的允许通过列表：

```
[SW1] interface GigabitEthernet 0/0/3
[SW1-GigabitEthernet0/0/3] port link-type trunk
[SW1-GigabitEthernet0/0/3] port trunk pvid vlan 1
[SW1-GigabitEthernet0/0/3] port trunk allow-pass vlan 10 20
```

注：SW2配置与SW1类似

验证配置



[SW1]**display vlan**
The total number of vlans is : 3

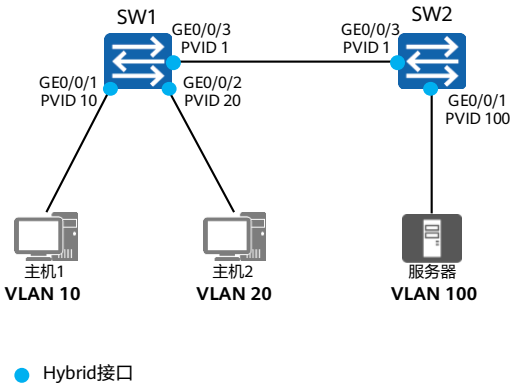
U: Up; D: Down; TG: Tagged; UT: Untagged;
MP: Vlan-mapping; ST: Vlan-stacking;
#: ProtocolTransparent-vlan; *: Management-vlan;

VID	Type	Ports	
1	common	UT:GE0/0/3(U)	.....
10	common	UT:GE0/0/1(U) TG:GE0/0/3(U)	
20	common	UT:GE0/0/2(U) TG:GE0/0/3(U)	

.....

- 命令：**display vlan**命令用来查看VLAN的相关信息。
- 输出信息：
  - Tagged/Untagged Port：手动加入本VLAN的接口，分为Tagged和Untagged方式。
  - VID或VLAN ID：VLAN编号。
  - Type或VLAN Type：VLAN类型，common指普通VLAN。
  - Ports：加入该VLAN的接口。

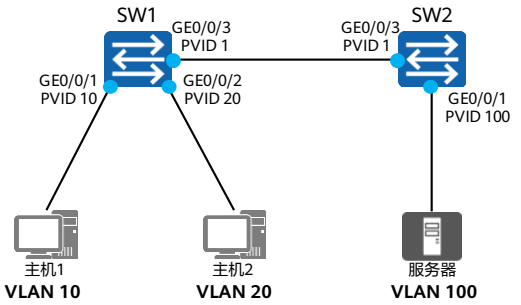
案例2：基于接口划分VLAN



- 组网需求：
  - 某企业的交换机连接有很多用户，且不同部门的用户都需要访问公司服务器。但是为了通信的安全性，企业希望不同部门的用户不能直接访问。
  - 可以在交换机上配置基于接口划分VLAN，并配置Hybrid接口，使得不同部门的用户不能直接进行二层通信，但都可以直接访问公司服务器。

- 配置思路：
  - 创建VLAN并将连接用户的接口加入VLAN，实现不同业务用户之间的二层流量隔离。
  - 配置SW1和SW2的各接口类型以及通过的VLAN，实现主机和服务器之间通过SW1和SW2通信。

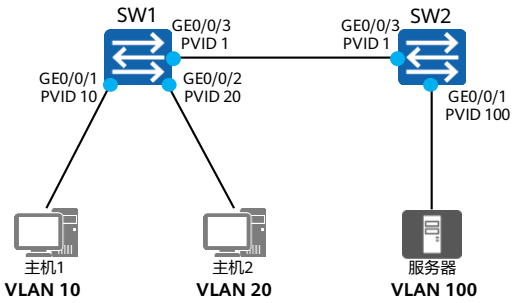
Hybrid接口的基础配置（1）



SW1的配置如下：

```
[SW1] vlan batch 10 20 100
[SW1] interface GigabitEthernet 0/0/1
[SW1-GigabitEthernet0/0/1] port link-type hybrid
[SW1-GigabitEthernet0/0/1] port hybrid pvid vlan 10
[SW1-GigabitEthernet0/0/1] port hybrid untagged vlan 10 100
[SW1-GigabitEthernet0/0/1] interface GigabitEthernet 0/0/2
[SW1-GigabitEthernet0/0/2] port link-type hybrid
[SW1-GigabitEthernet0/0/2] port hybrid pvid vlan 20
[SW1-GigabitEthernet0/0/2] port hybrid untagged vlan 20 100
[SW1-GigabitEthernet0/0/2] interface GigabitEthernet 0/0/3
[SW1-GigabitEthernet0/0/3] port link-type hybrid
[SW1-GigabitEthernet0/0/3] port hybrid tagged vlan 10 20 100
```

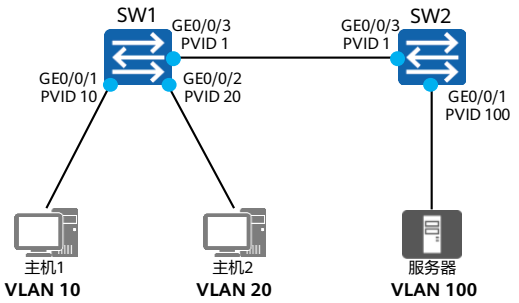
Hybrid接口的基础配置 (2)



SW2的配置如下：

```
[SW2] vlan batch 10 20 100
[SW2] interface GigabitEthernet 0/0/1
[SW2-GigabitEthernet0/0/1] port link-type hybrid
[SW2-GigabitEthernet0/0/1] port hybrid pvid vlan 100
[SW2-GigabitEthernet0/0/1] port hybrid untagged vlan 10  20 100
[SW2-GigabitEthernet0/0/1] interface GigabitEthernet 0/0/3
[SW2-GigabitEthernet0/0/3] port link-type hybrid
[SW2-GigabitEthernet0/0/3] port hybrid tagged vlan 10 20 100
```

验证配置




```
[SW1]display vlan
The total number of vlans is : 4
--------------------------------------------------------------------------------
U: Up;         D: Down;                TG: Tagged;  UT: Untagged;
MP: Vlan-mapping;                      ST: Vlan-stacking;
#: ProtocolTransparent-vlan;           *: Management-vlan;
--------------------------------------------------------------------------------
VID         Type        Ports
--------------------------------------------------------------------------------
1           common      UT:GE0/0/1(U)     GE0/0/2(U)     GE0/0/3(U) .....
10          common      UT:GE0/0/1(U)
                        TG:GE0/0/3(U)
20          common      UT:GE0/0/2(U)
                        TG:GE0/0/3(U)
100         common      UT:GE0/0/1(U)     GE0/0/2(U)
                        TG:GE0/0/3(U)
.....
```

VLAN的基础配置命令

1. 关联MAC地址与VLAN

```
[Huawei-vlan10] mac-vlan mac-address mac-address [ mac-address-mask | mac-address-mask-length ]
```

通过此命令配置MAC地址与VLAN关联。

- *mac-address*：指定与VLAN关联的MAC地址。格式为H-H-H。其中H为4位的十六进制数，可以输入1～4位，如00e0、fc01。当输入不足4位时，表示前面的几位为0，如：输入e0，等同于00e0。MAC地址不可设置为0000-0000-0000、FFFF-FFFF-FFFF和组播地址。
- *mac-address-mask*：指定MAC地址掩码。格式为H-H-H，其中H为1至4位的十六进制数。
- *mac-address-mask-length*：指定MAC地址掩码长度。整数形式，取值范围是1～48。

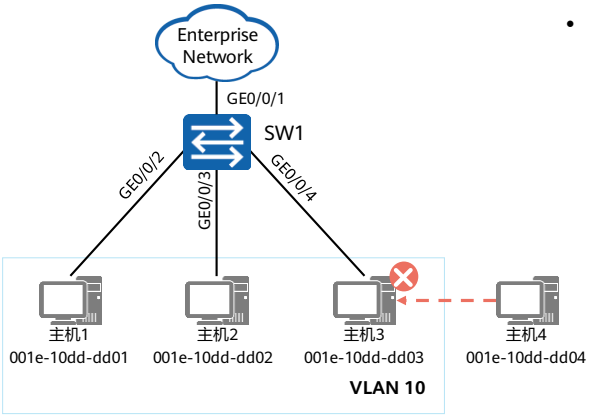
2. 使能MAC地址与VLAN

```
[Huawei-GigabitEthernet0/0/1] mac-vlan enable
```

通过此命令使能接口的MAC VLAN功能。

- 命令：**mac-vlan mac-address** *mac-address* [*mac-address-mask* | *mac-address-mask-length*]
  - *mac-address*：指定与VLAN关联的MAC地址。
    - 格式为H-H-H。其中H为4位的十六进制数，可以输入1～4位，如00e0、fc01。当输入不足4位时，表示前面的几位为0，如：输入e0，等同于00e0。
    - MAC地址不可设置为0000-0000-0000、FFFF-FFFF-FFFF和组播地址。
  - *mac-address-mask*：指定MAC地址掩码。
    - 格式为H-H-H，其中H为1至4位的十六进制数。
  - *mac-address-mask-length*：指定MAC地址掩码长度。
    - 整数形式，取值范围是1～48。
- 命令：**mac-vlan enable**，用来使能接口的MAC VLAN功能。

案例：基于MAC地址划分VLAN

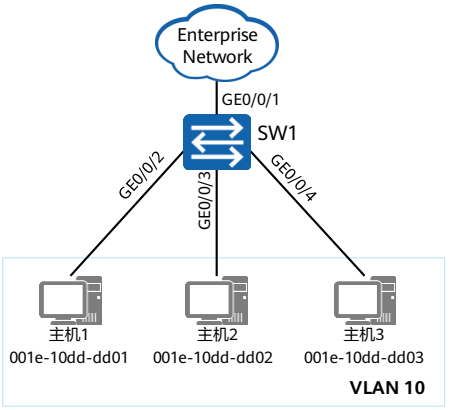


- 组网需求：
  - 某个公司的网络中，网络管理者将同一部门的员工划分到同一VLAN。为了提高部门内的信息安全，要求只有本部门员工的主机才可以访问公司网络。
  - 主机1、主机2、主机3为本部门员工的主机，要求这几台主机可以通过SW1访问公司网络，如换成其他主机则不能访问。
  - 可以配置基于MAC地址划分VLAN，将本部门员工主机的MAC地址与VLAN绑定，从而实现该需求。

  HUAWEI

- 配置思路：
  - 创建VLAN。
  - 配置各以太网接口以正确的方式加入VLAN。
  - 配置主机1、主机2、主机3的MAC地址与VLAN关联，实现根据报文中的源MAC地址确定VLAN。

创建VLAN，并关联MAC地址和VLAN



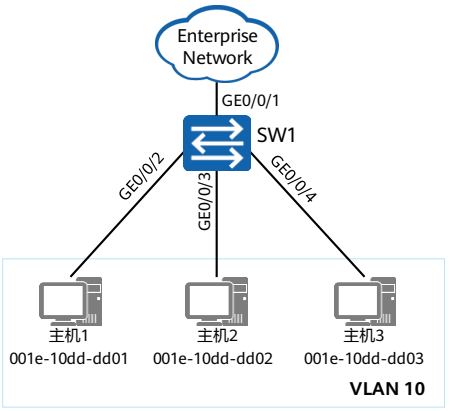
创建VLAN：

```
[SW1] vlan 10
[SW1-vlan10] quit
```

关联MAC地址和VLAN：

```
[SW1] vlan 10
[SW1-vlan10] mac-vlan mac-address 001e-10dd-dd01
[SW1-vlan10] mac-vlan mac-address 001e-10dd-dd02
[SW1-vlan10] mac-vlan mac-address 001e-10dd-dd03
[SW1-vlan10] quit
```

加入VLAN，并使能MAC VLAN功能



加入VLAN：

```
[SW1] interface gigabitethernet 0/0/1
[SW1-GigabitEthernet0/0/1] port link-type hybrid
[SW1-GigabitEthernet0/0/1] port hybrid tagged vlan 10

[SW1] interface gigabitethernet 0/0/2
[SW1-GigabitEthernet0/0/2] port link-type hybrid
[SW1-GigabitEthernet0/0/2] port hybrid untagged vlan 10
```

使能接口的基于MAC地址划分VLAN功能：

```
[SW1] interface gigabitethernet 0/0/2
[SW1-GigabitEthernet0/0/2] mac-vlan enable
[SW1-GigabitEthernet0/0/2] quit
```

注：GE0/0/3、GE0/0/4的配置与GE0/0/2类似

- 配置接口为Hybrid接口：在Access接口和Trunk接口上，只有基于MAC划分的VLAN和PVID相同时，才能使用MAC VLAN功能。所以基于MAC地址划分VLAN推荐在Hybrid口上配置。

验证配置

```
[SW1]display vlan
The total number of vlans is : 2
--------------------------------------------------------------------------------
U: Up;         D: Down;         TG: Tagged;  UT: Untagged;
MP: Vlan-mapping;               ST: Vlan-stacking;
#: ProtocolTransparent-vlan;    *: Management-vlan;
--------------------------------------------------------------------------------
VID        Type        Ports
--------------------------------------------------------------------------------
1          common      UT:GE0/0/1(U)            GE0/0/2(U)   GE0/0/3(U) .....
10         common      UT:GE0/0/2(U)            GE0/0/3(U)   GE0/0/4(U)
                       TG:GE0/0/1(U)
.....
```

```
[SW1]display mac-vlan mac-address all
--------------------------------------------------------------------------------
MAC Address          MASK           VLAN       Priority
--------------------------------------------------------------------------------
001e-10dd-dd01       ffff-ffff-ffff 10         0
001e-10dd-dd02       ffff-ffff-ffff 10         0
001e-10dd-dd03       ffff-ffff-ffff 10         0

Total MAC VLAN address count: 3
```

  HUAWEI

- 命令：**display mac-vlan** { **mac-address** { **all** | *mac-address* [*mac-address-mask* | *mac-address-mask-length*] } | **vlan** *vlan-id* }，用来查看基于MAC地址划分VLAN的配置信息。
  - all：显示所有MAC地址VLAN划分信息。
  - mac-address *mac-address*：显示指定MAC地址的VLAN划分信息。
    - 格式为H-H-H，其中H为1至4位的十六进制数。
  - *mac-address-mask*：MAC地址掩码。
    - 格式为H-H-H，其中H为1至4位的十六进制数。
  - *mac-address-mask-length*：MAC地址掩码长度。
    - 整数形式，取值范围是1～48。
  - vlan *vlan-id*：显示指定MAC-VLAN的配置信息。
    - 整数形式，取值范围是1～4094。
- 输出信息：
  - MAC Address：MAC地址。
  - MASK：MAC地址的掩码。
  - VLAN：基于MAC地址划分的VLAN。
  - Priority：指定MAC地址对应VLAN的802.1P优先级。

思考题

1. (多选) 下列关于VLAN的描述中，错误的是？（　　）
   A. VLAN技术可以将一个规模较大的冲突域隔离成若干个规模较小的冲突域
   B. VLAN技术可以将一个规模较大的二层广播域隔离成若干个规模较小的二层广播域
   C. 位于不同VLAN的计算机之间无法进行通信
   D. 位于同一VLAN中的计算机之间可以进行二层通信
2. 如果一个Trunk接口的PVID是5，且端口下配置port trunk allow-pass vlan 2 3，那么哪些VLAN的流量可以通过该Trunk接口进行传输？

1. AC
2. 执行了port trunk allow-pass vlan 2 3命令后，VLAN 5的数据帧不能在此接口上进行传输。VLAN 1的数据默认可以通过Trunk接口进行传输。所以VLAN 1，VLAN 2和VLAN 3的数据帧可以在Trunk接口上传输。

本章总结

- 本章节主要介绍了虚拟局域网（VLAN)的相关技术知识，包括：VLAN的作用，VLAN的标识及划分，VLAN的数据交互，VLAN的实际规划和应用，以及VLAN的相关基本配置。
- 通过VLAN技术，可以将物理的局域网划分成多个广播域，实现同一VLAN内的网络设备可以直接进行二层通信，不同VLAN内的设备不可以直接进行二层通信。


谢 谢
www.huawei.com

生成树

前言

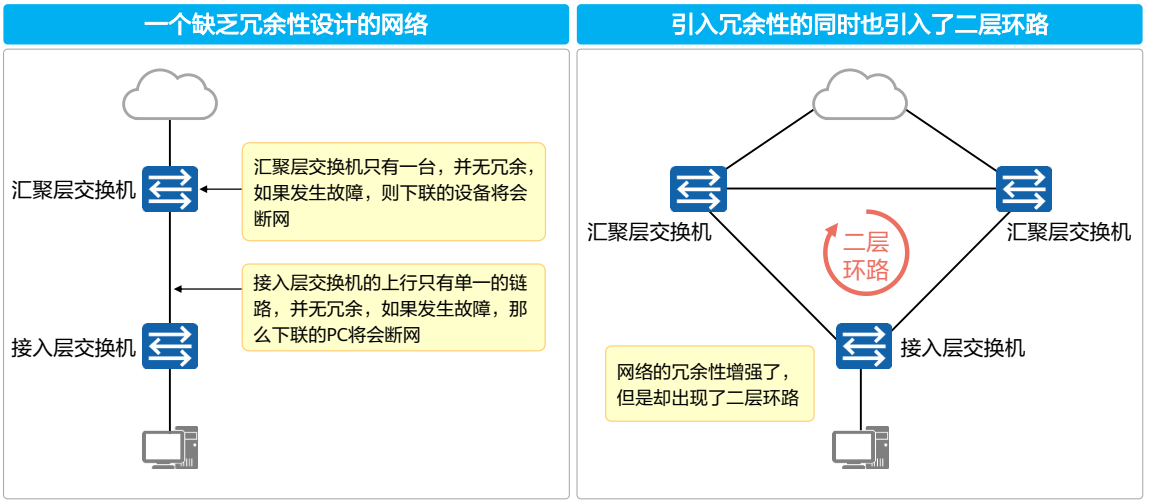
- 以太网交换网络中为了进行链路备份，提高网络可靠性，通常会使用冗余链路。但是使用冗余链路会在交换网络上产生环路，引发广播风暴以及MAC地址表不稳定等故障现象，从而导致用户通信质量较差，甚至通信中断。为解决交换网络中的环路问题，提出了生成树协议STP（Spanning Tree Protocol）。
- 运行STP协议的设备通过彼此交互信息发现网络中的环路，并有选择的对某个接口进行阻塞，最终将环形网络结构修剪成无环路的树形网络结构，从而防止报文在环形网络中不断循环，避免设备由于重复接收相同的报文造成处理能力下降。
- RSTP（Rapid Spanning Tree Protocol）协议基于STP协议，对原有的STP协议进行了更加细致的修改和补充，实现了网络拓扑快速收敛。

目标

- 学完本课程后，您将能够：
  - 描述园区交换网络中的二层环路产生原因及引发的问题。
  - 描述STP的基本概念与工作原理。
  - 区分STP与RSTP，并能够描述RSTP对STP的改进。
  - 完成STP的基础配置。
  - 了解除了生成树之外的其他消除交换网络二层环路的方法。

目录

技术背景：二层交换机网络的冗余性与环路

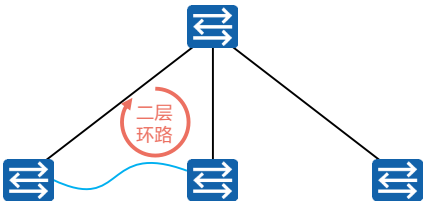


HUAWEI

- 随着局域网规模的不断扩大，越来越多的交换机被用来实现主机之间的互连。如图，接入层交换机单链路上联，则存在单链路故障，也就是如果这根上联链路发生故障，交换机下联用户就断网了。另一个问题的单点故障，也就是交换机如果宕机，交换机下联用户也就断网了。
- 为了解决此类问题，交换机在互连时一般都会使用冗余链路来实现备份。冗余链路虽然增强了网络的可靠性，但是也会产生环路，而环路会带来一系列的问题，继而导致通信质量下降和通信业务中断等问题。

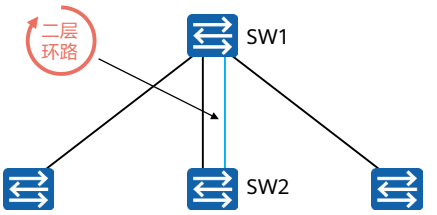
技术背景：人为错误导致的二层环路

人为错误导致的二层环路 案例1



在现实中，一些二层环路可能是由于人为的疏忽导致的，例如错误地连接设备之间的互联线缆等。

人为错误导致的二层环路 案例2

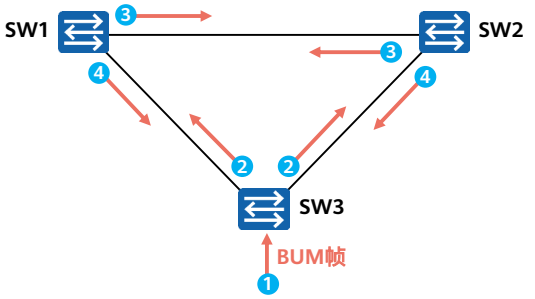


另一些二层环路可能是由于人为的配置错误导致的，在本例中，网络管理员未将SW1与SW2之间的链路绑定到一个逻辑链路（聚合链路）上，从而引入了二层环路。

- 在现实中，除了冗余链路会引起环路，还有一些人为错误导致的环路。

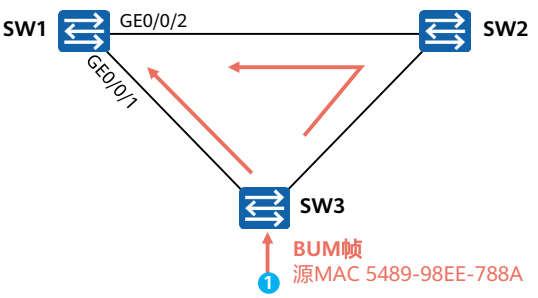
二层环路带来的问题

典型问题1：广播风暴



SW3收到BUM帧后将其进行泛洪，SW1及SW2收到后进一步泛洪，如此反复，最终导致整个网络资源被耗尽，网络瘫痪不可用。

典型问题2：MAC地址漂移

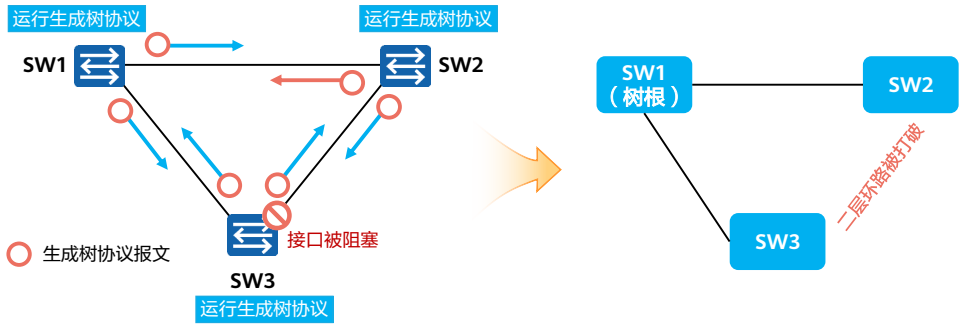


以SW1为例，5489-98EE-788A会不断地在GE0/0/1与GE0/0/2接口之间来回切换，这被称为MAC地址漂移现象。

BUM帧（Broadcast，Unknown unicast，Multicast）指定广播、未知单播及组播帧

- 问题一：广播风暴
  - 根据交换机的转发原则，如果交换机从一个端口上接收到的是一个广播帧，或者是一个目的MAC地址未知的单播帧，则会将这个帧向除源端口之外的所有其他端口转发。如果交换网络中有环路，则这个帧会被无限转发，此时便会形成广播风暴，网络中也会充斥着重复的数据帧。
  - 本例中，SW3收到了一个广播帧将其进行泛洪，SW1和SW2也会将此帧转发到除了接收此帧的其他所有端口，结果此帧又会被再次转发给SW3，这种循环会一直持续，于是便产生了广播风暴。交换机性能会因此急速下降，并会导致业务中断。
- 问题二：MAC地址表漂移
  - 交换机是根据所接收到的数据帧的源地址和接收端口生成MAC地址表项的。
  - 本例中，SW3收到一个广播帧泛洪，SW1从GE0/0/1接口接收到广播帧后学习且泛洪，形成MAC地址5489-98EE-788A与GE0/0/1的映射；SW2收到广播帧后学习且泛洪，SW1再次从GE0/0/2收到源MAC地址为5489-98EE-788A的广播帧并进行学习，5489-98EE-788A会不断地在GE0/0/1与GE0/0/2接口之间来回切换，这被称为MAC地址漂移现象。

初识生成树协议



在网络中部署生成树后，交换机之间会进行生成树协议报文的交互并进行无环拓扑计算，最终将网络中的某个（或某些）接口进行阻塞（Block），从而打破环路。

HUAWEI

- 在以太网中，二层网络的环路会带来广播风暴，MAC地址表震荡，重复数据帧等问题，为解决交换网络中的环路问题，提出了STP。
- STP通过构造一棵树来消除交换网络中的环路。
- 运行STP算法，判断网络中存在环路的地方并阻断冗余链路，将环路网络修剪成无环路的树型网络，从而避免了数据帧在环路网络中的增生和无穷循环。

生成树能够动态响应网络拓扑变化调整阻塞接口

交换机上运行的生成树协议会持续监控网络的拓扑结构，当网络拓扑结构发生变化时，生成树能感知到这些变化，并且自动做出调整。

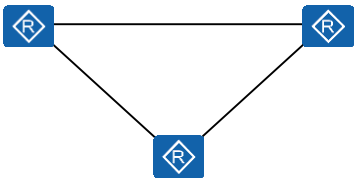
因此，生成树既能解决二层环路问题，也能为网络的冗余性提供一种方案。

HUAWEI

- 如图，交换机上运行STP协议，会通过报文监控网络的拓扑结构，正常情况下是将SW3上的一个接口进行阻塞（Block），从而打破环路，当监控到SW1与SW3之间出现链路故障，则恢复阻塞端口进入转发状态。

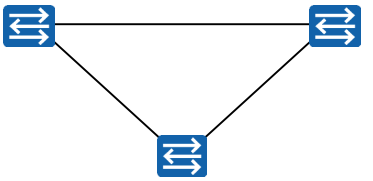
问答：二层及三层环路

三层环路（Layer 3 Loop）



- 常见根因：路由环路；
- 动态路由协议有一定的防环能力；
- IP报文头部中的TTL字段可用于防止报文被无止尽地转发。

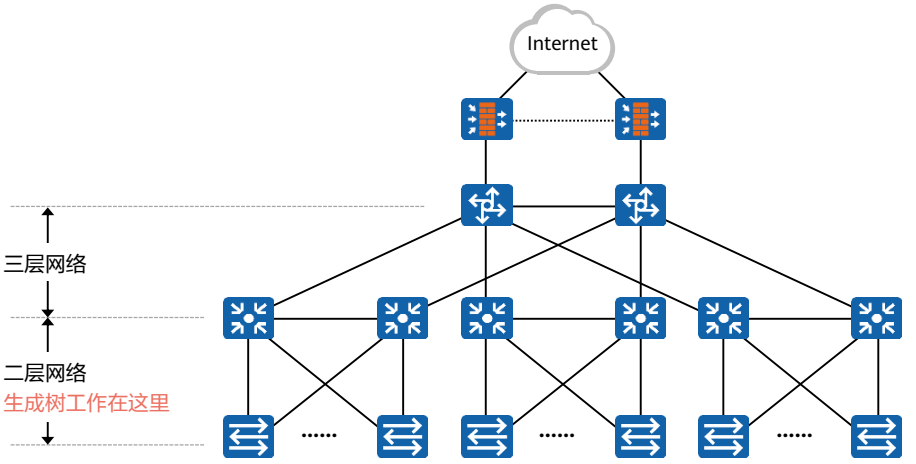
二层环路（Layer 2 Loop）



- 常见根因：网络中部署了二层冗余环境，或人为的误接线缆导致；
- 需借助特定的协议或机制实现二层防环；
- 二层帧头中并没有任何信息可用于防止数据帧被无止尽地转发。

- 常见环路主要分为二层环路和三层环路。
  - 二层环路主要因为网络中部署了二层冗余环境，或人为的误接线缆导致，可以通过借助特定的协议或机制实现二层防环；
  - 三层环路主要因为路由环路，可以通过动态路由协议防环和IP报文头部中的TTL字段用于防止报文被无止尽地转发。

生成树协议在园区网络中的应用位置



HUAWEI

- 生成树协议应用于园区网络的二层网络中，进行链路备份和消除环路。

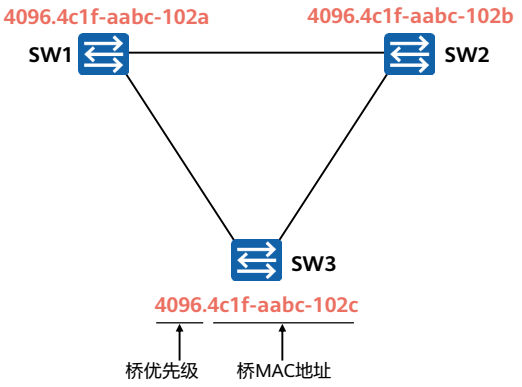
STP概述

- STP是一个用于局域网中消除环路的协议。
- 运行该协议的设备通过彼此交互信息而发现网络中的环路，并对某些接口进行阻塞以消除环路。
- STP在网络中运行后会持续监控网络的状态，当网络出现拓扑变更时，STP能够感知并且进行自动响应，从而使得网络状态适应新的拓扑结构，保证网络可靠性。
- 由于局域网规模的不断增长，生成树协议已经成为了当前最重要的局域网协议之一。

目录

桥ID 根桥 Cost RPC Port ID BPDU

STP的基本概念：桥ID



桥ID（Bridge ID，BID）

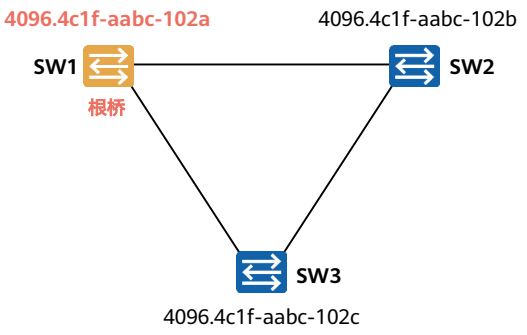
- IEEE 802.1D标准中规定BID由16位的桥优先级（Bridge Priority）与桥MAC地址构成。
- 每一台运行STP的交换机都拥有一个唯一的BID。
- BID桥优先级占据高16bit，其余的低48bit是桥MAC地址。
- 在STP网络中，BID最小的设备会被选举为根桥。

备注：此处网桥（Bridge），或者桥也就是交换机。

HUAWEI

- 在STP中，每一台交换机都有一个标示符，叫做Bridge ID或者桥ID，桥ID由16位的桥优先级（Bridge Priority）和48位的MAC地址构成。在STP网络中，桥优先级是可以配置的，取值范围是0～65535，默认值为32768,可以修改但是修改值必须为4096的倍数。优先级最高的设备（数值越小越优先）会被选举为根桥。如果优先级相同，则会比较MAC地址，MAC地址越小则越优先。
- 如图，需要在该网络中选举根桥，首先比较三台交换机的桥优先级，桥优先级都为4096，再比较三台交换机的MAC地址，谁小谁优先，最终选择SW1为根桥。

STP的基本概念：根桥



根桥（Root Bridge）

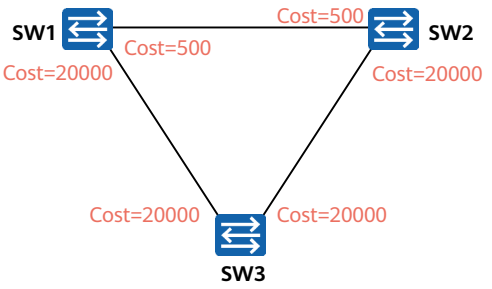
- STP的主要作用之一是在整个交换网络中计算出一棵无环的“树”（STP树）。
- 根桥是一个STP交换网络中的“树根”。
- STP开始工作后，会在交换网络中选举一个根桥，根桥是生成树进行拓扑计算的重要“参考点”，是STP计算得出的无环拓扑的“树根”。
- 在STP网络中，桥ID最小的设备会被选举为根桥。

  在BID的比较过程中，首先比较桥优先级，优先级的值越小，则越优先，拥有最小优先级值的交换机会成为根桥；如果优先级相等，那么再比较MAC地址，拥有最小MAC地址的交换机会成为根桥。

HUAWEI

- 树形的网络结构必须有树根，于是STP引入了根桥（Root Bridge）概念。
- 对于一个STP网络，根桥在全网中只有一个，它是整个网络的逻辑中心，但不一定是物理中心。根桥会根据网络拓扑的变化而动态变化。
- 网络收敛后，根桥会按照一定的时间间隔产生并向外发送配置BPDU，其他设备仅对该报文进行处理，传达拓扑变化记录，从而保证拓扑的稳定。

STP的基本概念：Cost



开销（Cost）

- 每一个激活了STP的接口都维护着一个Cost值，接口的Cost主要用于计算根路径开销，也就是到达根的开销。
- 接口的缺省Cost除了与其速率、工作模式有关，还与交换机使用的STP Cost计算方法有关。
- 接口带宽越大，则Cost值越小。
- 用户也可以根据需要通过命令调整接口的Cost。

HUAWEI

- 交换机的每个端口都有一个端口开销（Port Cost）参数，此参数表示该端口在STP中的开销值。默认情况下端口的开销和端口的带宽有关，带宽越高，开销越小。
- 华为交换机支持多种STP的路径开销计算标准，提供多厂商场景下最大程度的兼容性。缺省情况下，华为交换机使用IEEE 802.1t标准来计算路径开销。

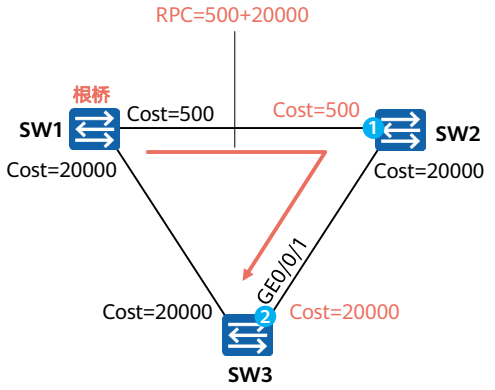
STP的基本概念：Cost计算方法

接口速率	接口模式	STP开销（推荐值）		
		IEEE 802.1d-1998标准	IEEE 802.1t标准	华为计算方法
100Mbps	Half-Duplex	19	200,000	200
	Full-Duplex	18	199,999	199
	Aggregated Link 2 Ports	15	100,000	180
1000Mbps	Full-Duplex	4	20,000	20
	Aggregated Link 2 Ports	3	10,000	18
10Gbps	Full-Duplex	2	2000	2
	Aggregated Link 2 Ports	1	1000	1
40Gbps	Full-Duplex	1	500	1
	Aggregated Link 2 Ports	1	250	1
100Gbps	Full-Duplex	1	200	1
	Aggregated Link 2 Ports	1	100	1
.....				

接口Cost是已经激活了STP的接口所维护的一个开销值，该值存在默认值，与接口的速率有关联，并且设备使用不同的算法时，相同的接口速率对应不同的Cost值。

桥ID 根桥 Cost RPC Port ID BPDU

STP的基本概念：RPC



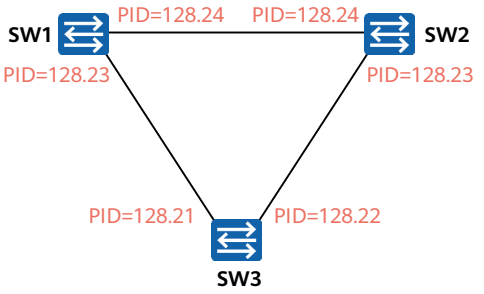
根路径开销（Root Path Cost）

- 在STP的拓扑计算过程中，一个非常重要的环节就是“丈量”交换机某个接口到根桥的“成本”，也即RPC。
- 一台设备从某个接口到达根桥的RPC等于从根桥到该设备沿途所有入方向接口的Cost累加。
- 在本例中，SW3从GE0/0/1接口到达根桥的RPC等于接口1的Cost加上接口2的Cost。

HUAWEI

- 从一个非根桥到达根桥的路径可能有多条，每一条路径都有一个总的开销值，此开销值是该路径上所有接收BPDU端口的端口开销总和（即BPDU的入方向端口），称为路径开销。非根桥通过对比多条路径的路径开销，选出到达根桥的最短路径，这条最短路径的路径开销被称为RPC，并生成无环树状网络。根桥的根路径开销是0。

STP的基本概念：Port ID



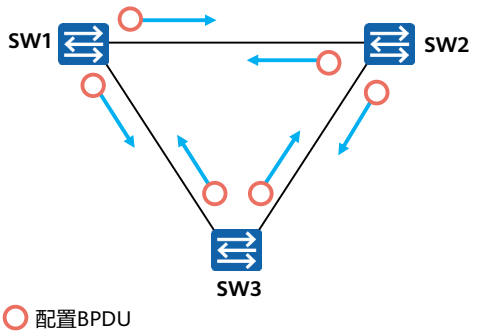
接口ID（Port ID，PID）

- 运行STP的交换机使用接口ID来标识每个接口，接口ID主要用于在特定场景下选举指定接口。
- 接口ID由两部分构成的，高4 bit是接口优先级，低12 bit是接口编号。
- 激活STP的接口会维护一个缺省的接口优先级，在华为交换机上，该值为128。用户可以根据实际需要，通过命令修改该优先级。

- 运行STP交换机的每个端口都有一个端口ID，端口ID由端口优先级和端口号构成。端口优先级取值范围是0到240，步长为16，即取值必须为16的整数倍。缺省情况下，端口优先级是128。端口ID可以用来确定端口角色。

桥ID 根桥 Cost RPC Port ID BPDU

STP的基本概念：BPDU



BPDU（Bridge Protocol Data Unit，网桥协议数据单元）

- BPDU是STP能够正常工作的根本。BPDU是STP的协议报文。
- STP交换机之间会交互BPDU报文，这些BPDU报文携带着一些重要信息，正是基于这些信息，STP才能够顺利工作。
- BPDU分为两种类型：
  - 配置BPDU（Configuration BPDU）
  - TCN BPDU（Topology Change Notification BPDU）
- 配置BPDU是STP进行拓扑计算的关键；TCN BPDU只在网络拓扑发生变更时才会被触发。

HUAWEI

- 为了计算生成树，交换机之间需要交换相关的信息和参数，这些信息和参数被封装在BPDU中。
- BPDU有两种类型：配置BPDU和TCN BPDU。
- 配置BPDU包含了桥ID、路径开销和端口ID等参数。STP协议通过在交换机之间传递配置BPDU来选举根交换机，以及确定每个交换机端口的角色和状态。在初始化过程中，每个桥都主动发送配置BPDU。在网络拓扑稳定以后，只有根桥主动发送配置BPDU，其他交换机在收到上游传来的配置BPDU后，才会发送自己的配置BPDU。
- TCN BPDU是指下游交换机感知到拓扑发生变化时向上游发送的拓扑变化通知。

配置BPDU的报文格式

PID	PVI	BPDU Type	Flags	Root ID	RPC	Bridge ID	Port ID	Message Age	Max Age	Hello Time	Forward Delay

字节	字段	描述
2	PID	协议ID ，对于STP而言，该字段的值总为0
1	PVI	协议版本ID，对于STP而言，该字段的值总为0
1	BPDU Type	指示本BPDU的类型，若值为0x00，则表示本报文为配置BPDU；若值为0x80，则为TCN BPDU
1	Flags	标志，STP只使用了该字段的最高及最低两个比特位，最低位是TC（Topology Change，拓扑变更）标志，最高位是TCA（Topology Change Acknowledgment，拓扑变更确认）标志
8	Root ID	根网桥的桥ID
4	RPC	根路径开销，到达根桥的STP Cost
8	Bridge ID	BPDU发送桥的ID
2	Port ID	BPDU发送网桥的接口ID（优先级+接口号）
2	Message Age	消息寿命，从根网桥发出BPDU之后的秒数，每经过一个网桥都加1，所以它本质上是到达根桥的跳数
2	Max Age	最大寿命，当一段时间未收到任何BPDU，生存期到达最大寿命时，网桥认为该接口连接的链路发生故障。默认20s
2	Hello Time	根网桥连续发送的BPDU之间的时间间隔，默认2s
2	Forward Delay	转发延迟，在侦听和学习状态所停留的时间间隔，默认15s

配置BPDU的比较原则

字段
协议ID
协议版本ID
类型
标志
根桥ID
根路径开销
网桥ID
接口ID
消息寿命
最大寿命
Hello时间
转发延迟

对于STP而言，最重要的工作就是在交换网络中计算出一个无环拓扑。在拓扑计算的过程中，一个非常重要的内容就是配置BPDU的比较。在配置BPDU中，有四个字段非常关键，它们是“根桥ID”、“根路径开销”、“网桥ID”以及“接口ID”，这四个字段便是交换机进行配置BPDU比较的关键内容。

STP按照如下顺序选择最优的配置BPDU：

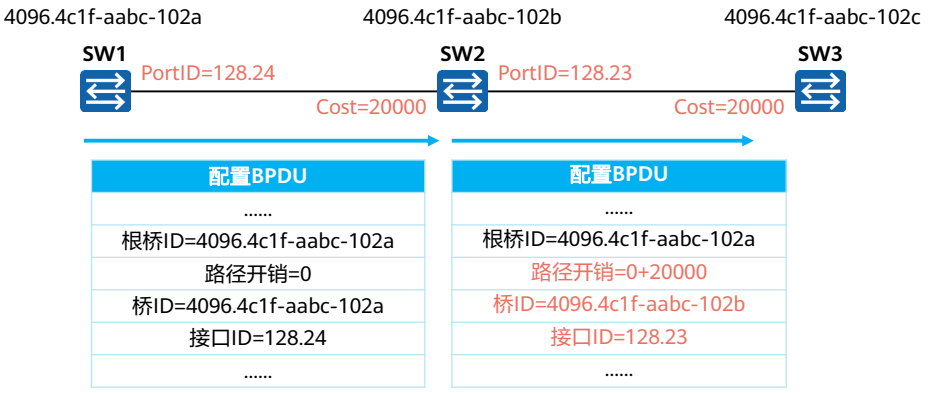
1. 最小的根桥ID
2. 最小的RPC
3. 最小的网桥ID
4. 最小的接口ID

在这四条原则中（每条原则都对应配置BPDU中的相应字段），第一条原则主要用于在网络中选举根桥，后面的原则主要用于选举根接口及指定接口。

HUAWEI

- STP操作：
  1. 选举一个根桥。
  2. 每个非根交换机选举一个根端口。
  3. 每个网段选举一个指定端口。
  4. 阻塞非根、非指定端口。
- STP中定义了三种端口角色：指定端口，根端口和预备端口。
  - 指定端口是交换机向所连网段转发配置BPDU的端口，每个网段有且只能有一个指定端口。一般情况下，根桥的每个端口总是指定端口。
  - 根端口是非根交换机去往根桥路径最优的端口。在一个运行STP协议的交换机上最多只有一个根端口，但根桥上没有根端口。
  - 如果一个端口既不是指定端口也不是根端口，则此端口为预备端口。预备端口将被阻塞。

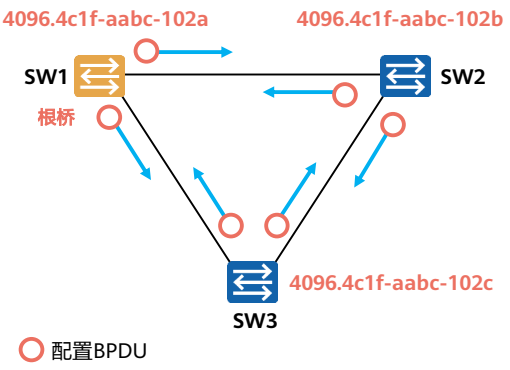
配置BPDU的转发过程



- 交换机在刚启动时都认为自己是根桥，互相发送配置BPDU进行STP运算。

选举根桥 选举根接口 选举指定接口 阻塞非指定接口

STP的计算过程（1）



在交换网络中选举一个根桥

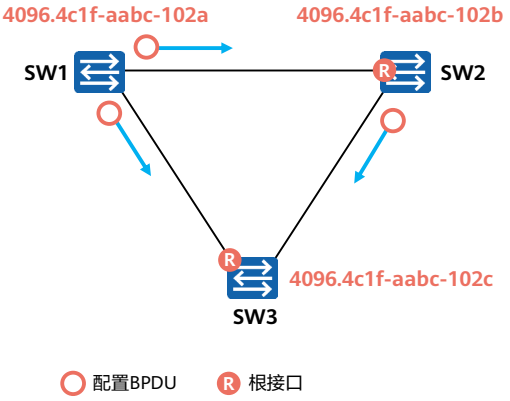
- STP在交换网络中开始工作后，每个交换机都会向网络中发送配置BPDU。配置BPDU中包含交换机自己的桥ID。
- 网络中拥有最小桥ID的交换机成为根桥。
- 在一个连续的STP交换网络中只会存在一个根桥。
- 根桥的角色是可抢占的。
- 为了确保交换网络的稳定，建议提前规划STP组网，并将规划为根桥的交换机的桥优先级设置为最小值0。

HUAWEI

- 什么是根桥?
  - 根桥是STP树的根节点。
  - 要生成一棵STP树，首先要确定出一个根桥。
  - 根桥是整个交换网络的逻辑中心，但不一定是它的物理中心。
  - 当网络的拓扑发生变化时，根桥也可能发生变化。（抢占）
- 选举过程：
  - STP交换机初始启动之后，都会认为自己是根桥，并在发送给其他交换机的BPDU中宣告自己为根桥。因此，此时BPDU中的根桥ID为各自设备的网桥ID。
  - 当交换机收到网络中其他设备发送来的BPDU后，会比较BPDU中的根桥ID和自己的BID。
  - 交换机不断交互BPDU，同时对BID进行比较，最终选举一台BID最小的交换机作为根桥，其他的则为非根桥。
  - 如图：根桥的选举先比较优先级，交换机SW1、2、3的优先级相等，则比较MAC地址，也优选最小的，所以SW1的BID最小，因此SW1为根桥，SW2和SW3为非根桥。
- 注意：
  - 根桥的角色可抢占。当有更优的BID的交换机加入网络时，网络会重新进行STP计算，选出新的根桥。

选举根桥 > 选举根接口 > 选举指定接口 > 阻塞非指定接口

STP的计算过程 (2)



在每台非根桥上选举一个根接口

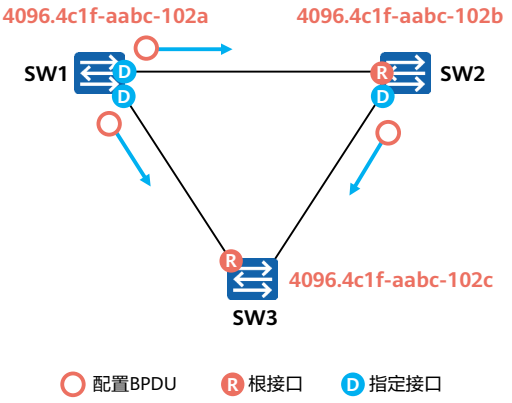
- 每一台非根桥交换机都会在自己的接口中选举出一个接口。
- 非根桥交换机上有且只会有一个根接口。
- 当非根桥交换机有多个接口接入网络中时，根接口是其收到最优配置BPDU的接口。
- 可以形象地理解为，根接口是每台非根桥上“朝向”根桥的接口。

HUAWEI

- 什么是根端口?
  - 一个非根桥设备上会有多个端口与网络相连，为了保证从某台非根桥设备到根桥设备的工作路径是最优且唯一的，就必须从该非根桥设备的端口中确定出一个被称为“根端口”的端口，由根端口来作为该非根桥设备与根桥设备之间进行报文交互的端口。
  - 在选举出根桥后，根桥仍然持续发送BPDU，而非根桥将持续不断的收到根桥发送的BPDU。因此，在所有非根桥上选举一个距离根桥“最近”的端口（根端口），在网络收敛后，根端口将不断的收到来自根桥的BPDU。
  - 即：根端口保证了交换机与根桥之间工作路径的唯一性和最优性。
- 注意：一个非根桥设备上，最多只能有一个根端口。
- 选举过程：
  1. 交换机有多个端口接入网络，各个端口都会收到BPDU报文，报文中会携带“RootID、RPC、BID、PID”等关键字段，端口会针对这些字段进行PK。
  2. 首先比较根路径开销（RPC），STP协议把根路径开销作为确定根端口的重要依据。RPC值越小，越优选，因此交换机会选RPC最小的端口作为根端口。
  3. 当RPC相同时，比较上行交换机的BID，即比较交换机各个端口收到的BPDU中的BID，值越小，越优选，因此交换机会选上行设备BID最小的端口作为根端口。
  4. 当上行交换机BID相同时，比较上行交换机的PID，即比较交换机各个端口收到的BPDU中的PID，值越小，越优先，因此交换机会选上行设备PID最小的端口作为根端口。
  5. 当上行交换机的PID相同时，则比较本地交换机的PID，即比较本端交换机各个端口各自的PID，值越小，越优先，因此交换机会选端口PID最小的端口作为根端口。

选举根桥 选举根接口 选举指定接口 阻塞非指定接口

STP的计算过程 (3)



在每条链路上选举一个指定接口

- 根接口选举出来后，非根桥会使用其在该接口上收到的最优BPDU进行计算，然后将计算得到的配置BPDU与除了根接口之外的其他所有接口所收到的配置BPDU进行比较：
  - 如果前者更优，则该接口为指定接口；
  - 如果后者更优，则该接口为非指定接口。
- 一般情况下，根桥的所有接口都是指定接口。

HUAWEI

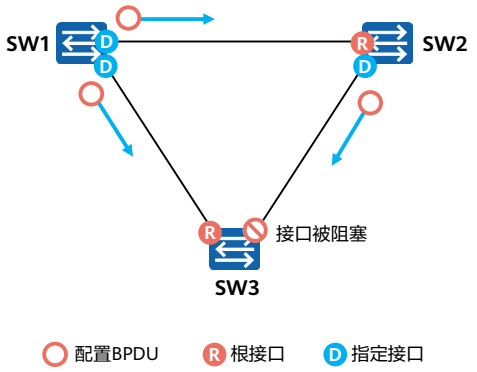
- 什么是指定端口?
  - 网络中的每个链路与根桥之间的工作路径必须是唯一的且最优的。当一个链路有两条及以上的路径通往根桥时（该链路连接了不同的交换机，或者该链路连接了同一台交换机的不同端口），与该链路相连的交换机（可能不止一台）就必须确定出一个唯一的指定端口。
  - 因此，每个链路（Link）选举一个指定端口，用于向这个链路发送BPDU。
- 注意：一般情况下，根桥上不存在任何根端口，只存在指定端口。
- 选举过程：

  指定端口也是通过比较RPC来确定的，选择RPC最小的作为指定端口，如果RPC相同，则比较BID和PID。

  1. 首先比较根路径开销（RPC），值越小，越优选，因此交换机会选RPC最小的端口作为指定端口。
  2. 若RPC相等，则比较链路两端交换机的BID，值越小，越优选，因此交换机会选BID最小的交换机的端口作为指定端口。
  3. 若BID相等，则比较链路两端端口的PID，值越小，越优选，因此交换机会选PID最小的交换机的端口作为指定端口。

选举根桥 选举根接口 选举指定接口 阻塞非指定接口

STP的计算过程 (4)

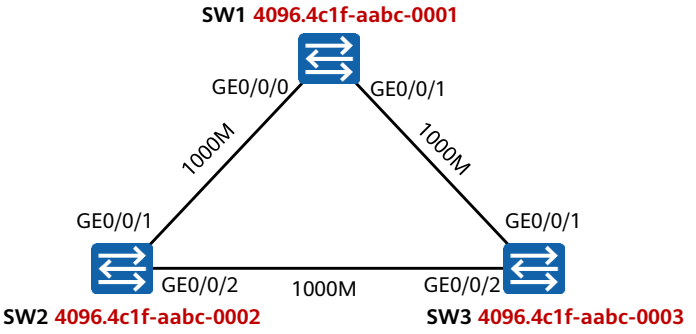


非指定接口被阻塞

- 一台交换机上，既不是根接口，又不是指定接口的接口被称为非指定接口。
- STP操作的最后一步是阻塞网络中的非指定接口。这一步完成后，网络中的二层环路就此消除。

- 什么是非指定端口（预备端口）？
- 在确定了根端口和指定端口之后，交换机上所有剩余的非根端口和非指定端口统称为预备端口。
- 阻塞非指定端口
  - STP会对这些非指定端口进行逻辑阻塞，即这些端口不能转发由终端计算机产生并发送的帧（用户数据帧）。
  - 一旦非指定端口被逻辑阻塞后，STP树（无环路工作拓扑）就生成了。
- 注意：
  - 非指定端口可以接收并处理BPDU。
  - 根端口和指定端口既可以接收和发送BPDU，也可以转发用户数据帧。

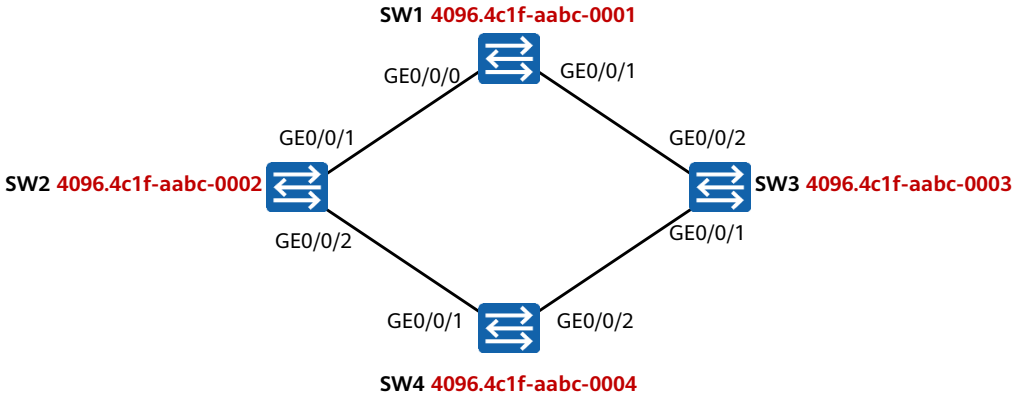
思考题1：识别以下拓扑中的根桥及各种接口角色



HUAWEI

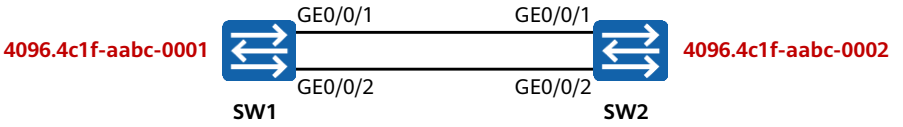
- 如图，首先选举根桥，三台交换机的桥优先级相同，则比较桥MAC地址，谁小谁优先，最终选举SW1为根桥；
- 其次选举根端口，SW2上GE0/0/1距离根桥最近，RPC最小，所以SW2的GE0/0/1为根端口，同理SW3的GE0/0/1也为根端口；
- 然后选举指定端口，SW1为根桥，所以SW1上的GE0/0/0和GE0/0/1端口为指定端口，SW2的GE0/0/2端口接收到SW3的配置BPDU，比较BID，SW2比SW3的BID更优，所以SW2的GE0/0/2端口为指定端口；
- 最终非根端口，非指定端口的SW3的GE0/0/2端口为预备端口。

思考题2：识别以下拓扑中的根桥及各种接口角色



- 如图，首先选举根桥，四台交换机的桥优先级相同，则比较桥MAC地址，谁小谁优先，最终选举SW1为根桥；
- 其次选举根端口，SW2上GE0/0/1距离根桥最近，RPC最小，所以SW2的GE0/0/1为根端口，同理SW3的GE0/0/2也为根端口，SW4的两个端口RPC相同，然后比较SW4的G0/0/1对应的交换机SW2的BID与G0/0/2对应的交换机SW3的BID，谁小谁优先，最终选举出SW4的GE0/0/1端口为根端口；
- 然后选举指定端口，SW1为根桥，所以SW1上的GE0/0/0和GE0/0/1端口为指定端口，SW2的GE0/0/2端口接收到SW4的配置BPDU，比较RPC，SW2比SW4的RPC更小，所以SW2的GE0/0/2端口为指定端口，同理可得SW3的GE0/0/1端口为指定端口；
- 最终非根端口，非指定端口的SW4的GE0/0/2端口为预备端口。

思考题3：识别以下拓扑中的根桥及各种接口角色



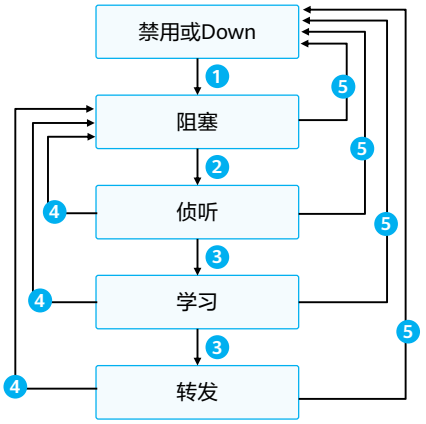
HUAWEI

- 如图，首先选举根桥，两台交换机的桥优先级相同，则比较桥MAC地址，谁小谁优先，最终选举SW1为根桥；
- 其次选举根端口，SW2上两个端口RPC相同，再比较两个接口对端的BID也相同，然后比较两个端口对端的PID，SW2的G0/0/1的对端PID：128.1，SW2的G0/0/2的对端PID：128.2，越小越优先，所以SW2的G0/0/1为根端口；
- 然后选举指定端口，SW1为根桥，所以SW1上的GE0/0/1和GE0/0/2端口为指定端口；
- 最终非根端口，非指定端口的SW2的GE0/0/2端口为预备端口。

STP的接口状态

状态名称	状态描述
禁用（Disable）	该接口不能收发BPDU，也不能收发业务数据帧，例如接口为down
阻塞（Blocking）	该接口被STP阻塞。处于阻塞状态的接口不能发送BPDU，但是会持续侦听BPDU，而且不能收发业务数据帧，也不会进行MAC地址学习
侦听（Listening）	当接口处于该状态时，表明STP初步认定该接口为根接口或指定接口，但接口依然处于STP计算的过程中，此时接口可以收发BPDU，但是不能收发业务数据帧，也不会进行MAC地址学习
学习（Learning）	当接口处于该状态时，会侦听业务数据帧（但是不能转发业务数据帧），并且在收到业务数据帧后进行MAC地址学习
转发（Forwarding）	处于该状态的接口可以正常地收发业务数据帧，也会进行BPDU处理。接口的角色需是根接口或指定接口才能进入转发状态

STP的接口状态迁移

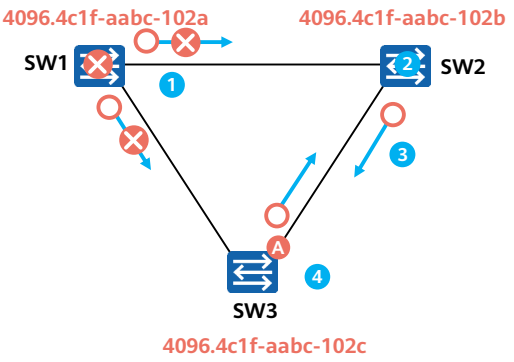


1. 接口初始化或激活，自动进入阻塞状态
2. 接口被选举为根接口或指定接口，自动进入侦听状态
3. 转发延迟计时器超时且接口依然为根接口或指定接口
4. 接口不再是根接口或指定接口或指定状态
5. 接口被禁用或者链路失效

HUAWEI

- 图中所示为STP的端口状态迁移机制，运行STP协议的设备上端口状态有5种：
  - Forwarding：转发状态。端口既可转发用户流量也可转发BPDU报文，只有根端口或指定端口才能进入Forwarding状态。
  - Learning：学习状态。端口可根据收到的用户流量构建MAC地址表，但不转发用户流量。增加Learning状态是为了防止临时环路。
  - Listening：侦听状态。端口可以转发BPDU报文，但不能转发用户流量。
  - Blocking：阻塞状态。端口仅仅能接收并处理BPDU，不能转发BPDU，也不能转发用户流量。此状态是预备端口的最终状态。
  - Disabled：禁用状态。端口既不处理和转发BPDU报文，也不转发用户流量。

拓扑变化 - 根桥故障



根桥故障恢复过程

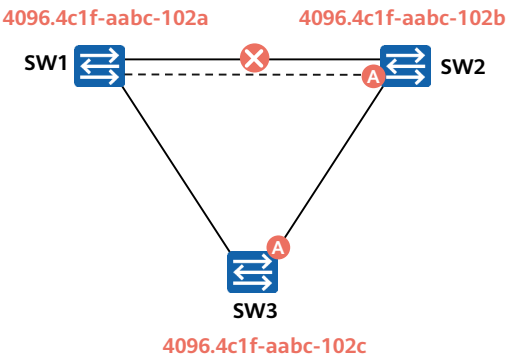
1. SW1根桥发生故障，停止发送BPDU报文。
2. SW2等待Max Age计时器（20 s）超时，从而导致已经收到的BPDU报文失效，又接收不到根桥发送的新的BPDU报文，从而得知上游出现故障。
3. 非根桥会互相发送配置BPDU，重新选举新的根桥。
4. 经过重新选举后，SW3的A端口经过两个Forward Delay（15 s）时间恢复转发状态。

- 非根桥会在BPDU老化之后开始根桥的重新选举。
- 根桥故障会导致50 s左右的恢复时间。

HUAWEI

- 根桥故障：
  - 在稳定的STP网络，非根桥会定期收到来自根桥的BPDU报文。
  - 如果根桥发生了故障，停止发送BPDU，下游交换机就无法收到来自根桥的BPDU报文。
  - 如果下游交换机一直收不到BPDU报文，Max Age计时器（缺省: 20s）就会超时，从而导致已经收到的BPDU报文失效，此时，非根桥会互相发送配置BPDU，重新选举新的根桥。
- 端口状态：
  - SW3的预备端口，20s后会从Blocking状态进入到Listening状态，再进入Learning状态，最终进入到Forwarding状态，进行用户流量的转发。
- 收敛时间：
  - 根桥故障会导致50s左右的恢复时间，等于Max Age加上2倍的Forward Delay收敛时间。

拓扑变化 - 物理链路故障



直连链路故障恢复过程

当交换机SW2网络稳定时检测到根端口的链路发生故障，则其备用端口会经过两倍的Forward Delay（15s）时间进入用户流量转发状态。

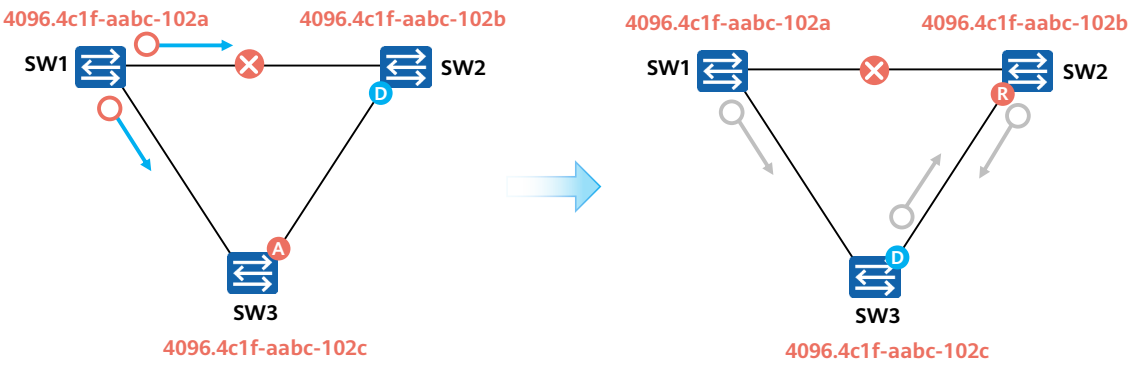
- SW2检测到直连链路物理故障后，会将预备端口转换为根端口。
- 直连链路故障，备用端口会经过30s后恢复转发状态。

HUAWEI

- 物理链路故障：
  - 当两台交换机间用两条链路互连时，其中一条是主用链路，另一条为备用链路。
  - 当网络稳定时，交换机SW2检测到根端口的链路发生故障（端口状态变为Down），则其备用端口会进入用户流量转发状态。
- 端口状态变化过程：
  - 备用端口会从Blocking状态，迁移到Listening-Learning-Forwarding状态。
  - 收敛时间：直连链路故障，备用端口会经过30s后恢复转发状态。

拓扑变化 - 非物理链路故障

- 非直连链路故障后，SW3的备用端口恢复到转发状态，非直连故障会导致50s左右的恢复时间。



HUAWEI

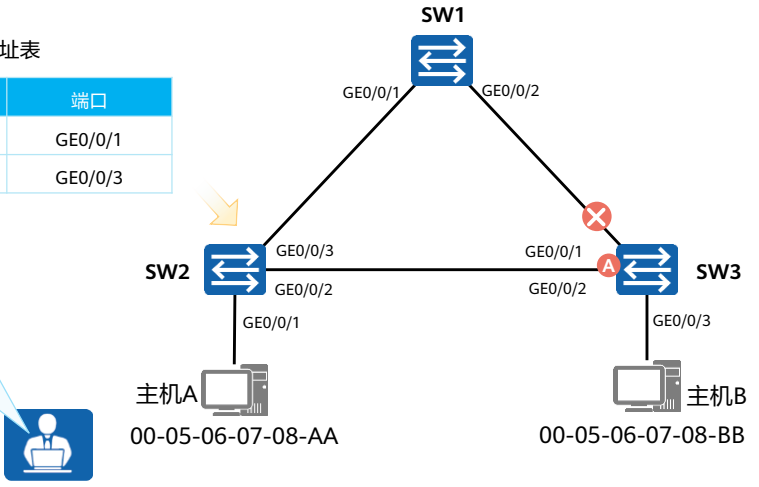
- 非物理链路故障
  - 在稳定的STP网络，非根桥会定期收到来自根桥的BPDU报文。
  - 若SW1与SW2之间的链路发生了某种故障（非物理故障），因此SW2一直收不到来自根桥SW1的BPDU报文，Max Age计时器（缺省: 20 s）就会超时，从而导致已经收到的BPDU报文失效。
  - 此时，非根桥SW2会认为根桥失效，并且认为自己是根桥，从而发送自己的配置BPDU给SW3，通知SW3自己是新的根桥。
  - SW3收到SW2发来的非最优的BPDU，会将从SW1收到的最优的BPDU转发给SW2。
  - 因此，SW2发现SW3发来的BPDU更优，就放弃宣称自己是根桥并重新确定端口角色。
- 端口状态：
  - SW3预备端口20s后会从Blocking状态进入到Listening状态，再进入Learning状态，最终进入到Forwarding状态，进行用户流量的转发。
- 收敛时间：
  - 非物理链路故障会导致50s左右的恢复时间，等于Max Age加上2倍的Forward Delay收敛时间。

拓扑改变导致MAC地址表错误（1）

MAC地址表

MAC	端口
00-05-06-07-08-AA	GE0/0/1
00-05-06-07-08-BB	GE0/0/3

如图，SW3的根端口发生故障，导致生成树拓扑重新收敛，在生成树拓扑完成收敛之后，从主机A到主机B的帧仍然不能到达目的地。这是因为交换机依赖MAC地址表转发数据帧，缺省情况下，MAC地址表项的老化时间是300秒。那么该怎么快速恢复转发?



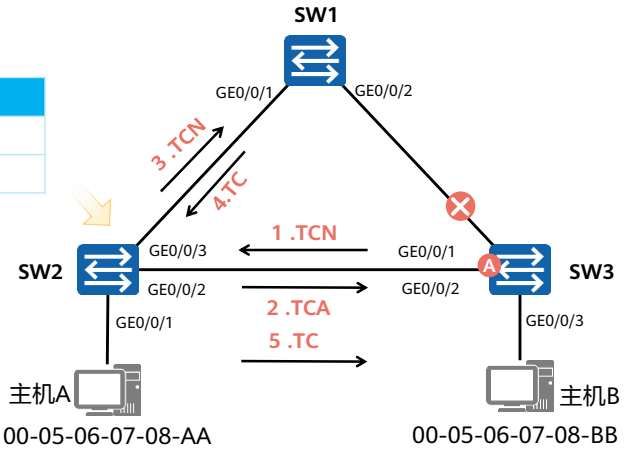
- 在交换网络中，交换机依赖MAC地址表转发数据帧。缺省情况下，MAC地址表项的老化时间是300秒。如果生成树拓扑发生变化，交换机转发数据的路径也会随着发生改变，此时MAC地址表中未及时老化掉的表项会导致数据转发错误，因此在拓扑发生变化后需要及时更新MAC地址表项。
- 本例中，SW2中的MAC地址表项定义了通过端口GigabitEthernet 0/0/1可以到达主机A，通过端口GigabitEthernet 0/0/3可以到达主机B。由于SW3的根端口产生故障，导致生成树拓扑重新收敛，在生成树拓扑完成收敛之后，从主机A到主机B的帧仍然不能到达目的地。这是因为MAC地址表项老化时间是300秒，主机A发往主机B的帧到达SW2后，SW2会继续通过端口GigabitEthernet 0/0/3转发该数据帧。

拓扑改变导致MAC地址表错误 (2)

MAC地址表

MAC	端口
00-05-06-07-08-AA	GE0/0/1
00-05-06-07-08-BB	GE0/0/2

- TCN BPDU在网络拓扑变化的时候产生。
- 报文格式：协议标识、版本号和类型。
- 拓扑变化：会使用到配置BPDU中Flags的TCA和TC位。



- 拓扑变化过程中，根桥通过TCN BPDU报文获知生成树拓扑里发生了故障。根桥生成TC用来通知其他交换机加速老化现有的MAC地址表项。
- 拓扑变更以及MAC地址表项更新的具体过程如下：
  - SW3感知到网络拓扑发生变化后，会不间断地向SW2发送TCN BPDU报文。
  - SW2收到SW3发来的TCN BPDU报文后，会把配置BPDU报文中的Flags的TCA位设置1，然后发送给SW3，告知SW3停止发送TCN BPDU报文。
  - SW2向根桥转发TCN BPDU报文。
  - SW1把配置BPDU报文中的Flags的TC位设置为1后发送，通知下游设备把MAC地址表项的老化时间由默认的300 s修改为Forward Delay的时间（默认为15 s）。
  - 最多等待15 s之后，SW2中的错误MAC地址表项会被自动清除。此后，SW2就能重新开始MAC表项的学习及转发操作。

目录

STP的基础配置命令（1）

1. 配置生成树工作模式

   [Huawei] **stp mode** { **stp** | **rstp** | **mstp** }

   交换机支持STP、RSTP和MSTP（Multiple Spanning Tree Protocol）三种生成树工作模式，默认情况工作在MSTP模式。

2. （可选）配置根桥

   [Huawei] **stp root primary**

   配置当前设备为根桥。缺省情况下，交换机不作为任何生成树的根桥。配置后该设备优先级数值自动为0，并且不能更改设备优先级。

3. （可选）备份根桥

   [Huawei] **stp root secondary**

   配置当前交换机为备份根桥。缺省情况下，交换机不作为任何生成树的备份根桥。配置后该设备优先级数值为4096，并且不能更改设备优先级。

STP的基础配置命令 (2)

1. （可选）配置交换机的STP优先级

```
[Huawei] stp priority priority
```

缺省情况下，交换机的优先级取值是32768。

2. （可选）配置接口路径开销

```
[Huawei] stp pathcost-standard { dot1d-1998 | dot1t | legacy }
```

配置接口路径开销计算方法。缺省情况下，路径开销值的计算方法为IEEE 802.1t（dot1t）标准方法。
同一网络内所有交换机的接口路径开销应使用相同的计算方法。

```
[Huawei-GigabitEthernet0/0/1] stp cost cost
```

设置当前接口的路径开销值。

STP的基础配置命令 (3)

1. （可选）配置接口优先级

   [Huawei-intf] **stp priority** *priority*

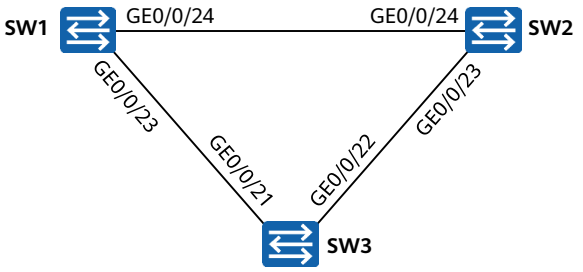
   配置接口的优先级。缺省情况下，交换机接口的优先级取值是128。

2. 启用STP/RSTP/MSTP

   [Huawei] **stp enable**

   使能交换机的STP/RSTP/MSTP功能。缺省情况下，设备的STP/RSTP/MSTP功能处于启用状态。

案例1：STP的基础配置



- 在上述三台交换机上部署STP，以便消除网络中的二层环路。
- 通过配置，将SW1指定为根桥，并使SW3的GE0/0/22接口被STP阻塞。

SW1的配置如下：

```
[SW1] stp mode stp
[SW1] stp enable
[SW1] stp priority 0
```

SW2的配置如下：

```
[SW2] stp mode stp
[SW2] stp enable
[SW2] stp priority 4096
```

SW3的配置如下：

```
[SW3] stp mode stp
[SW3] stp enable
```

案例1：STP的基础配置

在SW3上查看STP接口状态摘要：

<SW3> **display stp brief**

MSTID	Port	Role	STP State	Protection
0	GigabitEthernet0/0/21	ROOT	FORWARDING	NONE
0	GigabitEthernet0/0/22	ALTE	DISCARDING	NONE

目录

STP的不足之处

- STP协议虽然能够解决环路问题，但是由于网络拓扑收敛慢，影响了用户通信质量。如果网络中的拓扑结构频繁变化，网络也会随之频繁失去连通性，从而导致用户通信频繁中断，这是用户无法忍受的。
- STP没有细致区分接口状态和接口角色，不利于初学者学习及部署。
- 网络协议的优劣往往取决于协议是否对各种情况加以细致区分。
  - 从用户角度来讲，Listening、Learning和Blocking状态并没有区别，都同样不转发用户流量。
  - 从使用和配置角度来讲，接口之间最本质的区别并不在于接口状态，而是在于接口扮演的角色。
  - 根接口和指定接口可以都处于Listening状态，也可能都处于Forwarding状态。
- STP算法是被动的算法，依赖定时器等待的方式判断拓扑变化，收敛速度慢。
- STP算法要求在稳定的拓扑中，根桥主动发出配置BPDU报文，而其他设备进行处理，传遍整个STP网络。这也是导致拓扑收敛慢的主要原因之一。

RSTP概述

- IEEE 802.1w中定义的RSTP可以视为STP的改进版本，RSTP在许多方面对STP进行了优化，它的收敛速度更快，而且能够兼容STP。
- RSTP引入了新的接口角色，其中替代接口的引入使得交换机在根接口失效时，能够立即获得新的路径到达根桥。备份端口作为指定端口的备份，帮助链路上的网桥快速获得到根桥的备份路径。RSTP的状态规范根据端口是否转发用户流量和学习MAC地址把原来的5种状态缩减为3种。 另外，RSTP还引入了边缘接口的概念，这使得交换机连接终端设备的接口在初始化之后能够立即进入转发状态，提高了工作效率。

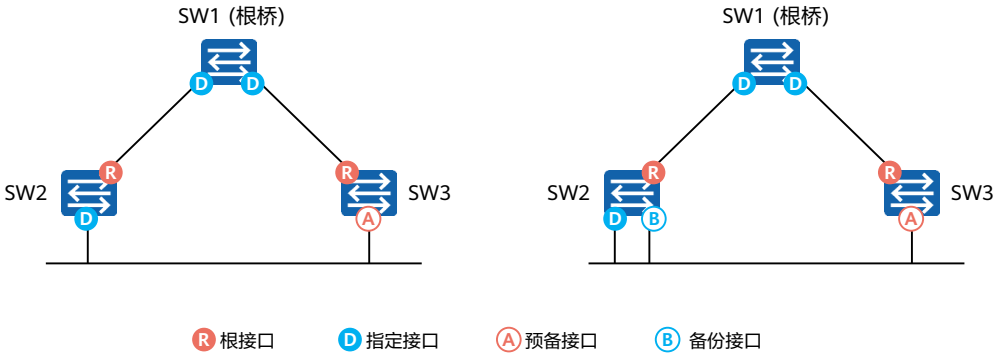
- IEEE于2001年发布的802.1w标准定义了快速生成树协议RSTP（Rapid Spanning-Tree Protocol），RSTP在STP基础上进行了改进，实现了网络拓扑快速收敛。
- RSTP（快速生成树）是从STP演化而来的，基本思想一样；当交换网络拓扑结构发生变化时， RSTP可以通过Proposal/Agreement机制更快地恢复网络的连通性。
- 根据STP的不足，RSTP删除了3种端口状态，新增加了2种端口角色，并且把端口属性充分的按照状态和角色解耦；此外，RSTP还增加了相应的一些增强特性和保护措施，实现网络的稳定和快速收敛。
- RSTP是可以与STP实现后向兼容的，但在实际中，并不推荐这样的做法，原因是RSTP会失去其快速收敛的优势，而STP慢速收敛的缺点会暴露出来。
- RSTP对STP的其他改进:
  - 配置BPDU的处理发生变化：
    - 拓扑稳定后，配置BPDU报文的发送方式进行了优化；
    - 使用更短的BPDU超时计时；
    - 对处理次等BPDU的方式进行了优化；
  - 配置BPDU格式的改变，充分利用了STP协议报文中的Flag字段，明确了接口角色。
  - RSTP拓扑变化处理：相比于STP进行了优化，加速针对拓扑变更的反应速度。

RSTP对STP的其他改进

- 配置BPDU的处理发生变化：
  - 拓扑稳定后，配置BPDU报文的发送方式进行了优化
  - 使用更短的BPDU超时计时
  - 对处理次等BPDU的方式进行了优化
- 配置BPDU格式的改变，充分利用了STP协议报文中的Flag字段，明确了接口角色
- RSTP拓扑变化处理：相比于STP进行了优化，加速针对拓扑变更的反应速度

端口角色不同

- 通过接口角色的增补，简化了生成树协议的理解及部署

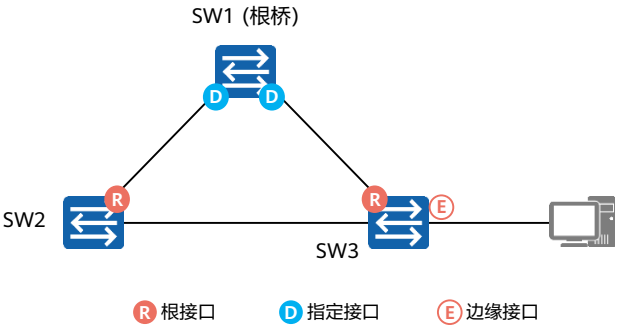


RSTP的接口角色共有4种：根接口、指定接口、预备接口和备份接口

- 从配置BPDU报文发送角度来看：
  - 预备（Alternate）接口就是由于学习到其它网桥发送的配置BPDU报文而阻塞的接口。
  - 备份（Backup）接口就是由于学习到自己发送的配置BPDU报文而阻塞的接口。
- 从用户流量角度来看：
  - Alternate接口提供了从指定桥到根的另一条可切换路径，作为根接口的备份接口。
  - Backup接口作为指定接口的备份，提供了另一条从根桥到相应网段的备份通路。

边缘端口

- 如果指定端口位于整个域的边缘，不再与任何交换设备连接，这种端口叫做边缘端口。



边缘端口一般与用户终端设备直接连接，可以由Disabled状态直接转到Forwarding状态。

HUAWEI

- 在STP中用户终端接入交换设备端口状态由Disabled状态转到Forwarding状态需要经过状态迁移的延迟时间，那么用户在这段时间无法上网，如果网络频繁变化，用户上网状态非常不稳定，时断时续。
- 边缘端口一般与用户终端设备直接连接，不与任何交换设备连接。边缘端口正常情况下接收不到配置BPDU报文，不参与RSTP运算，可以由Disabled状态直接转到Forwarding状态，且不经历时延，就像在端口上将STP禁用了一样。但是，一旦边缘端口收到配置BPDU报文，就丧失了边缘端口属性，成为普通STP端口，并重新进行生成树计算，从而引起网络震荡。

端口状态不同

- RSTP的状态规范把原来的5种状态缩减为3种。
  - 如果不转发用户流量也不学习MAC地址，那么接口状态就是Discarding状态。
  - 如果不转发用户流量但是学习MAC地址，那么接口状态就是Learning状态。
  - 如果既转发用户流量又学习MAC地址，那么接口状态就是Forwarding状态。

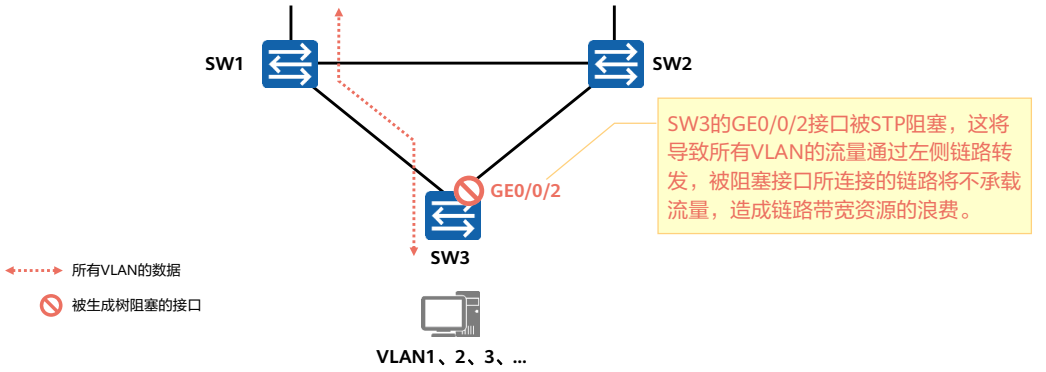
STP接口状态	RSTP接口状态	接口在拓扑中的角色
Forwarding	Forwarding	包括根接口、指定接口
Learning	Learning	包括根接口、指定接口
Listening	Discarding	包括根接口、指定接口
Blocking	Discarding	包括Alternate接口、Backup接口
Disabled	Discarding	包括Disable接口

- RSTP把原来STP的5种端口状态简化成了3种。
  - Discarding状态，端口既不转发用户流量也不学习MAC地址。
  - Learning状态，端口不转发用户流量但是学习MAC地址。
  - Forwarding状态，端口既转发用户流量又学习MAC地址。

目录

STP/RSTP的缺陷：所有的VLAN共享一棵生成树

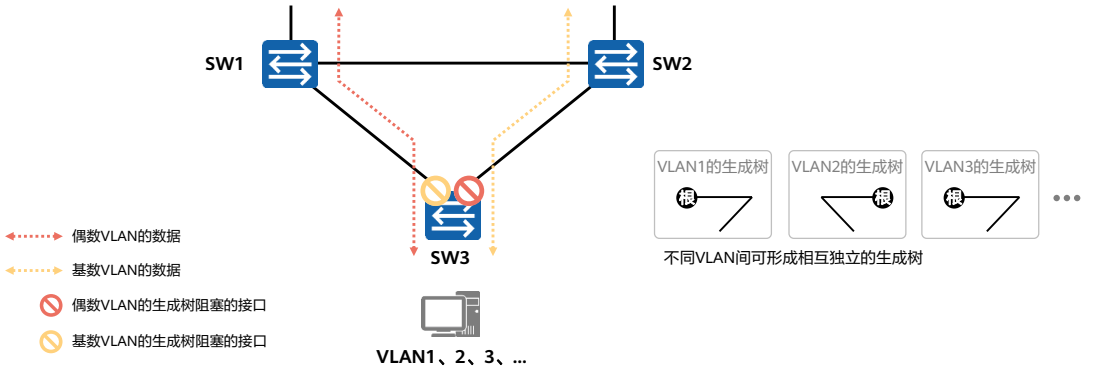
- RSTP在STP基础上进行了改进，实现了网络拓扑快速收敛。
- 但RSTP和STP还存在同一个缺陷：由于局域网内所有的VLAN共享一棵生成树，因此无法在VLAN间实现数据流量的负载均衡，链路被阻塞后将不承载任何流量，还有可能造成部分VLAN的报文无法转发。



HUAWEI

VBST：基于VLAN的生成树

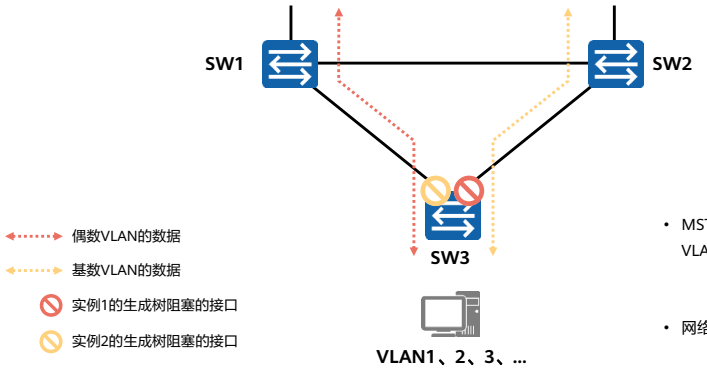
- 华为公司提出了VBST（VLAN-Based Spanning Tree）生成树解决方案。该解决方案中，生成树的形成是基于VLAN的，不同VLAN间可形成相互独立的生成树，不同VLAN内的流量沿着各自的生成树转发，进而可实现流量的负载分担。



- 企业网中部署VBST：
  - 可消除网络中可能存在的通信环路。
  - 可实现链路的复用和流量的负载分担，进而有效地提高链路带宽的利用率。
  - 配置和维护简单，进而可降低配置和维护成本。
- 但是如果网络中VLAN的数量较多，为每个VLAN执行独立的生成树计算将耗费交换机大量的资源。

MSTP：多生成树

- 为了弥补STP和RSTP的缺陷，IEEE于2002年发布的802.1s标准定义了MSTP。
- MSTP兼容STP和RSTP，既可以快速收敛，又提供了数据转发的多个冗余路径，在数据转发过程中实现VLAN数据的负载均衡。



- MSTP将VLAN映射到一个生成树的实例，若干个VLAN可共用一棵生成树。例如：
  - 将偶数VLAN映射到实例1
  - 将基数VLAN映射到实例2
- 网络中将只维护2棵生成树。

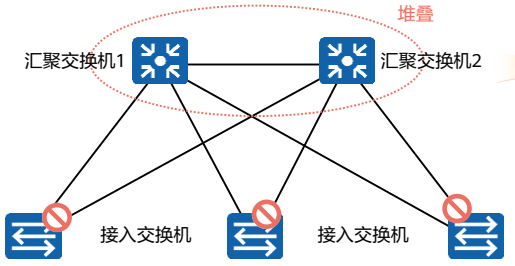
HUAWEI

MSTP概述

- MSTP把一个交换网络划分成多个域，每个域内形成多棵生成树，生成树之间彼此独立。
- 每棵生成树叫做一个多生成树实例MSTI（Multiple Spanning Tree Instance）。
- 所谓生成树实例就是多个VLAN的集合所对应的生成树。
- 通过将多个VLAN捆绑到一个实例，可以节省通信开销和资源占用率。
- MSTP各个实例拓扑的计算相互独立，在这些实例上可以实现负载均衡。
- 可以把多个相同拓扑结构的VLAN映射到一个实例里，这些VLAN在接口上的转发状态取决于接口在对应实例的状态。

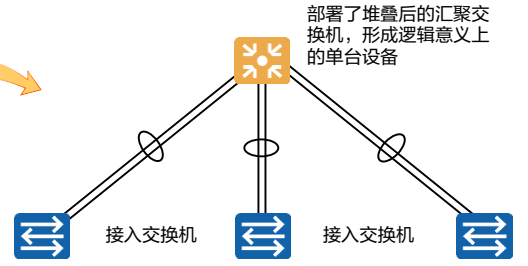
堆叠与园区网络树形结构组网形态

传统STP组网



两台汇聚交换机分别与接入层交换机构成了三角形的二层环路，网络中不得不部署STP，而STP将阻塞网络中的接口，造成链路带宽无法充分利用。

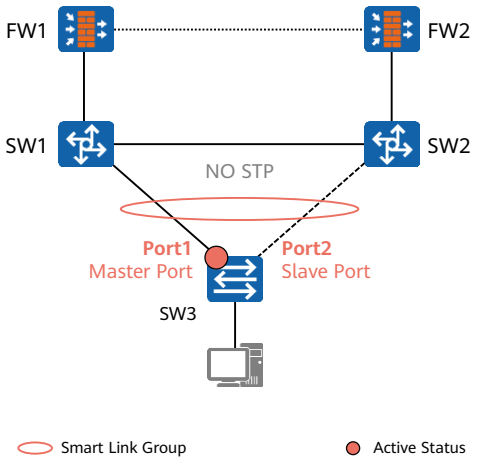
交换机堆叠组网



将汇聚交换机部署堆叠，形成逻辑意义上的单台设备，从而简化网络拓扑，此外，通过在汇聚交换机与接入交换机之间部署链路聚合，可将拓扑进一步简化为“树形结构”，消除二层环路，同时充分提高链路带宽利用率

- 堆叠iStack（Intelligent Stack），是指将多台支持堆叠特性的交换机设备组合在一起，从逻辑上组合成一台整体交换设备。
- 堆叠系统建立之前，每台交换机都是单独的实体，有自己独立的IP地址和MAC地址，对外体现为多台交换机，用户需要独立的管理所有的交换机；堆叠建立后堆叠成员对外体现为一个统一的逻辑实体，用户使用一个IP地址对堆叠中的所有交换机进行管理和维护，如图所示。通过交换机堆叠，可以实现网络大数据量转发和网络高可靠性，同时简化网络管理。

Smart Link



- Smart Link是一种为双上行组网量身定做的解决方案：
  - 在双向行的设备上部署，当网络正常时，两条上行链路中，一条处于活跃状态，而另一条则处于备份状态（不承载业务流量）。如此一来二层环路就此打破。
  - 当主用链路发生故障后，流量会在毫秒级的时间内迅速切换到备用链路上，保证了数据的正常转发。
  - Smart Link配置简单，便于用户操作。
  - 无需协议报文交互，收敛速度及可靠性大大提升。

- 如图所示Switch3采用双上行方式分别连接到FW1和FW2，这样Switch3到达上行的链路就可以有两条。在Switch3上配置Smart Link，正常情况下，可实现Port2所在链路作为Port1所在链路的备份。若Port1所在的链路发生故障，Smart Link会自动将数据流量切换到Port2所在链路，保证业务不中断。

思考题

1. （单选）以下关于STP接口状态的说法，错误的是（　）。

A.被阻塞的接口不会侦听，也不发送BPDU。

B.处于Learning状态的接口会学习MAC地址，但是不会转发数据。

C.处于Listening状态的接口会持续侦听BPDU。

D.被阻塞的接口如果一定时间内收不到BPDU，则会自动切换到Listening状态。

1. A

本章总结

- 生成树是一个用于局域网中消除环路的协议。运行该协议的设备通过彼此交互信息而发现网络中的环路，并对某些接口进行阻塞以消除环路。由于局域网规模的不断增长，生成树协议已经成为了当前最重要的局域网协议之一。
- 在以太网交换网中部署生成树协议后，如果网络中出现环路，生成树协议通过拓扑计算，可实现：
  - 消除环路：通过阻塞冗余链路消除网络中可能存在的网络通信环路。
  - 链路备份：当前活动的路径发生故障时，激活冗余备份链路，恢复网络连通性。
- STP（Spanning-Tree Protocol）作为一种存在已久的协议，已经无法满足现代园区网络的需求，但是了解STP的工作原理，有助于为进一步熟悉并掌握RSTP及MSTP的原理与部署做好铺垫。


谢 谢
www.huawei.com

实现VLAN间通信

前言

- 传统交换二层组网中，默认所有网络都处于同一个广播域，这带了诸多问题。VLAN（Virtual Local Area Network，虚拟局域网）技术的提出，满足了二层组网隔离广播域需求，使得属于不同VLAN的网络无法互访，但不同VLAN之间又存在着相互访问的需求。
- 本章主要描述如何实现不同VLAN之间的相互通信。

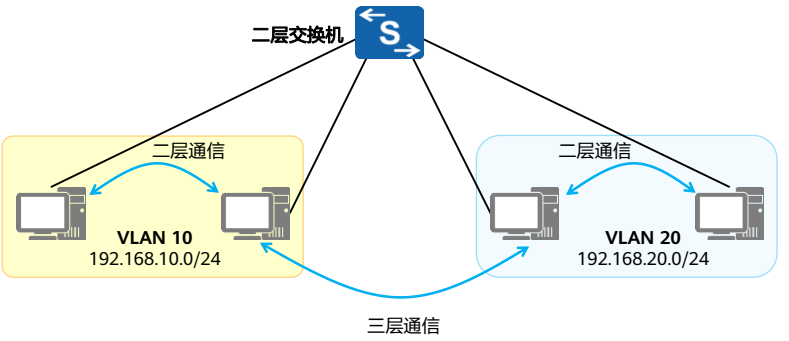
目标

- 学完本课程后，您将能够：
  - 了解如何实现VLAN间通信
  - 掌握如何使用路由器（物理接口、子接口）实现VLAN间通信
  - 掌握如何使用三层交换机实现VLAN间通信
  - 掌握报文三层转发过程

目录

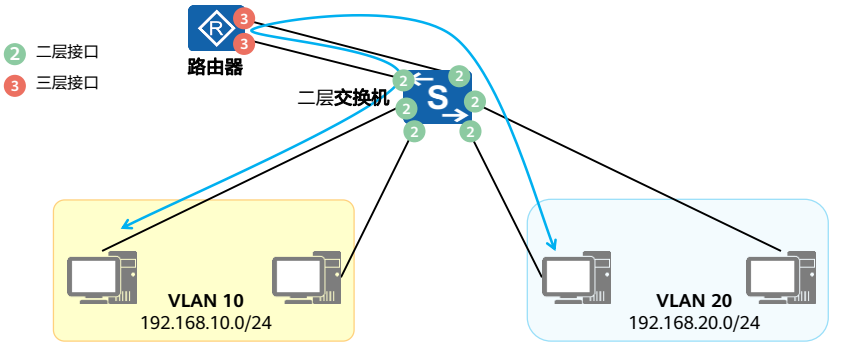
VLAN间通信（1）

- 实际网络部署中一般会将不同IP地址段划分到不同的VLAN。
- 同VLAN且同网段的PC之间可直接进行通信，无需借助三层转发设备，该通信方式被称为二层通信。
- VLAN之间需要通过三层通信实现互访，三层通信需借助三层设备。



VLAN间通信 (2)

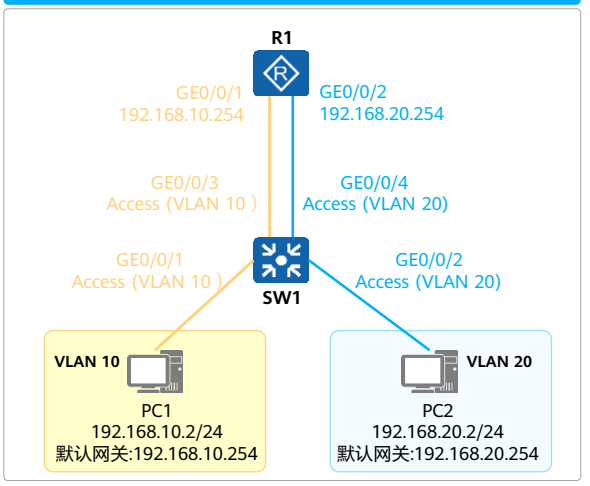
- 常见的三层设备：路由器、三层交换机、防火墙等。
- 将二层交换机与路由器的三层接口互联，由三层设备进行路由转发来实现通信。



目录

使用路由器物理接口

物理连接图

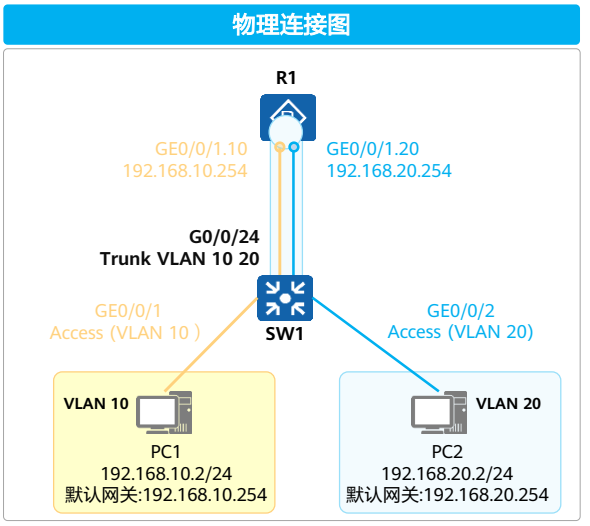


- 路由器三层接口作为网关，转发本网段前往其它网段的流量。
- 路由器三层接口无法处理携带VLAN Tag的数据帧，因此交换机上联路由器的接口需配置为Access。
- 路由器的一个物理接口作为一个VLAN的网关，因此存在一个VLAN就需要占用一个路由器物理接口。
- 路由器作为三层转发设备其接口数量较少，方案的可扩展性太差。

HUAWEI

- 在二层交换机上配置VLAN，每个VLAN单独使用一个交换机接口与路由器互联。
- 路由器使用两个物理接口，分别作为VLAN 10及VLAN 20内PC的默认网关，使用路由器的物理接口实现VLAN之间的通信。

使用路由器子接口

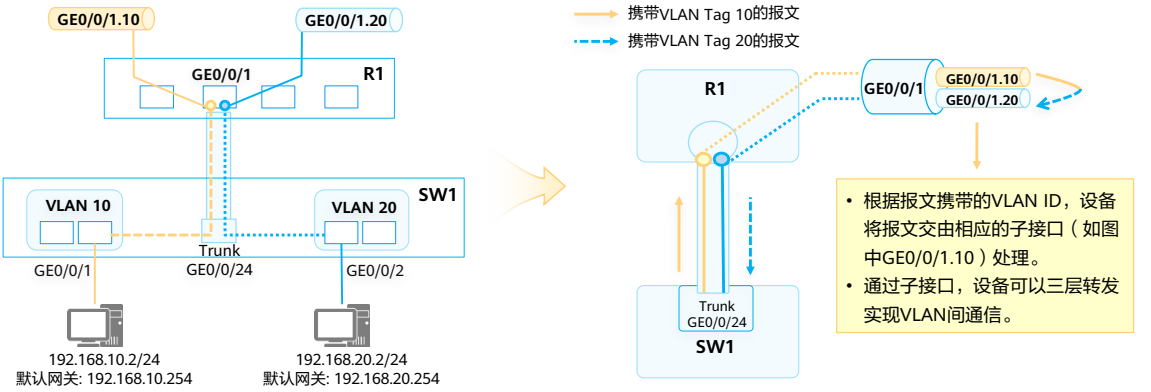


- 子接口（Sub-Interface）是基于路由器以太网接口所创建的逻辑接口，以物理接口ID+子接口ID进行标识，子接口同物理接口一样可进行三层转发。
- 子接口不同于物理接口，可以终结携带VLAN Tag的数据帧。
- 基于一个物理接口创建多个子接口，将该物理接口对接到交换机的Trunk接口，即可实现使用一个物理接口为多个VLAN提供三层转发服务。

- R1使用一个物理接口（GE0/0/1）与交换机SW1对接，并基于该物理接口创建两个子接口：GE0/0/1.10及GE0/0/1.20，分别使用这两个子接口作为VLAN 10及VLAN 20的默认网关。
- 由于三层子接口不支持VLAN报文，当它收到VLAN报文时，会将VLAN报文当成是非法报文而丢弃。因此，需要在子接口上将VLAN Tag剥掉，也就是需要VLAN终结（VLAN Termination）。

子接口处理流程

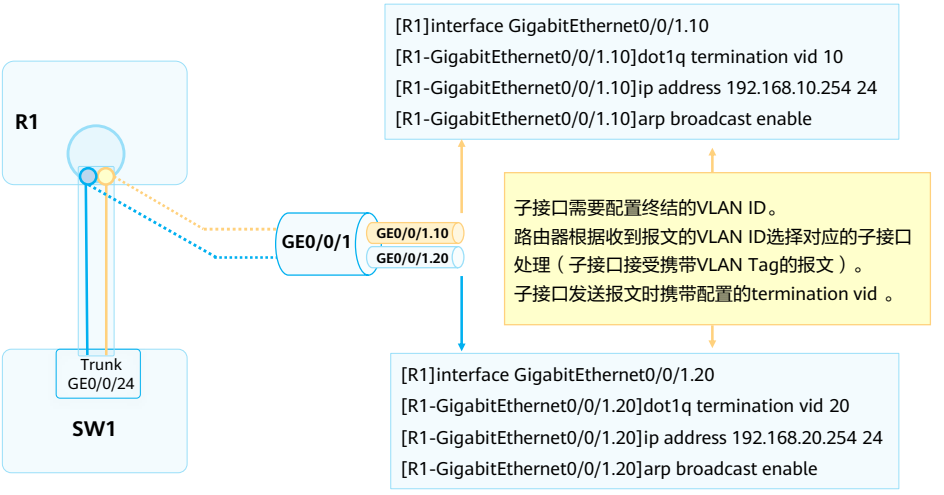
- 交换机连接路由器的接口类型配置为Trunk，根据报文的VLAN Tag不同，路由器将收到的报文交由对应的子接口处理。



  HUAWEI

- 子接口终结VLAN的实质包含两个方面：
  - 对接口接收到报文，剥除VLAN标签后进行三层转发或其他处理。
  - 对接口发出的报文，又将相应的VLAN标签添加到报文中后再发送。

子接口配置示例

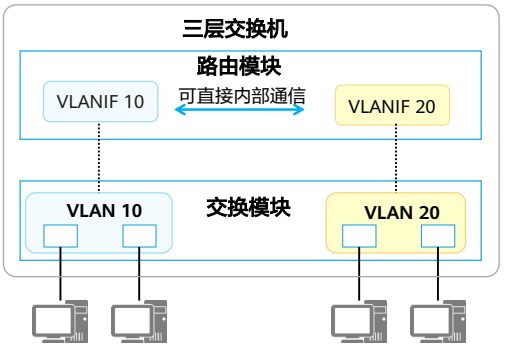


  HUAWEI

- **interface** *interface-type interface-number.sub-interface number*命令用来创建子接口。sub-interface number代表物理接口内的逻辑接口通道。一般情况下，为了方便记忆，子接口ID与所要终结的VLAN ID相同。
- **dot1q termination vid**命令用来配置子接口Dot1q终结的单层VLAN ID。缺省情况，子接口没有配置dot1q终结的单层VLAN ID。**arp broadcast enable**命令用来使能终结子接口的ARP广播功能。缺省情况下，终结子接口没有使能ARP广播功能。终结子接口不能转发广播报文，在收到广播报文后它们直接把该报文丢弃。为了允许终结子接口能转发广播报文，可以通过在子接口上执行此命令。

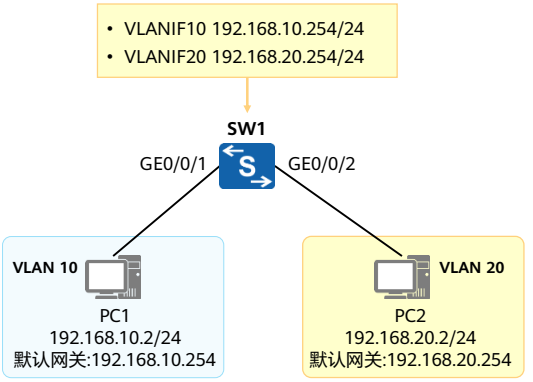
目录

三层交换机和VLANIF接口



- 二层交换机（Layer 2 Switch）指的是只具备二层交换功能的交换机。
- 三层交换机（Layer 3 Switch）除了具备二层交换机的功能，还支持通过三层接口（如VLANIF接口）实现路由转发功能。
- VLANIF接口是一种三层的逻辑接口，支持VLAN Tag的剥离和添加，因此可以通过VLANIF接口实现VLAN之间的通信。
- VLANIF接口编号与所对应的VLAN ID相同，如VLAN 10对应VLANIF 10。

VLANIF配置示例



基础配置：

```
[SW1]vlan batch 10 20
[SW1] interface GigabitEthernet 0/0/1
[SW1-GigabitEthernet0/0/1] port link-type access
[SW1-GigabitEthernet0/0/1] port default vlan 10
[SW1] interface GigabitEthernet 0/0/2
[SW1-GigabitEthernet0/0/2] port link-type access
[SW1-GigabitEthernet0/0/2] port default vlan 20
```

配置VLANIF：

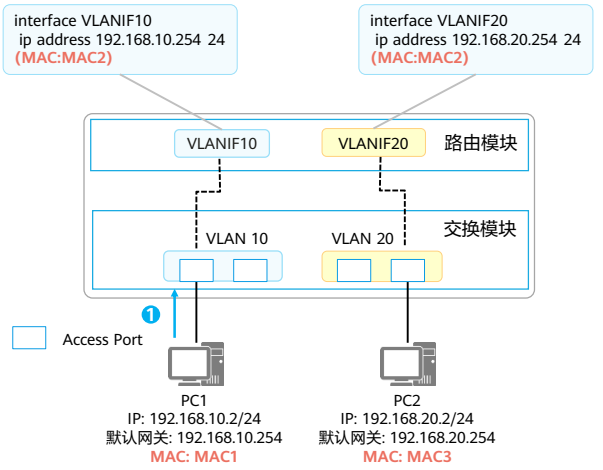
```
[SW1]interface Vlanif 10
[SW1-Vlanif10]ip address 192.168.10.254 24
[SW1]interface Vlanif 20
[SW1-Vlanif20]ip address 192.168.20.254 24
```

- 配置需求:
  两台PC分别属于VLAN 10、VLAN 20。通过三层交换机完成两台PC之间的相互通信。

HUAWEI

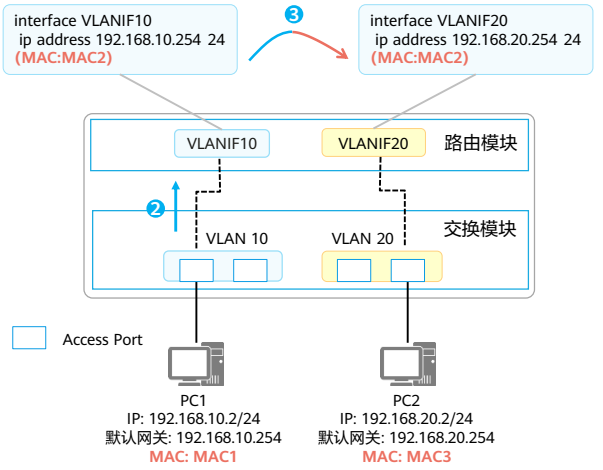
- **interface vlanif** *vlan-id*命令用来创建VLANIF接口并进入到VLANIF接口视图。*vlan-id*表示与VLANIF接口相关联的VLAN编号。VLANIF接口的IP地址作为主机的网关IP地址，和主机的IP地址必须位于同一网段。

VLANIF转发流程（1）



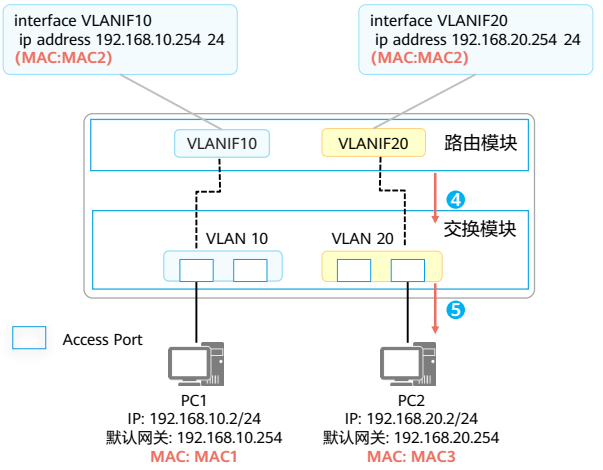
- 假设PC、三层交换机上都已存在相应的ARP或MAC表项。
- PC1与PC2之间通信过程如下：
- PC1通过本地IP地址、本地掩码、对端IP地址进行计算，发现目的设备PC2与自身不在同一个网段，判断该通信为三层通信，将去往PC2的流量发给网关。PC1发送的数据帧：源MAC = MAC1，目的MAC = MAC2。

VLANIF转发流程 (2)



- 交换机收到PC1发送的去往PC2的报文，经解封装发现目的MAC为VLANIF10接口的MAC地址，所以将报文交给路由模块继续处理。
- 路由模块解析发现目的IP为192.168.20.2，非本地接口存在的IP地址，因此需要对该报文三层转发。查找路由表后，匹配中VLANIF20产生的直连路由。

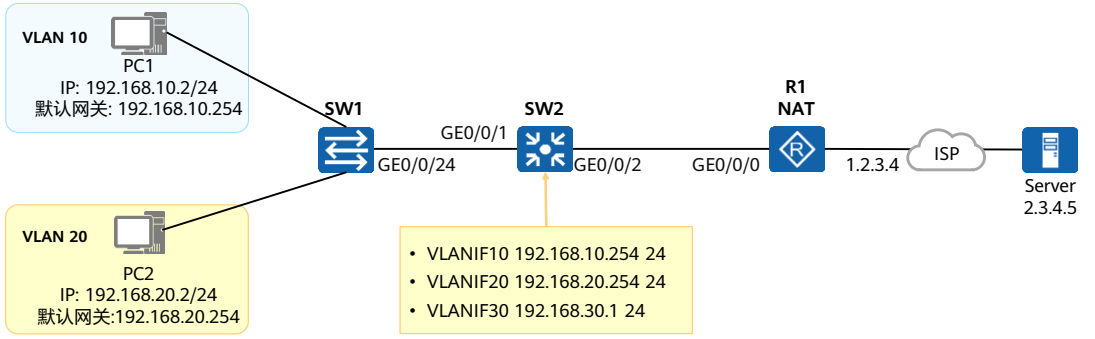
VLANIF转发流程 (3)



- 因为匹配的为直连路由，说明已经到达最后一跳，所以交换机在ARP表中查找192.168.20.2，获取192.168.20.2的MAC地址，交由交换模块重新封装为数据帧。
- 交换模块查找MAC地址表以明确报文出接口、是否需要携带VLAN Tag。最终交换模块发送的数据帧：源MAC = MAC2，目的MAC = MAC3，VLAN Tag = None。

目录

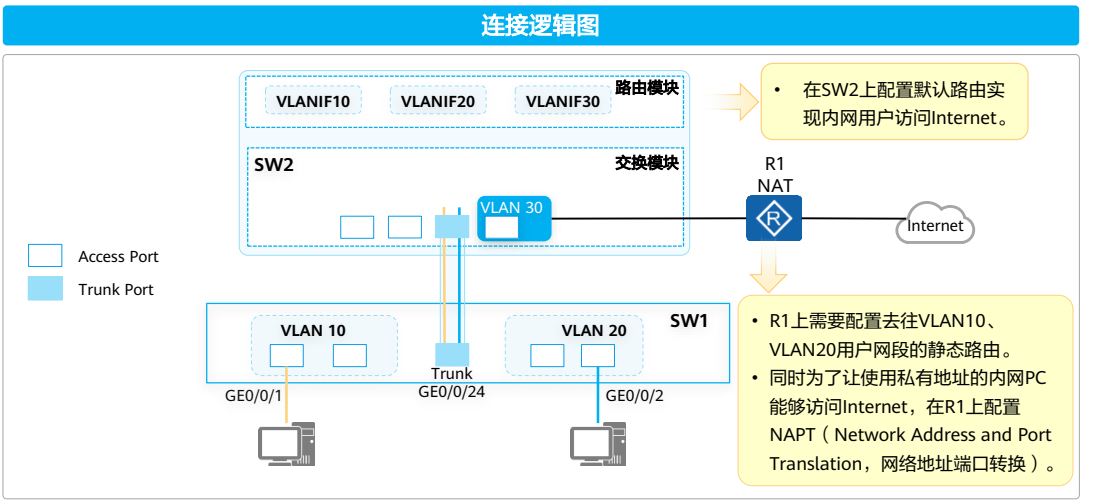
网络拓扑



以该拓扑为例，讲解VLAN 10内PC1到Internet上的服务器（2.3.4.5）的通信过程。

HUAWEI

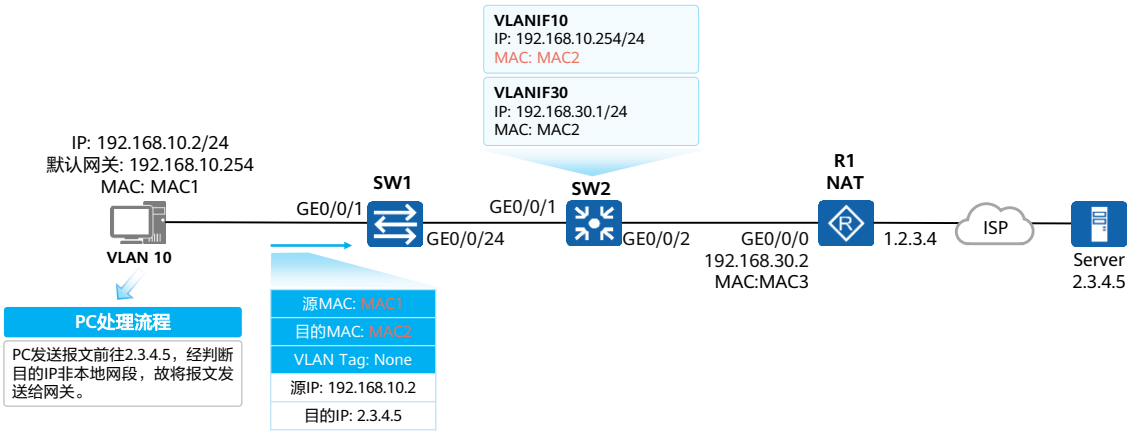
连接逻辑图



  HUAWEI

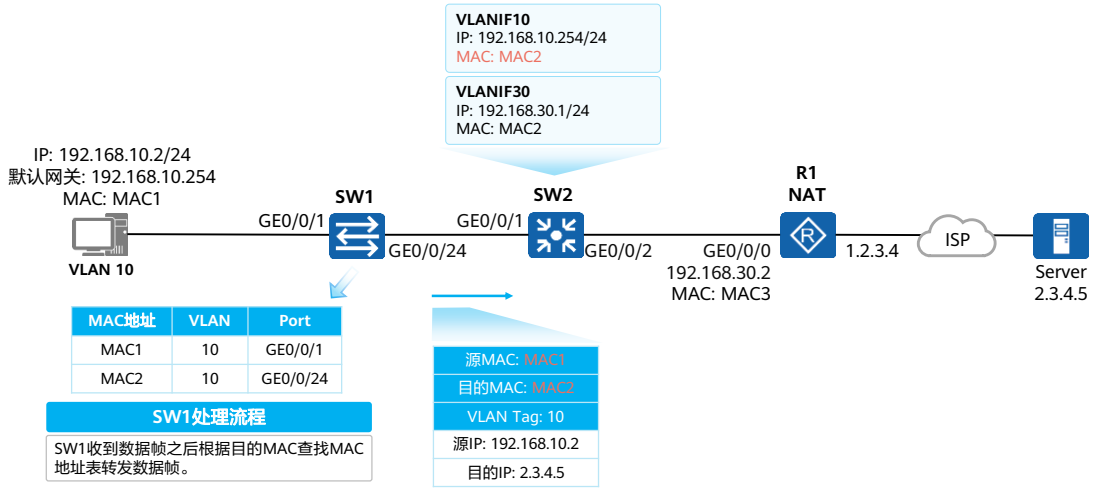
- NAPT（Network Address Port Translation，网络地址端口转换）：将IP数据报文头中的IP地址、端口号转换为另一个IP地址、端口号的过程，主要用于实现内部网络（私有IP地址）访问外部网络（公有IP地址）的功能，NAPT支持多个内部地址映射到同一个公有地址上，可以实现使用一个公有地址支持内网多个内部地址同时访问外部网络。

通信过程（1）



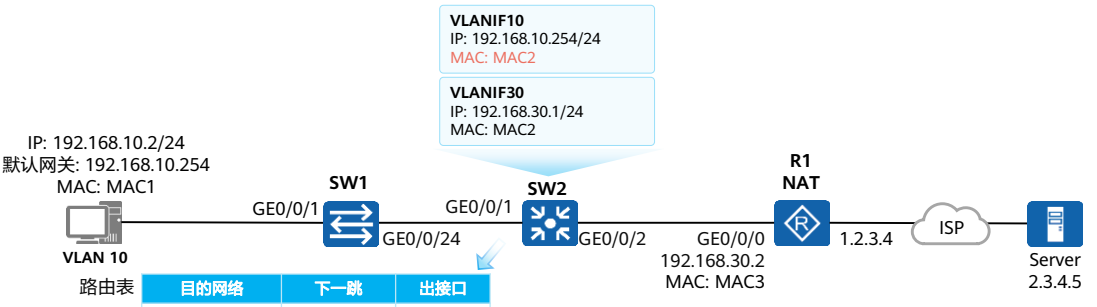
- 假设所有设备上都已存在相应的ARP或MAC表项。

通信过程 (2)



HUAWEI

通信过程 (3)



路由表

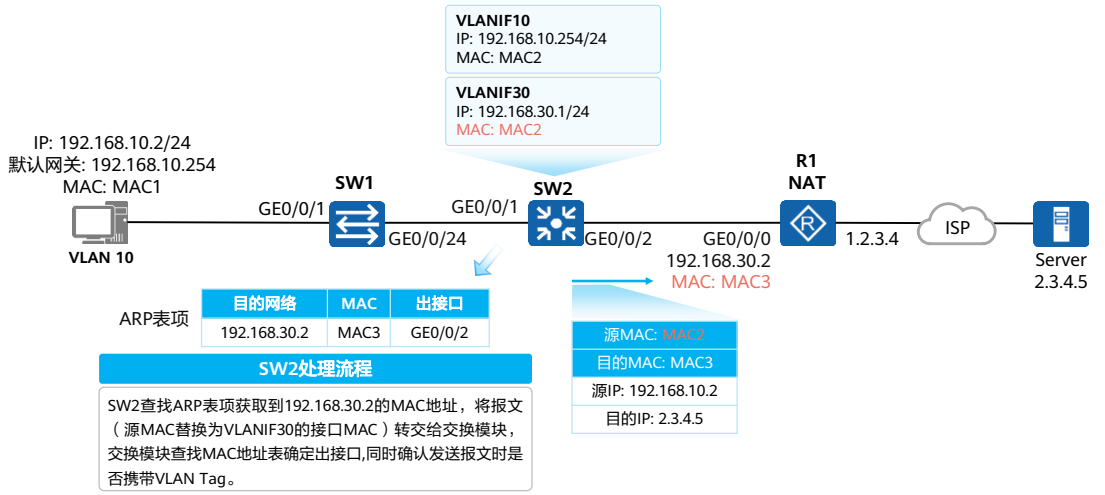
目的网络	下一跳	出接口
0.0.0.0/0	192.168.30.2	VLANIF30

SW2处理流程

SW2收到数据帧之后，查看目的MAC为自身接口VLANIF10的MAC地址，交由路由模块在路由表中查找2.3.4.5。

SW2路由查找结果为匹配缺省路由，出接口为VLANIF30、下一跳为192.168.30.2，查找ARP获取192.168.30.2的MAC地址。

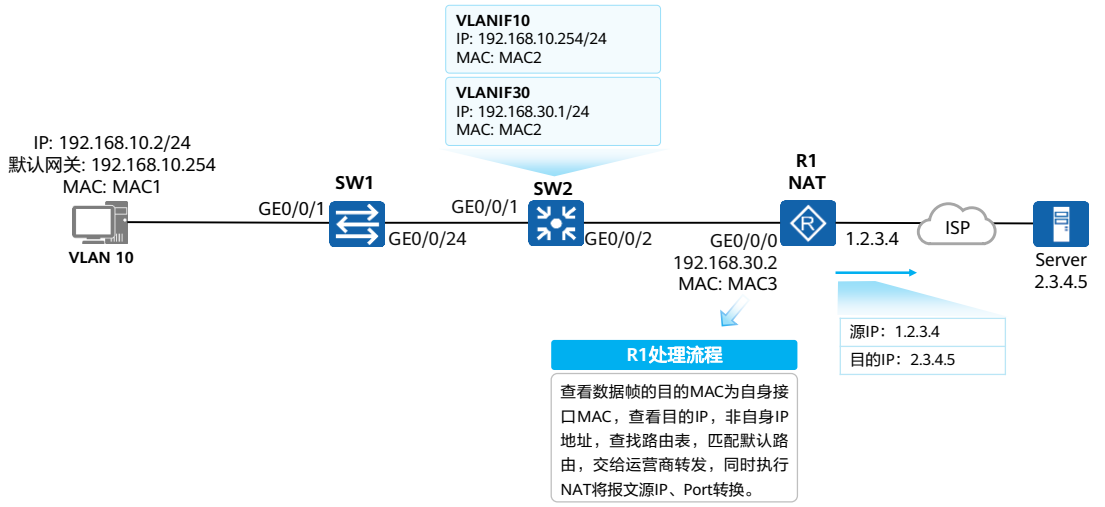
通信过程 (4)



SW2处理流程

SW2查找ARP表项获取到192.168.30.2的MAC地址，将报文（源MAC替换为VLANIF30的接口MAC）转交给交换模块，交换模块查找MAC地址表确定出接口,同时确认发送报文时是否携带VLAN Tag。

通信过程 (5)



- NAT（Network Address Translation，网络地址转换）：将IP数据报文头中的IP地址转换为另一个IP地址。

思考题

1. 通过子接口实现VLAN间通信时，交换机连接路由器的接口需要做哪些配置?
2. 报文经过三层转发时，报文内容有哪些变化?

1. 将接口配置为Trunk或者Hybrid，放通终端对应的VLAN（携带VLAN-Tag）。
2. 源目IP在转发过程中保持不变（无NAT场景），但是源目MAC会改变，三层转发时每经过一个三层设备进行三层转发，源目MAC都会发生变化。

本章总结

- 本章介绍了三种实现VLAN间通信的方式：通过路由器实现、通过子接口实现、通过VLANIF实现。
- 本章还详细介绍了三层交换机的通信过程，在通信过程中的设备处理机制、报文头部的变化。

更多信息

- 二、三层接口对比

二层接口（Layer2 Interface）	三层接口（Layer3 Interface）
二层接口不能配置IP地址	三层接口可以配置IP地址
二层接口不具备MAC地址	三层接口具备MAC地址
当二层接口收到数据帧时，设备在其MAC地址表中查询该帧的目的MAC地址，找到匹配的MAC地址表项后按照该表项的指示转发帧；如果没有找到匹配的MAC地址表项，则将帧进行泛洪。	三层接口收到数据帧后，如果数据帧的目的MAC地址与设备的本地MAC地址相同，则将数据帧解除封装，然后在路由表中查询数据包的目的IP地址，找到匹配的路由表项后按照该表项的指示转发包；如果没有找到匹配的表项，则将包丢弃。
典型的二层接口如二层交换机（只具备二层交换能力的交换机）的物理接口；大部分三层交换机（同时具备二层及三层交换能力的交换机）的物理接口缺省为二层接口。	典型的三层接口如路由器的三层接口。 某些三层交换机的物理接口可以切换成三层模式。 此外除了物理三层接口，还存在逻辑三层接口，例如交换机的VLANIF，或者网络设备上的逻辑子接口，如GE0/0/1.10。
二层接口并不隔离广播域，当二层接口收到广播帧时，会将数据帧进行泛洪。	三层接口隔离广播域，当三层接口收到广播帧时，缺省不会进行泛洪，而是直接终结。


谢 谢
www.huawei.com

以太网链路聚合与交换机堆叠、集群

前言

- 随着业务的发展和园区网络规模的不断扩大，用户对于网络的带宽、可靠性要求越来越高。传统解决方案通过升级设备方式提高网络带宽，同时通过部署冗余链路并辅以STP（Spanning Tree Protocol，生成树协议）协议实现高可靠。传统解决方案存在灵活度低、故障恢复时间长、配置复杂等缺点。
- 本章节将介绍通过链路聚合技术与堆叠、集群技术实现网络带宽提升与高可靠性保障。

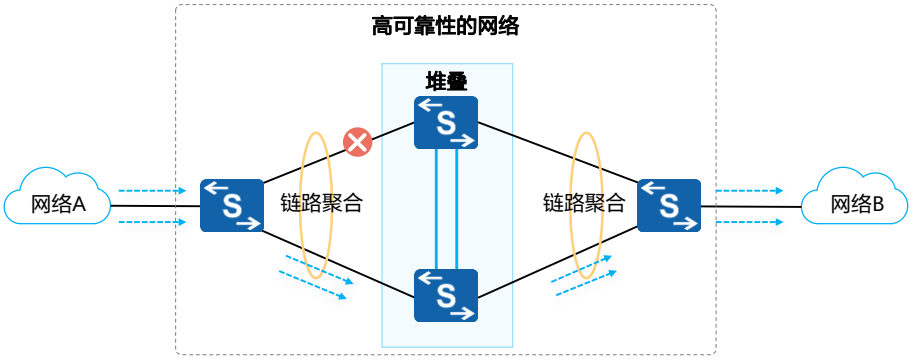
目标

- 学完本课程后，您将能够：
  - 了解链路聚合的作用
  - 了解链路聚合的分类
  - 理解LACP模式的链路聚合协商过程
  - 了解iStack和CSS的优点与原理
  - 了解链路聚合与堆叠技术常见应用与组网

目录

网络的可靠性

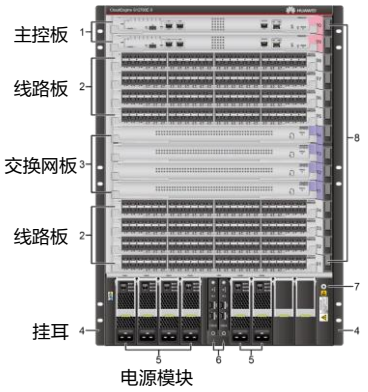
- 网络的可靠性指当设备或者链路出现单点或者多点故障时保证网络服务不间断的能力。
- 网络的可靠性可以从单板、设备、链路多个层面实现。



HUAWEI

- 随着网络的快速普及和应用的日益深入，各种增值业务得到了广泛部署，网络中断可能导致大量业务异常、造成重大经济损失。因此，作为承载业务主体的基础网络，其可靠性成为备受关注的焦点。

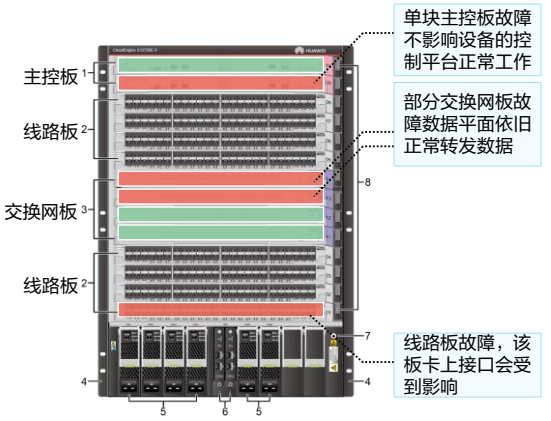
单板可靠性 (1)



S12700E-8机框正面结构

- 框式交换机由机框、电源模块、风扇模块、主控板、交换网板（SFU）、线路板（LPU）构成。
- 机框：为各种板卡、模块提供插槽，实现板卡间的通信。
- 电源模块：设备的供电系统
- 风扇模块：设备的散热系统
- 主控板（MPU，Main Processing Unit）：负责整个系统的控制平面和管理平面。
- 交换网板（SFU，Switch Fabric Unit）：负责整个系统的数据平面。数据平面提供高速无阻塞数据通道，实现各个业务模块之间的业务交换功能。
- 线路板（LPU，Line Processing Unit）：线路处理单元是物理设备上用于提供数据转发功能的模块，提供不同速率的光口、电口。

单板可靠性 (2)

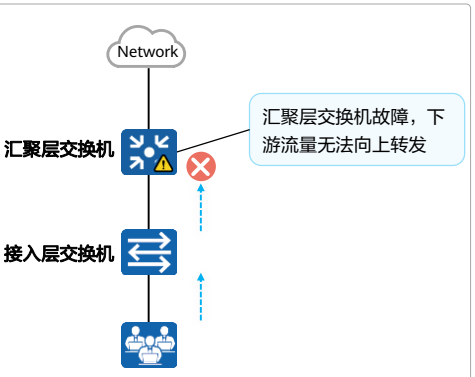


S12700E-8机框正面结构

- 以S12700E-8为例，设备提供8个线路板槽位、4个交换网板槽位、2个主控板槽位、6个电源模块槽位、4个风扇模块槽位。
- 框式交换机配置多个主控板、交换网板可保证设备自身的可靠性，单个槽位的交换网板、主控板损坏不影响设备的正常运行。
- 框式交换机的线路板损坏后，该板卡上的接口无法正常转发数据。

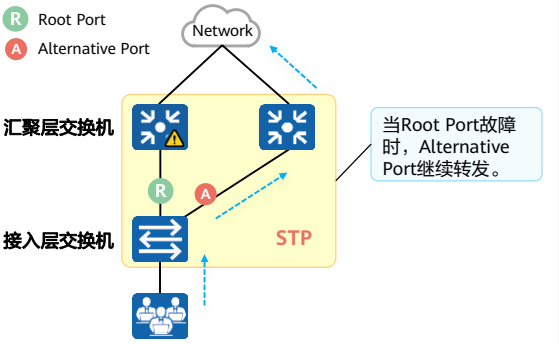
设备可靠性

无备份



设备无冗余设计的网络中，下游交换机采用单上行接入，上行交换机的接口故障或设备故障会导致下游网络全部中断。

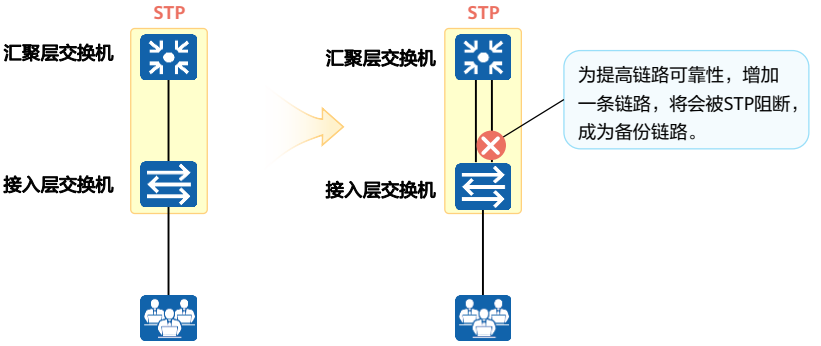
主备备份



设备冗余设计的网络中，下游交换机双上行接入，采用链路一主一备的方式，主链路上行接口、设备故障可以切换到备份链路，通过备份设备转发。

HUAWEI

链路可靠性



- 为保证设备间链路可靠性，在设备间部署多条物理线路，为防止环路STP只保留一条链路转发流量，其余链路成为备份链路。

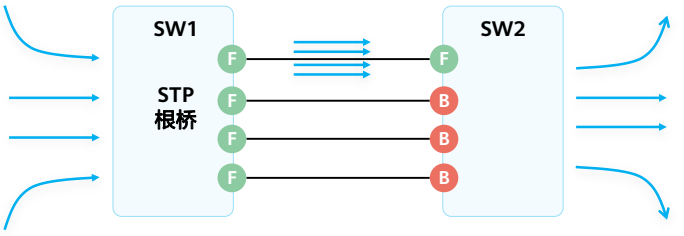
目录

提升链路带宽

- 设备之间存在多条链路时，由于STP的存在，实际只会有一条链路转发流量，设备间链路带宽无法得到提升。

F 转发流量的接口

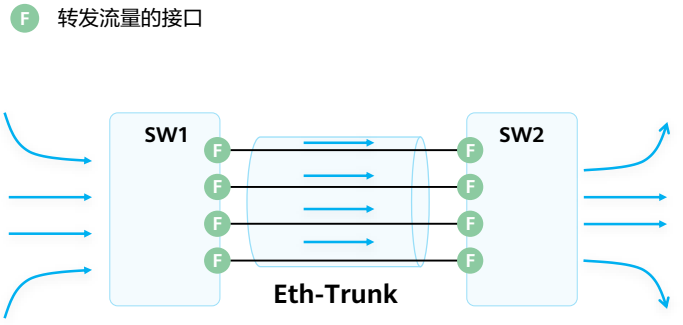
B STP阻塞端口，不转发流量



HUAWEI

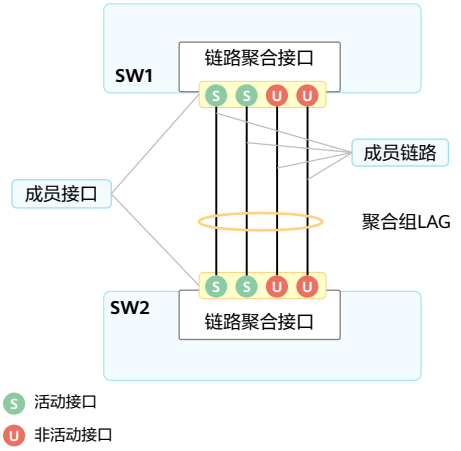
以太网链路聚合

- 以太网链路聚合Eth-Trunk：简称链路聚合，通过将多个物理接口捆绑成为一个逻辑接口，可以在不进行硬件升级的条件下，达到增加链路带宽的目的。



HUAWEI

链路聚合基本术语/概念



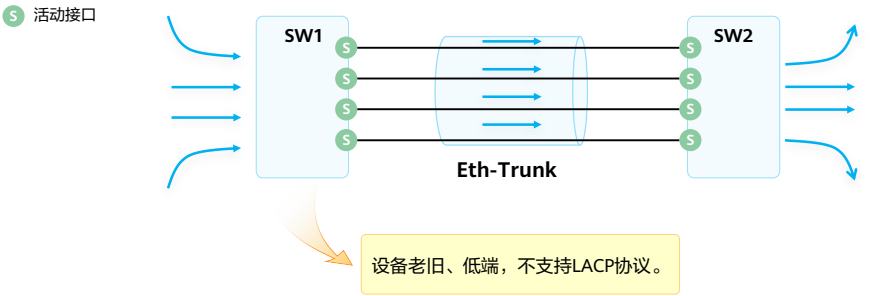
- 聚合组（Link Aggregation Group，LAG）：若干条链路捆绑在一起所形成的的逻辑链路。每个聚合组唯一对应着一个逻辑接口，这个逻辑接口又被称为链路聚合接口或Eth-Trunk接口。
- 成员接口和成员链路：组成Eth-Trunk接口的各个物理接口称为成员接口。成员接口对应的链路称为成员链路。
- 活动接口和活动链路：活动接口又叫选中（Selected）接口，是参与数据转发的成员接口。活动接口对应的链路被称为活动链路（Active link）
- 非活动接口和非活动链路：又叫非选中（Unselected）接口，是不参与转发数据的成员接口。非活动接口对应的链路被称为非活动链路（Inactive link）。
- 聚合模式：根据是否开启LACP（Link Aggregation Control Protocol，链路聚合控制协议），链路聚合可以分为手工模式和LACP模式。
- 其他概念：活动接口上限阈值和活动接口下限阈值。

HUAWEI

- 链路聚合接口可以作为普通的以太网接口来使用，与普通以太网接口的差别在于：转发的时候链路聚合组需要从成员接口中选择一个或多个接口来进行数据转发。
- 一个聚合组内要求成员接口以下参数相同：
  - 接口速率
  - 双工模式
  - VLAN配置：接口类型都是Trunk或者Access，如果为Access接口的default VLAN需要一致，如果为Trunk接口，接口放通的VLAN、缺省VLAN需要一致。

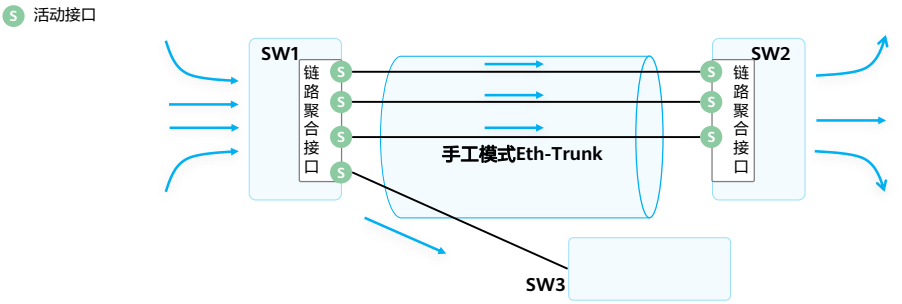
目录

手工模式



- 手工模式： Eth-Trunk的建立、成员接口的加入均由手动配置，双方系统之间不使用LACP进行协商。
- 正常情况下所有链路都是活动链路，该模式下所有活动链路都参与数据的转发，平均分担流量，如果某条活动链路故障，链路聚合组自动在剩余的活动链路中平均分担流量。
- 当聚合的两端设备中存在一个不支持LACP协议时，可以使用手工模式。

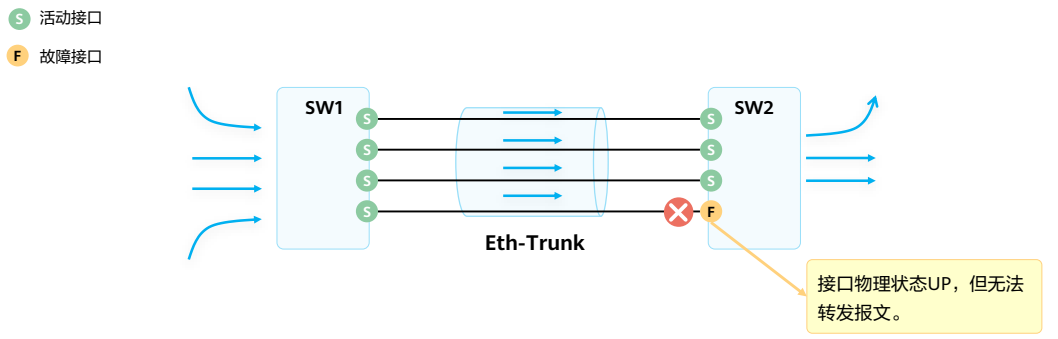
手工模式缺陷 (1)



- 为了使链路聚合接口正常工作，必须保证本端链路聚合接口中所有成员接口的对端接口：
  - 属于同一设备
  - 加入同一链路聚合接口
- 手工模式下，设备间没有报文交互，因此只能通过管理员人工确认。

- 在上图示例中SW1将四个接口加入到同一个聚合接口，但是其中一个接口的对端为SW3，而不是SW2，导致部分流量被负载分担到SW3，从而导致通信异常。

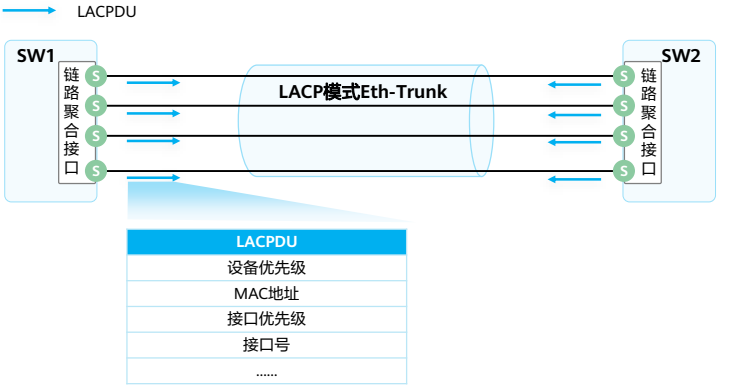
手工模式缺陷 (2)



- 手动模式下，设备只能通过物理层状态判断对端接口是否正常工作。

目录

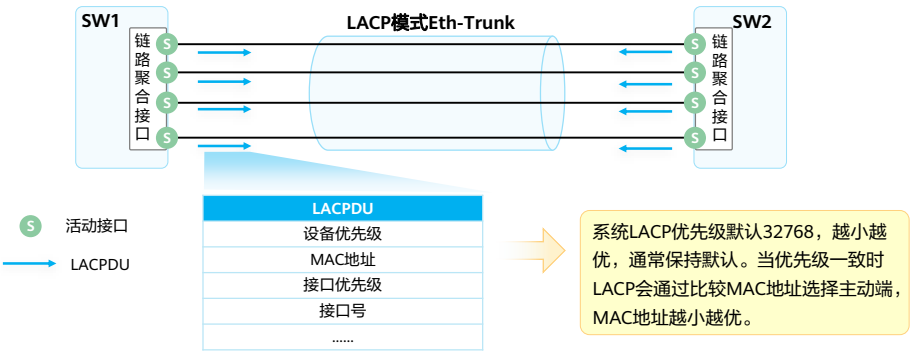
LACPDU



- LACP模式：采用LACP协议的一种链路聚合模式。设备间通过链路聚合控制协议数据单元（Link Aggregation Control Protocol Data Unit，LACPDU）进行交互，通过协议协商确保对端是同一台设备、同一个聚合接口的成员接口。
- LACPDU报文中包含设备优先级、MAC地址、接口优先级、接口号等。

系统优先级

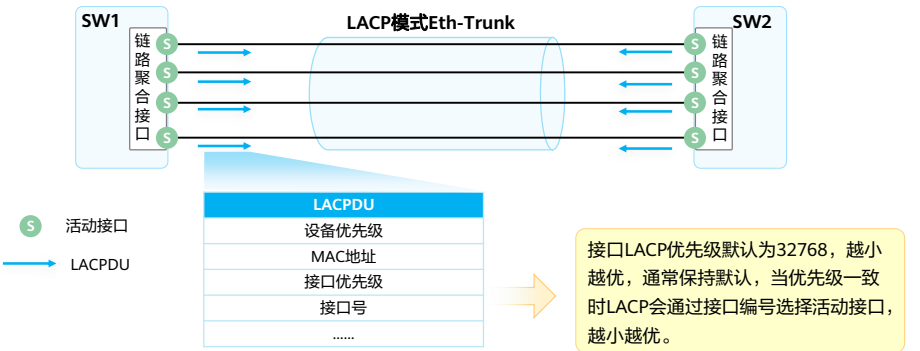
- LACP模式下，两端设备所选择的活动接口数目必须保持一致，否则链路聚合组就无法建立。此时可以使其中一端成为主动端，另一端（被动端）根据主动端选择活动接口。
- 通过系统LACP优先级确定主动端，值越小优先级越高。



HUAWEI

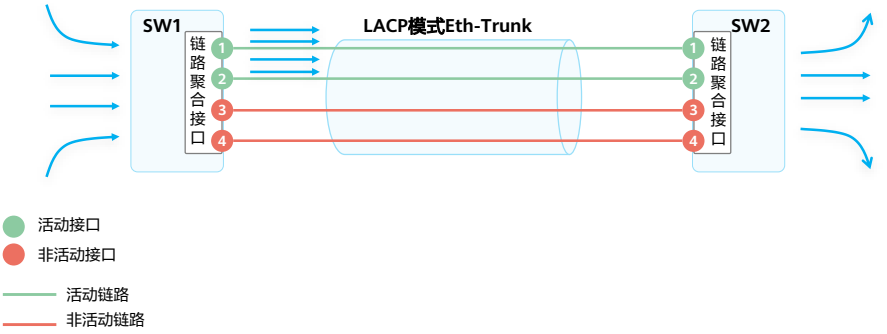
接口优先级

- 选出主动端后，两端都会以主动端的接口优先级来选择活动接口，优先级高的接口将优先被选为活动接口。接口LACP优先级值越小，优先级越高。



最大活动接口数 (1)

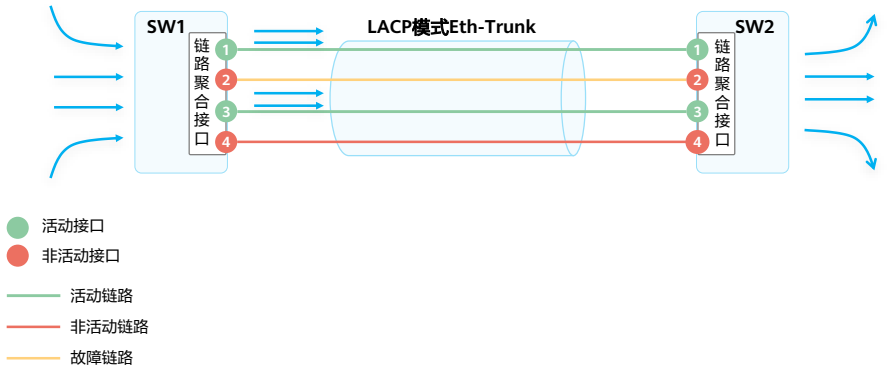
- LACP模式支持配置最大活动接口数目，当成员接口数目超过最大活动接口数目时会通过比较接口优先级、接口号选举出较优的接口成为活动接口，其余的则成为备份端口（非活动接口），同时对应的链路分别成为活动链路、非活动链路。交换机只会从活动接口中发送、接收报文。



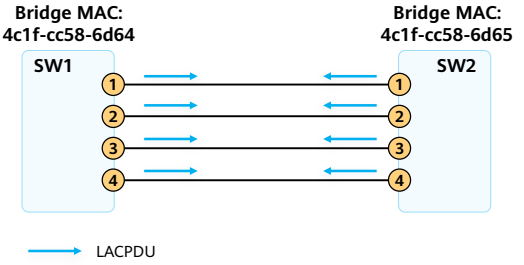
活动接口
非活动接口
活动链路
非活动链路

最大活动接口数 (2)

- 当活动链路中出现链路故障时，可以从非活动链路中找出一条优先级最高（接口优先级、接口编号比较）的链路替换故障链路，实现总体带宽不发生变化、业务的不间断转发。



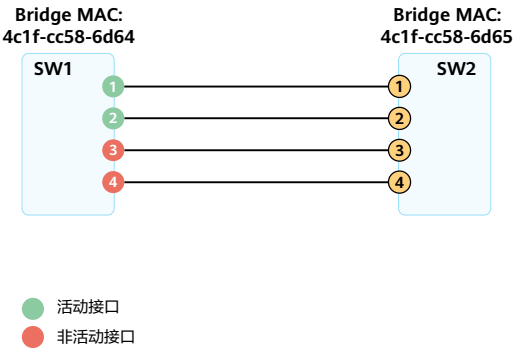
活动链路选举 (1)



- SW1、SW2配置LACP模式的链路聚合。两端都设置最大活跃接口数为2。
- 通过LACPDU选举出优先级较高的交换机SW1，作为LACP协商过程的主动端。

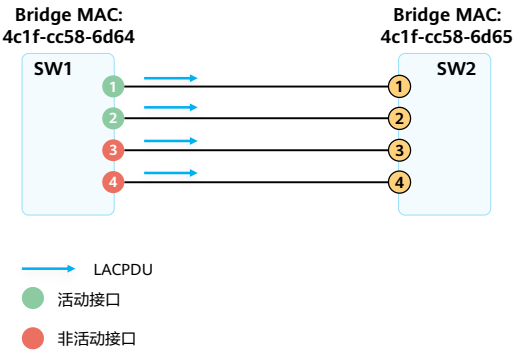
- SW1、SW2配置LACP模式的链路聚合,将四个接口加入Eth-Trunk中，接口编号分别为1、2、3、4。SW1、SW2配置Eth-Trunk最大活动接口数目为2，其余配置保持默认（系统优先级、接口优先级）。
- SW1、SW2分别从成员接口1、2、3、4对外发送LACPDU。
- SW1、SW2收到对端发送的LACPDU，比较系统优先级，都为默认的32768，继续比较MAC地址，SW1 MAC：4c1f-cc58-6d64，SW2 MAC：4c1f-cc58-6d65，SW1拥有更小的MAC地址，优选成为LACP选举的主动端。

活动链路选举 (2)



- SW1在本端通过比较接口优先级、接口编号选举出活动接口，其中1、2号接口在相同的接口优先级下拥有更小的接口编号，成为活动接口。

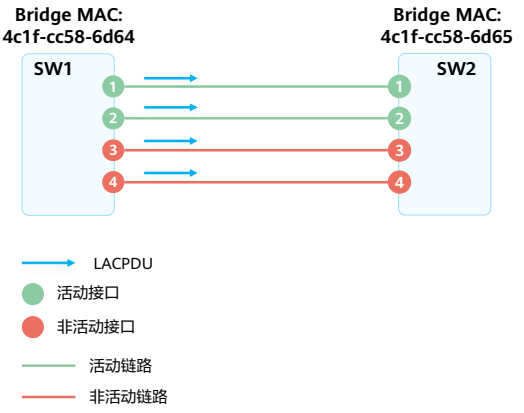
活动链路选举 (3)



- SW1通过LACPDU将本端活动端口选举结果告知对端。

- LACP通过LACPDU中的三个flags来标识该端口的状态，如果是活跃端口如下三个flags的值将会是1：
  - Synchronization
  - Collecting
  - Distributing
- 如果是非活跃端口，该三个flags字段的值将为0。

活动链路选举 (4)

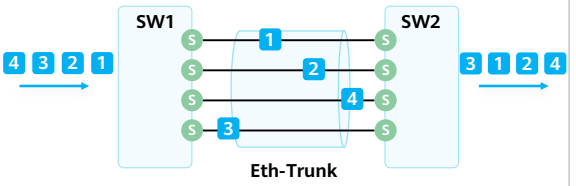


- SW2依据SW1的选举结果，明确本端的活动接口，同时对应的链路成为活动链路。
- 至此，Eth-Trunk的活动链路选举过程完成。

负载分担

基于包的负载分担

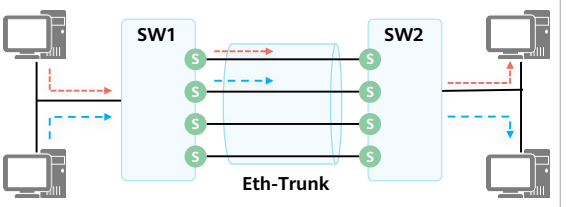
S 活动接口



在使用Eth-Trunk转发数据时，由于聚合组两端设备之间有多条物理链路，如果每个数据帧在不同的链路上转发，则有可能导致数据帧到达对端时间不一致，从而引起数据乱序。

基于流的负载分担

S 活动接口

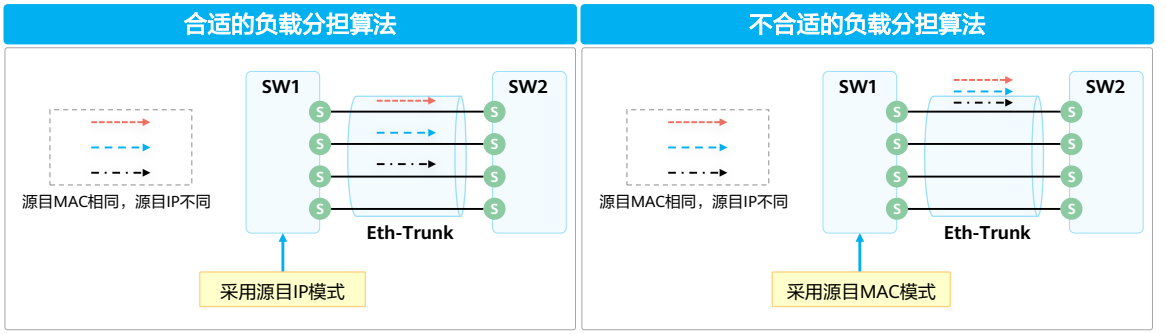


Eth-Trunk推荐采用逐流负载分担的方式，即一条相同的流负载到一条链路，这样既保证了同一数据流的数据帧在同一条物理链路转发，又实现了流量在聚合组内各物理链路上的负载分担。

HUAWEI

负载分担模式

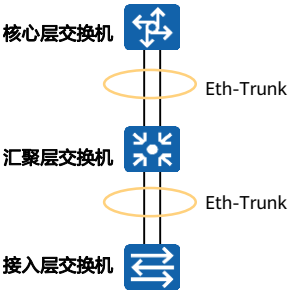
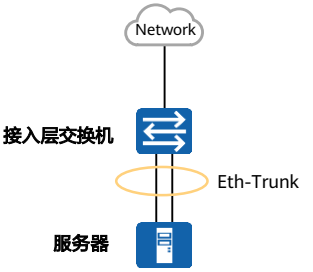
- Eth-trunk支持基于报文的IP地址或MAC地址来进行负载分担，可以配置不同的模式（本地有效，对出方向报文生效）将数据流分担到不同的成员接口上。
- 常见的模式有：源IP、源MAC、目的IP、目的MAC、源目IP、源目MAC。
- 实际业务中用户需要根据业务流量特征选择配置合适的负载分担方式。业务流量中某种参数变化越频繁，选择与此参数相关的负载分担方式就越容易实现负载均衡。



- 如果报文的IP地址变化较频繁，那么选择基于源IP、目的IP或者源目IP的负载分担模式更有利于流量在各物理链路间合理的负载分担；
- 如果报文的MAC地址变化较频繁，IP地址比较固定，那么选择基于源MAC、目的MAC或源目MAC的负载分担模式更有利于流量在各物理链路间合理的负载分担。
- 如果负载分担模式选择的和实际业务特征不相符，可能会导致流量分担不均，部分成员链路负载很高，其余的成员链路却很空闲，如在报文源目IP变化频繁但是源目MAC固定的场景下选择源目MAC模式，那将会导致所有流量都分担在一条成员链路上。

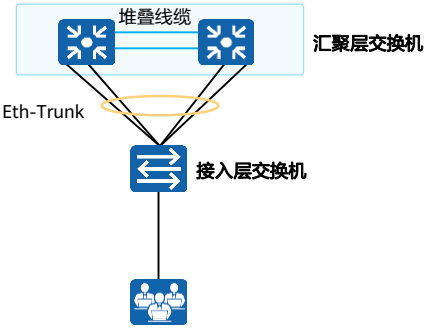
目录

典型使用场景 (1)

交换机之间	交换机与服务器之间
	
为保证交换机之间的链路带宽以及可靠性，可以在交换机之间部署多条物理链路并使用Eth-Trunk。	为了提高服务器的接入带宽和可靠性，将两个或者更多的物理网卡聚合成一个网卡组，与交换机建立链路聚合。

典型使用场景 (2)

交换机与堆叠系统



堆叠系统使得两台交换机成为一台逻辑上的设备，交换机与堆叠系统通过链路聚合互联可以组建高可靠、无环的网络。

防火墙双机热备心跳线

防火墙双机热备组网中使用心跳线来检测对端设备的状态，为防止单端口、单链路故障导致的状态监测错误可以部署Eth-Trunk，使用Eth-Trunk作为检测状态的心跳线。

目录

配置命令介绍 (1)

1. 创建链路聚合组

   [Huawei] **interface eth-trunk** *trunk-id*

   创建Eth-Trunk接口，并进入Eth-Trunk接口视图。

2. 配置链路聚合模式

   [Huawei-Eth-Trunk1] **mode** *{***lacp** / **manual load-balance** *}*

   Mode lacp配置链路聚合模式为lacp模式，mode manual load-balance配置链路聚合模式为手工模式。
   注意：需要保持两端链路聚合模式一致。

3. 将接口加入链路聚合组中（以太网接口视图）

   [Huawei-GigabitEthernet0/0/1] **eth-trunk** *trunk-id*

   在接口视图下，把接口加入到Eth-Trunk中。

配置命令介绍 (2)

4. 将接口加入链路聚合组中（Eth-Trunk视图）

   [Huawei-Eth-Trunk1] **trunkport** *interface-type { interface-number}*

   在Eth-Trunk视图中将接口加入到链路聚合组中。3、4两种方式都可以将接口加入到链路聚合组中。

5. 使能允许不同速率端口加入同一Eth-Trunk接口的功能

   [Huawei-Eth-Trunk1] **mixed-rate link enable**

   缺省情况下，设备未使能允许不同速率端口加入同一Eth-Trunk接口的功能，只能相同速率的接口加入到同一个Eth-Trunk接口中。

6. 配置系统LACP优先级

   [Huawei] **lacp priority** *priority*

   系统LACP优先级值越小优先级越高，缺省情况下，系统LACP优先级为32768。

配置命令介绍 (3)

7. 配置接口LACP优先级

[Huawei-GigabitEthernet0/0/1] **lacp priority** *priority*

在接口视图下配置接口LACP优先级。缺省情况下，接口的LACP优先级是32768。接口优先级取值越小，接口的LACP优先级越高。
只有在接口已经加入到链路聚合中才可以配置该命令。

8. 配置最大活动接口数

[Huawei-Eth-Trunk1] **max active-linknumber** *{number}*

配置时需注意保持本端和对端的最大活动接口数一致，只有LACP模式支持配置最大活动接口数。

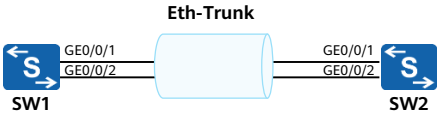
9. 配置最小活动接口数

[Huawei-Eth-Trunk1] **least active-linknumber** *{number}*

本端和对端设备的活动接口数下限阈值可以不同，手动模式、LACP模式都支持配置最小活动接口数。
配置最小活动接口数目的是为了保证最小带宽，当前活动链路数目小于下限阈值时，Eth-Trunk接口的状态转为Down。

- 不同型号交换机的可设置的最大活动接口数并不一致，如S6720HI、S6730H、S6730S和S6730S-S链路聚合组活动接口数的上限阈值是32，而S6720LI、S6720S-LI、S6720SI和S6720S-SI链路聚合组活动接口数的上限阈值是16。具体数值查阅产品手册确定。
- 设置最小活动接口数目是为了保证最小带宽，当带宽过小时一些对链路带宽有要求的业务将会出现异常，此时切断Eth-Trunk通过网络自身的高可靠性将业务切换到其他路径，从而保证业务的正常运行。

手工模式链路聚合配置举例



- 案例需求描述：
  - SW1、SW2都连接着VLAN10、VLAN20的网络。
  - SW1和SW2之间通过两根以太网链路互联，为了提供链路冗余以及保证传输可靠性，在SW1、SW2之间配置手工模式的链路聚合。

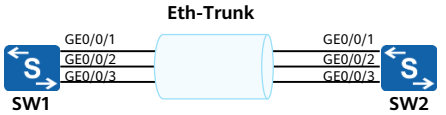
SW1的配置如下：

```
[SW1] interface eth-trunk 1
[SW1-Eth-Trunk1] trunkport gigabitethernet 0/0/1 to 0/0/2
[SW1-Eth-Trunk1] port link-type trunk
[SW1-Eth-Trunk1] port trunk allow-pass vlan 10 20
```

SW2的配置如下：

```
[SW2] interface eth-trunk 1
[SW2-Eth-Trunk1] trunkport gigabitethernet 0/0/1 to 0/0/2
[SW2-Eth-Trunk1] port link-type trunk
[SW2-Eth-Trunk1] port trunk allow-pass vlan 10 20
```

LACP模式链路聚合配置举例 (1)

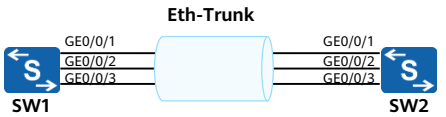


- 案例需求描述：
  - SW1、SW2都连接着VLAN10、VLAN20的网络。
  - SW1和SW2之间通过三根以太网链路互联，为了提供链路冗余以及保证传输可靠性，在SW1、SW2之间配置LACP模式的链路聚合，并且手动调整优先级让SW1成为主动端，并配置最大活跃端口为2，另外一条链路作为备份。

SW1的配置如下：

```
[SW1] interface eth-trunk 1
[SW1-Eth-Trunk1] mode lacp
[SW1-Eth-Trunk1] max active-linknumber 2
[SW1-Eth-Trunk1] trunkport gigabitethernet 0/0/1 to 0/0/3
[SW1-Eth-Trunk1] port link-type trunk
[SW1-Eth-Trunk1] port trunk allow-pass vlan 10 20
[SW1-Eth-Trunk1] quit
[SW1] lacp priority 30000
```

LACP模式链路聚合配置举例 (2)



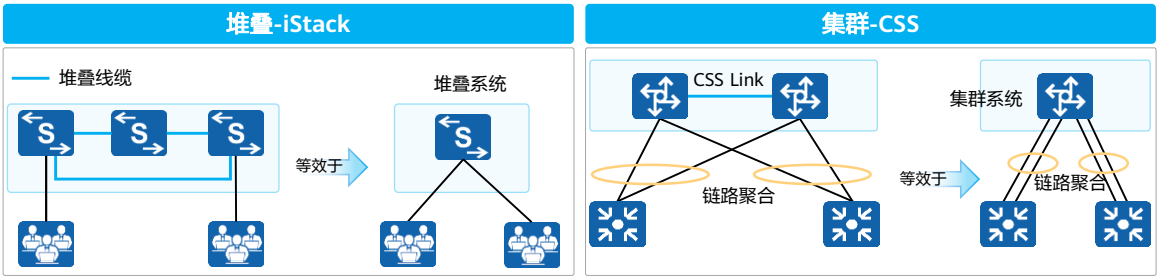
- 案例需求描述：
  - SW1、SW2都连接着VLAN10、VLAN20的网络。
  - SW1和SW2之间通过三根以太网链路互联，为了提供链路冗余以及保证传输可靠性，在SW1、SW2之间配置LACP模式的链路聚合，并且手动调整优先级让SW1成为主动端，并配置最大活跃端口为2，另外一条链路作为备份。

SW2的配置如下：

```
[SW2] interface eth-trunk 1
[SW2-Eth-Trunk1] mode lacp
[SW2-Eth-Trunk1] max active-linknumber 2
[SW2-Eth-Trunk1] trunkport gigabitethernet 0/0/1 to 0/0/3
[SW2-Eth-Trunk1] port link-type trunk
[SW2-Eth-Trunk1] port trunk allow-pass vlan 10 20
[SW2-Eth-Trunk1] quit
```

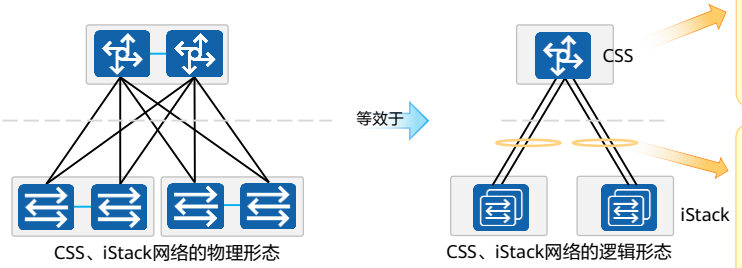
目录

什么是堆叠、集群



- 堆叠（iStack）：多台支持堆叠特性的交换机通过堆叠线缆连接在一起，从逻辑上变成一台交换设备，作为一个整体参与数据转发。
- 集群（Cluster Switch System，CSS ）：将两台支持集群特性的交换机设备组合在一起，从逻辑上组合成一台交换设备。
- 集群只支持两台设备，一般框式交换机支持CSS，盒式设备支持iStack。

堆叠、集群的优势



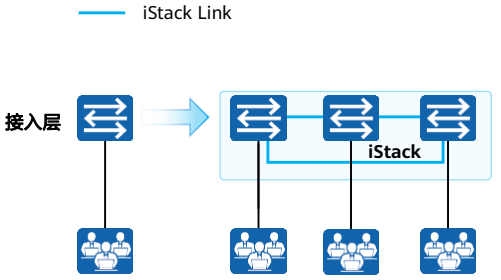
- 逻辑上一台设备，简化运维，方便管理。
- 一台物理设备故障，其他设备可以接管转发、控制平台，避免了单点故障。

- 跨设备的链路聚合，物理上的无环网络，无需再部署STP。
- 链路聚合中的链路全部有效使用，链路利用率100%。

- 交换机多虚一：堆叠交换机对外表现为一台逻辑交换机，控制平面合一，统一管理。
- 转发平面合一：堆叠内物理设备转发平面合一，转发信息共享并实时同步。
- 跨设备链路聚合：跨物理设备的链路被聚合成一个Eth-Trunk端口，和下游设备实现互联。

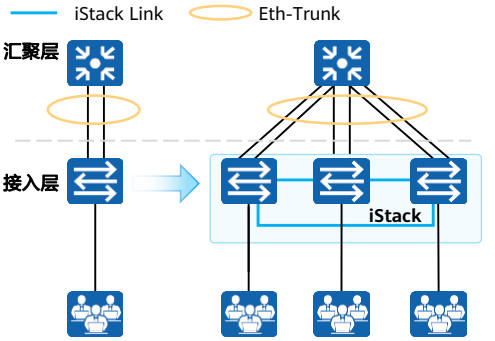
实际应用（1）

扩展端口



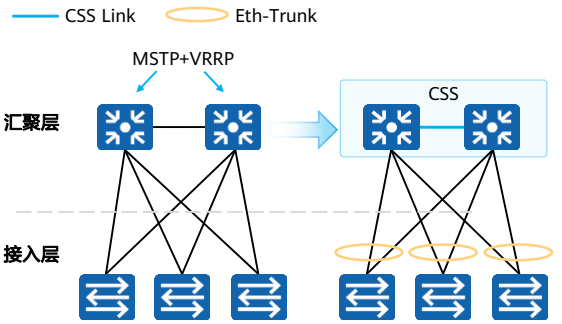
- 当接入的用户数增加到原交换机端口密度不能满足接入需求时，可以增加新交换机与原交换机组成堆叠系统扩展端口数量。

扩展带宽、冗余备份



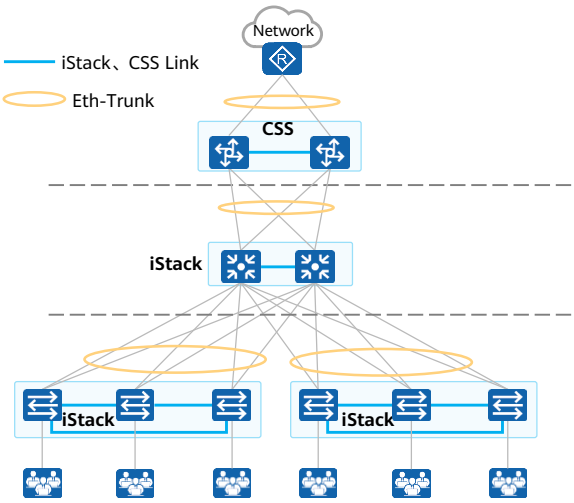
- 需要增大交换机上行带宽时，可以增加新交换机与原交换机组成堆叠系统，将成员交换机的多条物理链路配置成一个聚合组，提高交换机的上行带宽，同时形成设备间备份和链路的跨设备冗余备份，增加可靠性。

实际应用 (2)



- 两台设备组成集群，虚拟成单一的逻辑设备。简化后的组网不再需要使用MSTP、VRRP等协议，简化了网络配置，同时依靠跨设备的链路聚合，实现快速收敛，提高了可靠性。

推荐架构



核心层

- 核心使用CSS集群组网，上下行采用Eth-Trunk，构建高可靠、无环的网络。

汇聚层

- 汇聚交换机采用iStack，上下行采用Eth-Trunk，构建高可靠、无环的网络。

接入层

- 地理位置接近的接入设备（如一个楼宇内的接入交换机）使用iStack虚拟化成为一台逻辑上的设备，端口数量充足，简化了管理。
- 使用Eth-Trunk和汇聚层互联，逻辑上网络结构简单，不再需要使用STP、VRRP。具有高可靠性、高上行带宽、快速收敛的优势。

思考题

1. 基于包和基于流的负载分担有何区别?
2. LACP模式如何选举主动端?
3. CSS、iStack有何优势?

1. 基于包每个数据包负载到不同的链路，有可能导致报文乱序，基于流一条相同的流负载到一个相同的链路，不会发生报文乱序，但是单条流无法利用整个聚合接口的逻辑带宽。
2. 比较系统优先级，越小越优。如果系统优先级相同则继续比较桥MAC，越小越优。优先级高者成为主动端。
3. 简化网络管理、提高网络可靠性、能够充分利用网络链路带宽、使用跨设备的Eth-Trunk可以构建物理上无环的网络。

本章总结

- 为提高链路可靠性、链路利用率、链路带宽可以使用链路聚合技术，按照聚合方式不同可以分为静态聚合和LACP模式聚合。
- LACP模式采用报文协商，可以实现活动链路的备份，在链路出现故障时将备份链路选举为活动链路继续参与转发。
- 为保证报文到达的顺序，链路聚合的负载分担采用基于流的形式。
- 使用iStack、CSS技术可以简化网络管理、简化网络结构、提高网络可靠性。


谢 谢
www.huawei.com

ACL原理与配置

前言

- 随着网络的飞速发展，网络安全和网络服务质量QoS（Quality of Service)问题日益突出。访问控制列表 (ACL, Access Control List)是与其紧密相关的一个技术。
- ACL可以通过对网络中报文流的精确识别，与其他技术结合，达到控制网络访问行为、防止网络攻击和提高网络带宽利用率的目的，从而切实保障网络环境的安全性和网络服务质量的可靠性。
- 在本章节中，将介绍ACL的基本原理和基本作用，ACL的不同种类及特点，ACL的基本组成和匹配顺序，通配符的使用方法和ACL的相关配置。

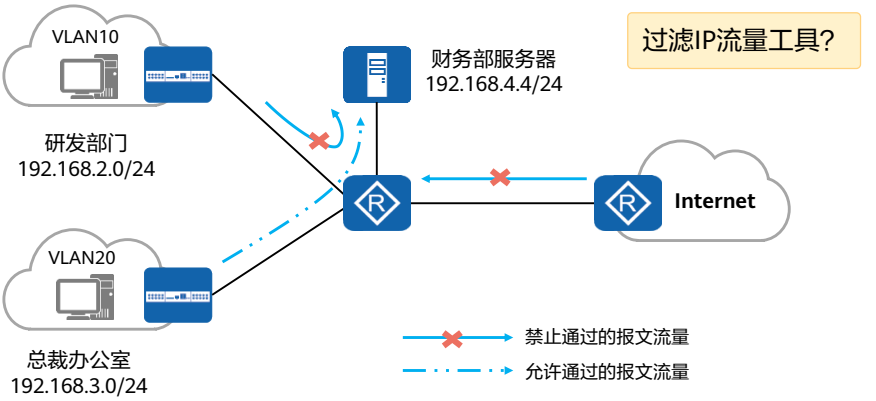
- 注：
  - 不同的网络设备厂商在ACL技术的实现上各不相同，本章节对于ACL技术的描述是针对华为网络设备上所实现的ACL技术而言的。
  - 局域网，LAN (Local Area Network) 是连接住宅、学校、实验室、大学校园或办公大楼等有限区域内计算机的计算机网络 。

目标

- 学完本课程后，您将能够：
  - 描述ACL的基本原理和基本作用
  - 区分ACL的不同种类及特点
  - 描述ACL规则的基本组成结构和匹配顺序
  - 掌握ACL中通配符的使用方法
  - 完成ACL的基本组网配置

目录

技术背景：需要一个工具，实现流量过滤

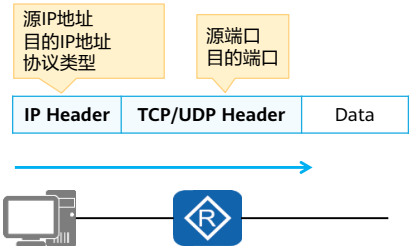


- 某公司为保证财务数据安全，禁止研发部门访问财务服务器，但总裁办公室不受限制。

- 随着网络的飞速发展，网络安全和网络服务质量QoS（Quality of Service）问题日益突出。
  - 园区重要服务器资源被随意访问，园区机密信息容易泄露，造成安全隐患。
  - Internet病毒肆意侵略园区内网，内网环境的安全性堪忧。
  - 网络带宽被各类业务随意挤占，服务质量要求最高的语音、视频业务的带宽得不到保障，造成用户体验差。
- 以上种种问题，都对正常的网络通信造成了很大的影响。因此，提高网络安全性和服务质量迫在眉睫，我们需要对网络进行控制。比如，需要借助一个工具帮助实现一些流量的过滤。

ACL概述

- ACL是由一系列permit或deny语句组成的、有序规则的列表。
- ACL是一个匹配工具，能够对报文进行匹配和区分。



ACL应用

- 匹配IP流量
- 在Traffic-filter中被调用
- 在NAT（Network Address Translation）中被调用
- 在路由策略中被调用
- 在防火墙的策略部署中被调用
- 在QoS中被调用
- 其他.....

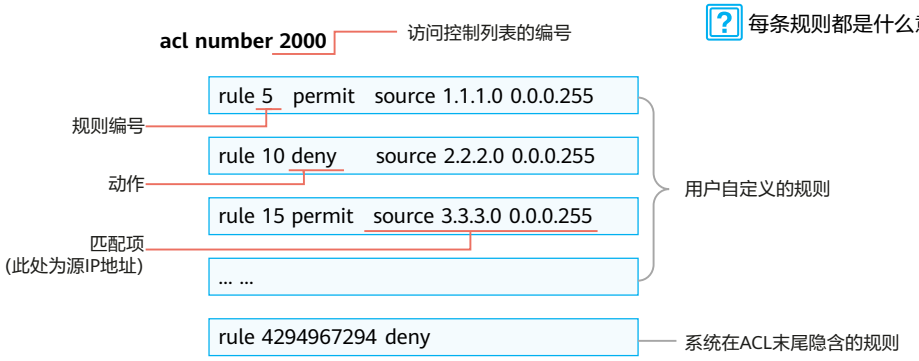
HUAWEI

- 通过ACL可以实现对网络中报文流的精确识别和控制，达到控制网络访问行为、防止网络攻击和提高网络带宽利用率的目的，从而切实保障网络环境的安全性和网络服务质量的可靠性。
  - ACL是由permit或deny语句组成的一系列有顺序的规则的集合；它通过匹配报文的相关字段实现对报文的分类。
  - ACL是能够匹配一个IP数据包中的源IP地址、目的IP地址、协议类型、源目的端口等元素的基础性工具；ACL还能够用于匹配路由条目。
- 在本章课程中主要通过流量过滤来介绍ACL。

目录

ACL的组成

- ACL由若干条permit或deny语句组成。每条语句就是该ACL的一条规则，每条语句中的permit或deny就是与这条规则相对应的处理动作。



  HUAWEI

- ACL的组成：
  - ACL编号：在网络设备上配置ACL时，每个ACL都需要分配一个编号，称为ACL编号，用来标识ACL。不同分类的ACL编号范围不同，这个后面具体讲。
  - 规则：前面提到了，一个ACL通常由若干条“permit/deny”语句组成，每条语句就是该ACL的一条规则。
  - 规则编号：每条规则都有一个相应的编号，称为规则编号，用来标识ACL规则。可以自定义，也可以系统自动分配。ACL规则的编号范围是0～4294967294，所有规则均按照规则编号从小到大进行排序。
  - 动作：每条规则中的permit或deny，就是与这条规则相对应的处理动作。permit指“允许”，deny指“拒绝”，但是ACL一般是结合其他技术使用，不同的场景，处理动作的含义也有所不同。
    - 比如：ACL如果与流量过滤技术结合使用（即流量过滤中调用ACL），permit就是“允许通行”的意思，deny就是“拒绝通行”的意思。
  - 匹配项：ACL定义了极其丰富的匹配项。例子中体现的源地址，ACL还支持很多其他规则匹配项。例如，二层以太网帧头信息（如源MAC、目的MAC、以太帧协议类型）、三层报文信息（如目的地址、协议类型）以及四层报文信息（如TCP/UDP端口号）等。
- 提问：rule 5 permit source 1.1.1.0 0.0.0.255 是什么意思？这个在后续课程中会介绍。

规则编号

```
acl number 2000
                规则编号
 rule    5      deny     source  10.1.1.1  0
 rule    10     deny     source  10.1.1.2  0
 rule    15     permit   source  10.1.1.0  0.0.0.255
          步长=5
```

如果希望增加1条规则，该如何处理？

rule 11 deny source 10.1.1.3 0

```
acl number 2000
 rule    5      deny     source  10.1.1.1  0
 rule    10     deny     source  10.1.1.2  0
 rule    11     deny     source  10.1.1.3  0
 rule    15     permit   source  10.1.1.0  0.0.0.255
```

规则编号与步长

- **规则编号（Rule ID）：**

  一个ACL中的每一条规则都有一个相应的编号。

- **步长（Step）：**

  步长是系统自动为ACL规则分配编号时，每个相邻规则编号之间的差值，缺省值为5。步长的作用是为了方便后续在旧规则之间，插入新的规则。

- **Rule ID分配规则：**

  系统为ACL中首条未手工指定编号的规则分配编号时，使用步长值（例如步长=5，首条规则编号为5）作为该规则的起始编号；为后续规则分配编号时，则使用大于当前ACL内最大规则编号且是步长整数倍的最小整数作为规则编号。

- 规则编号和步长的概念：
  - 规则编号：每条规则都有一个相应的编号，称为规则编号，用来标识ACL规则。可以自定义，也可以系统自动分配。
  - 步长：系统自动为ACL规则分配编号时，每个相邻规则编号之间会有一个差值，这个差值称为“步长”。缺省步长为5，所以规则编号就是5/10/15...以此类推。
    - 如果手工指定了一条规则，但未指定规则编号，系统就会使用大于当前ACL内最大规则编号且是步长整数倍的最小整数作为规则编号。
    - 步长可以调整，如果将步长改为2，系统则会自动从当前步长值开始重新排列规则编号，规则编号就变成2、4、6...。
- 那步长的作用是什么？直接rule 1/2/3/4...为什么不可以？
  - 先来看一个小题目：如果希望增加一条规则，该如何处理？
  - 可以在rule 10和rule 15之间，手工加入一条rule 11。
  - 因此，设置一定长度的步长的作用，是方便后续在旧规则之间插入新的规则。

通配符 (1)

acl number 2000

					通配符
rule	5	deny	source	10.1.1.1	0
rule	10	deny	source	10.1.1.2	0
rule	15	permit	source	10.1.1.0	0.0.0.255

通配符 (Wildcard)

- 通配符是一个32比特长度的数值，用于指示IP地址中，哪些比特位需要严格匹配，哪些比特位无需匹配。
- 通配符通常采用类似网络掩码的点分十进制形式表示，但是含义却与网络掩码完全不同。

- 匹配规则：

  “0”表示“严格匹配”；“1”表示“任意”

如何匹配192.168.1.1/24对应网段的地址?

	严格匹配			任意	
192.168.1.1	11000000	10101000	00000000	00000001	192.168.1.0/24网段
0.0.0.255	00000000	00000000	00000000	11111111	

- 当进行IP地址匹配的时候，后面会跟着32位掩码位，这32位称为通配符。
- 通配符，也是点分十进制格式，换算成二进制后，“0”表示“匹配”，“1”表示“不关心”。
- 具体看下这2条规则：
  - rule 5: 拒绝源IP地址为10.1.1.1报文通过——因为通配符为全0，所以每一位都要严格匹配，因此匹配的是主机IP地址10.1.1.1；
  - rule 15:允许源IP地址为10.1.1.0/24网段地址的报文通过——因为通配符：0.0.0.11111111，后8位为1，表示不关心，因此10.1.1.xxxxxxxx 的后8位可以为任意值，所以匹配的是10.1.1.0/24网段。
- 例子：如果要精确匹配192.168.1.1/24这个IP地址对应的网段地址，通配符是多少?
  - 可以得出：网络位需要严格匹配，主机位无所谓，因此通配符为“0.0.0.255”。

通配符 (2)

- 匹配192.168.1.0/24这个子网中的奇数IP地址，例如192.168.1.1、192.168.1.3、192.168.1.5等。

  HUAWEI

- 如果想匹配192.168.1.0/24网段中的奇数IP地址，通配符该怎么写呢?
  - 我们先来看一看，奇数IP地址都有哪些：192.168.1.1、192.168.1.5、192.168.1.11.....
  - 后八位写成二进制：192.168.1.00000001、192.168.1.00000101、192.168.1.00001011.....
  - 可以看出共同点：最后8位的高7位是任意值，最低位固定为1，因此答案是：192.168.1.1 0.0.0.254（0.0.0.11111110）
- 这就得出了通配符的一个特点：通配符中的1或者0是可以不连续的。

- 还有两个特殊的通配符：
  - 当通配符全为0来匹配IP地址时，表示精确匹配某个IP地址；
  - 当通配符全为1来匹配0.0.0.0地址时，表示匹配了所有IP地址。

ACL的分类与标识

- 基于ACL规则定义方式的分类

分类	编号范围	规则定义描述
基本ACL	2000~2999	仅使用报文的源IP地址、分片信息和生效时间段信息来定义规则。
高级ACL	3000~3999	可使用IPv4报文的源IP地址、目的IP地址、IP协议类型、ICMP类型、TCP源/目的端口号、UDP源/目的端口号、生效时间段等来定义规则。
二层ACL	4000~4999	使用报文的以太网帧头信息来定义规则，如根据源MAC地址、目的MAC地址、二层协议类型等。
用户自定义ACL	5000~5999	使用报文头、偏移位置、字符串掩码和用户自定义字符串来定义规则。
用户ACL	6000~9999	既可使用IPv4报文的源IP地址或源UCL（User Control List）组，也可使用目的IP地址或目的UCL组、IP协议类型、ICMP类型、TCP源端口/目的端口、UDP源端口/目的端口号等来定义规则。

- 基于ACL标识方法的分类

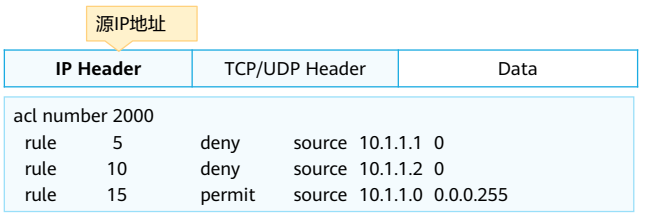
分类	规则定义描述
数字型ACL	传统的ACL标识方法。创建ACL时，指定一个唯一的数字标识该ACL。
命名型ACL	通过名称代替编号来标识ACL。

- 基于ACL规则定义方式的划分，可分为：
  - 基本ACL、高级ACL、二层ACL、用户自定义ACL和用户ACL。
- 基于ACL标识方法的划分，则可分为：
  - 数字型ACL和命名型ACL。
- 注意：用户在创建ACL时可以为其指定编号，不同的编号对应不同类型的ACL。同时，为了便于记忆和识别，用户还可以创建命名型ACL，即在创建ACL时为其设置名称。命名型ACL，也可以是“名称 数字”的形式，即在定义命名型ACL时，同时指定ACL编号。如果不指定编号，系统则会自动为其分配一个数字型ACL的编号。
- 本课程的ACL分类以华为S系列交换机为例。

基本ACL与高级ACL

- 基本ACL

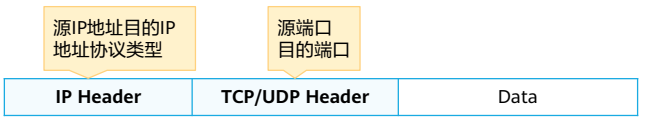
编号范围：
2000-2999




```
acl number 2000
 rule      5        deny      source  10.1.1.1  0
 rule      10       deny      source  10.1.1.2  0
 rule      15       permit    source  10.1.1.0  0.0.0.255
```

- 高级ACL

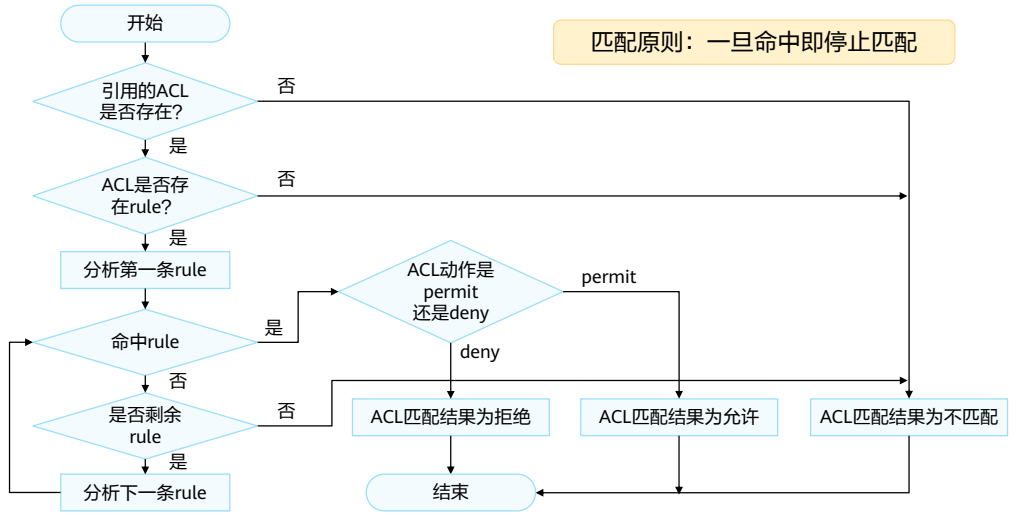
编号范围：
3000-3999




```
acl number 3000
 rule   5   permit  ip    source   10.1.1.0   0.0.0.255   destination   10.1.3.0   0.0.0.255
 rule   10  permit  tcp   source   10.1.2.0   0.0.0.255   destination   10.1.3.0   0.0.0.255   destination-port eq 21
```


- 基本ACL：
  - 主要针对IP报文的源IP地址进行匹配，基本ACL的编号范围是2000-2999。
  - 比如这个例子，创建的是acl 2000，就意味着创建的是基本ACL。
- 高级ACL：
  - 可以根据IP报文中的源IP地址、目的IP地址、协议类型，TCP或UDP的源目端口号等元素进行匹配，可以理解为：基本ACL是高级ACL的一个子集，高级ACL可以比基本ACL定义出更精确、更复杂、更灵活的规则。

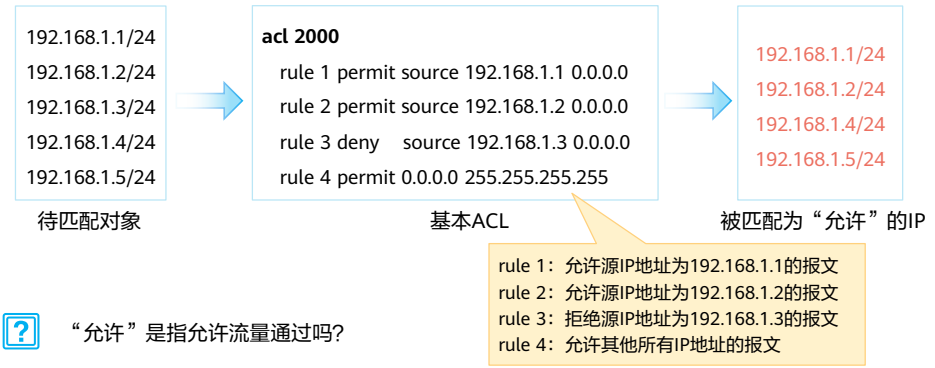
ACL的匹配机制



- ACL的匹配机制概括来说就是：
  - 配置ACL的设备接收报文后，会将该报文与ACL中的规则逐条进行匹配，如果不能匹配上，就会继续尝试去匹配下一条规则。
  - 一旦匹配上，则设备会对该报文执行这条规则中定义的处理动作，并且不再继续尝试与后续规则匹配。
- 匹配流程：首先系统会查找设备上是否配置了ACL。
  - 如果ACL不存在，则返回ACL匹配结果为：不匹配。
  - 如果ACL存在，则查找设备是否配置了ACL规则。
    - 如果规则不存在，则返回ACL匹配结果为：不匹配。
    - 如果规则存在，则系统会从ACL中编号最小的规则开始查找。
      - 如果匹配上了permit规则，则停止查找规则，并返回ACL匹配结果为：匹配（允许）。
      - 如果匹配上了deny规则，则停止查找规则，并返回ACL匹配结果为：匹配（拒绝）。
      - 如果未匹配上规则，则继续查找下一条规则，以此循环。如果一直查到最后一条规则，报文仍未匹配上，则返回ACL匹配结果为：不匹配。
- 从整个ACL匹配流程可以看出，报文与ACL规则匹配后，会产生两种匹配结果：“匹配”和“不匹配”。
  - 匹配（命中规则）：指存在ACL，且在ACL中查找到了符合匹配条件的规则。不论匹配的动作是“permit”还是“deny”，都称为“匹配”，而不是只是匹配上permit规则才算“匹配”。
  - 不匹配（未命中规则）：指不存在ACL，或ACL中无规则，再或者在ACL中遍历了所有规则都没有找到符合匹配条件的规则。以上三种情况，都叫做“不匹配”。
- 匹配原则：一旦命中即停止匹配。

ACL的匹配顺序及匹配结果

- 配置顺序（config模式）
  - 系统按照ACL规则编号从小到大的顺序进行报文匹配，规则编号越小越容易被匹配。

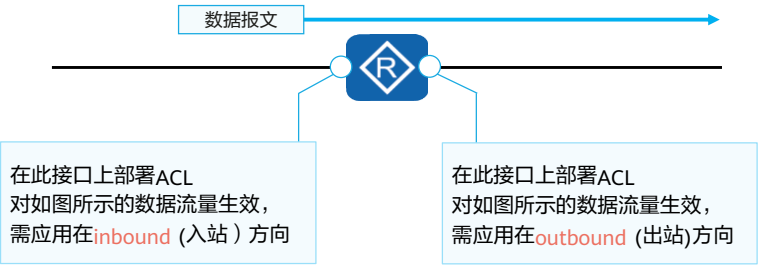


“允许”是指允许流量通过吗?

- 一条ACL可以由多条“deny或permit”语句组成，每一条语句描述一条规则，这些规则可能存在包含关系，也可能有重复或矛盾的地方，因此ACL的匹配顺序是十分重要的。
- 华为设备支持两种匹配顺序：自动排序（auto模式）和配置顺序（config模式）。缺省的ACL匹配顺序是config模式。
  - 自动排序，是指系统使用“深度优先”的原则，将规则按照精确度从高到低进行排序，并按照精确度从高到低的顺序进行报文匹配。——这个比较复杂，这里就不具体展开了，感兴趣的同学可以课后查看资料。
  - 配置顺序，系统按照ACL规则编号从小到大的顺序进行报文匹配，规则编号越小越容易被匹配。——这个就是我们前面提到的匹配顺序。
    - 如果后面又添加了一条规则，则这条规则会被加入到相应的位置，报文仍然会按照从小到大的顺序进行匹配。
- 匹配结果：（如图所示，以192.168.1.3/24为例）
  - 首先理解ACL 2000的含义：
    - rule 1：允许源IP地址为192.168.1.1的报文
    - rule 2：允许源IP地址为192.168.1.2的报文
    - rule 3：拒绝源IP地址为192.168.1.2的报文
    - rule 4：允许其他所有IP地址的报文

    - 当源IP地址为192.168.1.3的报文经过配置了ACL的设备时：
        - 首先查看rule 1，发现不匹配；
        - 继续查看rule 2，发现仍不匹配；
        - 继续查看rule 3，发现匹配，且是“拒绝”动作。
- 注意：ACL技术总是与其他技术结合在一起使用的，因此，所结合的技术不同，“允许(permit)”及“拒绝 (deny)”的内涵和作用也会不同。例如，当ACL技术与流量过滤技术结合使用时，permit就是“允许通行”的意思，deny就是“拒绝通行”的意思。

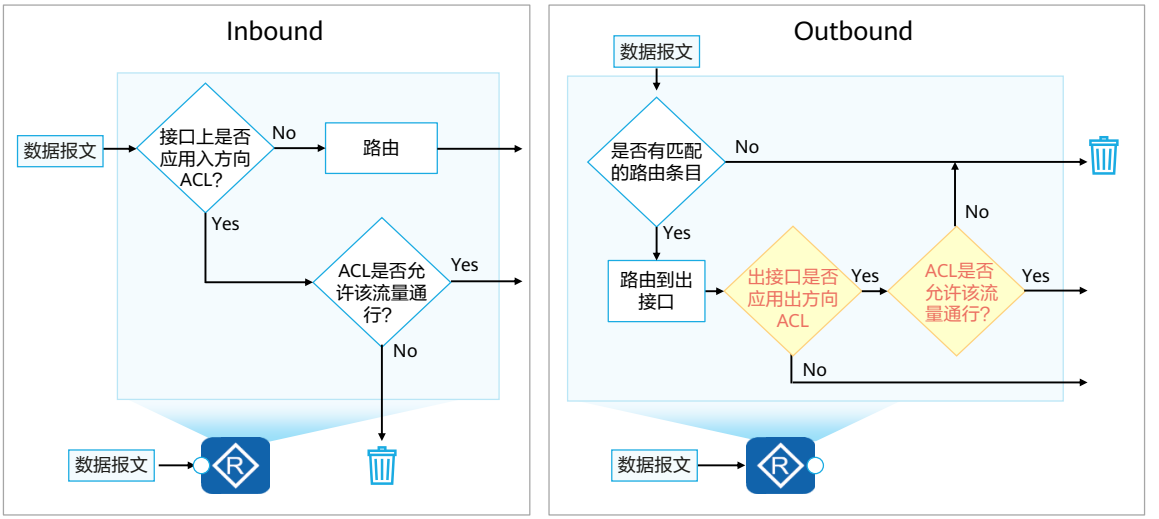
ACL的匹配位置



在此接口上部署ACL
对如图所示的数据流量生效，
需应用在inbound (入站）方向

在此接口上部署ACL
对如图所示的数据流量生效，
需应用在outbound (出站)方向

入站 (Inbound)及出站 (Outbound)方向



HUAWEI

目录

基本ACL的基础配置命令

1. 创建基本ACL

[Huawei] **acl** [**number**] *acl-number* [**match-order config**]

使用编号（2000～2999）创建一个数字型的基本ACL，并进入基本ACL视图。

[Huawei] **acl name** *acl-name* { **basic** | *acl-number* } [**match-order config**]

使用名称创建一个命名型的基本ACL，并进入基本ACL视图。

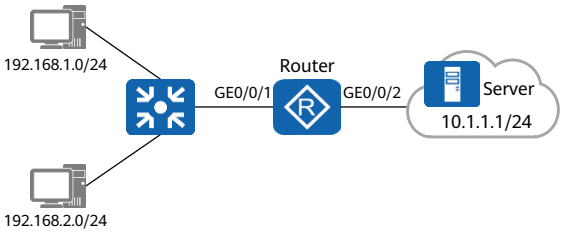
2. 配置基本ACL的规则

[Huawei-acl-basic-2000] **rule** [*rule-id*] { **deny** | **permit** } [**source** { *source-address source-wildcard* | **any** } | **time-range** *time-name*]

在基本ACL视图下，通过此命令来配置基本ACL的规则。

- 创建基本ACL
- [Huawei] **acl** [**number**] *acl-number* [**match-order config**]
  - *acl-number*：指定访问控制列表的编号。
  - **match-order config**：指定ACL规则的匹配顺序，config表示配置顺序。
- [Huawei] **acl name** *acl-name* { **basic** | *acl-number* } [**match-order config**]
  - *acl-name*：指定创建的ACL的名称。
  - **basic**：指定ACL的类型为基本ACL。
- 配置基本ACL规则
- [Huawei-acl-basic-2000] **rule** [*rule-id*] { **deny** | **permit** } [**source** { *source-address source-wildcard* | **any** } | **time-range** *time-name*]
  - *rule-id*：指定ACL的规则ID。
  - **deny**：指定拒绝符合条件的报文。
  - **permit**：指定允许符合条件的报文。
  - **source** { *source-address source-wildcard* | **any** }：指定ACL规则匹配报文的源地址信息。如果不配置，表示报文的任何源地址都匹配。其中：
    - *source-address*：指定报文的源地址。
    - *source-wildcard*：指定源地址通配符。
    - **any**：表示报文的任意源地址。相当于source-address为0.0.0.0或者source-wildcard为255.255.255.255。
  - **time-range** *time-name*：指定ACL规则生效的时间段。其中，time-name表示ACL规则生效时间段名称。如果不指定时间段，表示任何时间都生效。

案例：使用基本ACL过滤数据流量



- 配置需求：

  在Router上部署基本ACL后，ACL将试图穿越Router的源地址为192.168.1.0/24网段的数据包过滤掉，并放行其他流量，从而禁止192.168.1.0/24网段的用户访问Router右侧的服务器网络。

1、Router已完成IP地址和路由的相关配置

2、在Router上创建基本ACL，禁止192.168.1.0/24网段访问服务器网络：

```
[Router] acl 2000
[Router-acl-basic-2000] rule deny source 192.168.1.0  0.0.0.255
[Router-acl-basic-2000] rule permit source any
```

3、由于从接口GE0/0/1进入Router，所以在接口GE0/0/1的入方向配置流量过滤：

```
[Router] interface GigabitEthernet 0/0/1
[Router-GigabitEthernet0/0/1] traffic-filter inbound acl 2000
[Router-GigabitEthernet0/0/1] quit
```


HUAWEI

- 配置思路：
  - 配置基本ACL和流量过滤，使设备可以对特定网段的报文进行过滤。
- 配置步骤：
  1. 如图完成路由器的IP地址和路由相关配置。
  2. 创建基本ACL 2000并配置ACL规则，拒绝192.168.1.0/24网段的报文通过，允许其他网段的报文通过。
  3. 配置流量过滤。
- 注：
  - **traffic-filter**命令，用来在接口上配置基于ACL对报文进行过滤。
  - 命令格式：**traffic-filter** { **inbound** | **outbound** } **acl** { *acl-number* | **name** *acl-name* }
    - **inbound**：指定在接口入方向上配置报文过滤。
    - **outbound**：指定在接口出方向上配置报文过滤。
    - **acl**：指定基于IPv4 ACL对报文进行过滤。

高级ACL的基础配置命令 (1)

1. 创建高级ACL

[Huawei] **acl** [**number**] *acl-number* [**match-order config**]

使用编号（3000～3999）创建一个数字型的高级ACL，并进入高级ACL视图。

[Huawei] **acl name** *acl-name* { **advance** | *acl-number* } [**match-order config**]

使用名称创建一个命名型的高级ACL，进入高级ACL视图。

- 创建高级ACL
- [Huawei] **acl** [**number**] *acl-number* [**match-order config**]
  - *acl-number*：指定访问控制列表的编号。
  - **match-order config**：指定ACL规则的匹配顺序，config表示配置顺序。
- [Huawei] **acl name** *acl-name* { **advance** | *acl-number* } [**match-order config**]
  - *acl-name*：指定创建的ACL的名称。
  - **advance**：指定ACL的类型为高级ACL。

高级ACL的基础配置命令 (2)

2. 配置基本ACL的规则
根据IP承载的协议类型不同，在设备上配置不同的高级ACL规则。对于不同的协议类型，有不同的参数组合。
   - 当参数protocol为IP时，高级ACL的命令格式为

     rule [*rule-id*] { **deny** | **permit** } **ip** [**destination** { *destination-address destination-wildcard* | **any** } | **source** { *source-address source-wildcard* | **any** } | **time-range** *time-name* | [**dscp** *dscp* | [**tos** *tos* | **precedence** *precedence*]]]

     在高级ACL视图下，通过此命令来配置高级ACL的规则。

   - 当参数protocol为TCP时，高级ACL的命令格式为

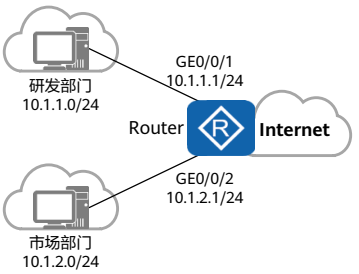
     rule [*rule-id*] { **deny** | **permit** } { *protocol-number* | **tcp** } [**destination** { *destination-address destination-wildcard* | **any** } | **destination-port** { **eq** port | **gt** port | **lt** port | **range** port-start port-end } | **source** { *source-address source-wildcard* | **any** } | **source-port** { **eq** port | **gt** port | **lt** port | **range** port-start port-end } | **tcp-flag** { **ack** | **fin** | **syn** } * | **time-range** time-name] *

     在高级ACL视图下，通过此命令来配置高级ACL的规则。

- 配置高级ACL的规则
- 当参数protocol为IP时：
  - **rule** [*rule-id*] { **deny** | **permit** } **ip** [**destination** { *destination-address destination-wildcard* | **any** } | **source** { *source-address source-wildcard* | **any** } | **time-range** *time-name* | [**dscp** *dscp* | [**tos** *tos* | **precedence** *precedence*]]]
    - **ip** ： 指定ACL规则匹配报文的协议类型为IP。
    - **destination** { *destination-address destination-wildcard* | **any** }：指定ACL规则匹配报文的目的地址信息。如果不配置，表示报文的任何目的地址都匹配。
    - **dscp** *dscp*：指定ACL规则匹配报文时，区分服务代码点（Differentiated Services Code Point），取值为：0~63。
    - **tos** *tos*：指定ACL规则匹配报文时，依据服务类型字段进行过滤，取值为：0~15。
    - **precedence** *precedence*：指定ACL规则匹配报文时，依据优先级字段进行过滤。precedence表示优先级字段值，取值为：0~7。

- 当参数protocol为TCP时：
  - **rule** [*rule-id*] { **deny** | **permit** } { *protocol-number* | **tcp** } [**destination** { *destination-address destination-wildcard* | **any** } | **destination-port** { **eq** *port* | **gt** *port* | **lt** *port* | **range** *port-start port-end* } | **source** { *source-address source-wildcard* | **any** } | **source-port** { **eq** port | **gt** port | **lt** port | **range** *port-start port-end* } | **tcp-flag** { **ack** | **fin** | **syn** } * | **time-range** *time-name*] *
    - *protocol-number* | **tcp**：指定ACL规则匹配报文的协议类型为TCP。可以采用数值6表示指定TCP协议。
    - **destination-port** { **eq** *port* | **gt** *port* | **lt** *port* | **range** *port-start port-end*}：指定ACL规则匹配报文的UDP或者TCP报文的目的端口，仅在报文协议是TCP或者UDP时有效。如果不指定，表示TCP/UDP报文的任何目的端口都匹配。其中：
      - **eq** *port*：指定等于目的端口；
      - **gt** *port*：指定大于目的端口；
      - **lt** *port*：指定小于目的端口；
      - **range** *port-start port-end*：指定源端口的范围。
    - **tcp-flag**：指定ACL规则匹配报文的TCP报文头中SYN Flag。

案例：使用高级ACL限制不同网段的用户互访（1）



配置需求：

- 某公司通过Router实现各部门之间的互连。为方便管理网络，管理员为公司的研发部和市场部规划了两个网段的IP地址。
- 现要求Router能够限制两个网段之间互访，防止公司机密泄露。

1、Router已完成IP地址和路由的相关配置。

2、创建高级ACL 3001并配置ACL规则，拒绝研发部访问市场部的报文：

```
[Router] acl 3001
[Router-acl-adv-3001] rule deny ip source 10.1.1.0 0.0.0.255
destination 10.1.2.0 0.0.0.255
[Router-acl-adv-3001] quit
```

3、创建高级ACL 3002并配置ACL规则，拒绝市场部访问研发部的报文：

```
[Router] acl 3002
[Router-acl-adv-3002] rule deny ip source 10.1.2.0 0.0.0.255
destination 10.1.1.0 0.0.0.255
[Router-acl-adv-3002] quit
```


HUAWEI

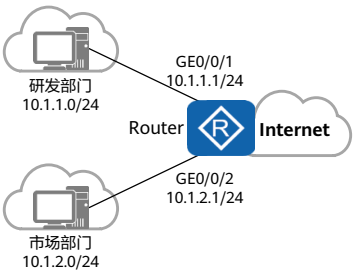
- 配置思路：

  配置高级ACL和流量过滤，使设备可以对研发部与市场部互访的报文进行过滤。

- 配置步骤：

  1. 如图完成路由器的IP地址和路由的相关配置。
  2. 创建高级ACL 3001并配置ACL规则，拒绝研发部访问市场部的报文通过。
  3. 创建高级ACL 3002并配置ACL规则，拒绝市场部访问研发部的报文通过。

案例：使用高级ACL限制不同网段的用户互访 (2)



配置需求：

- 某公司通过Router实现各部门之间的互连。为方便管理网络，管理员为公司的研发部和市场部规划了两个网段的IP地址。
- 现要求Router能够限制两个网段之间互访，防止公司机密泄露。

4、由于研发部和市场部互访的流量分别从接口GE0/0/1和GE0/0/2进入Router，所以在接口GE0/0/1和GE0/0/2的入方向配置流量过滤：

```
[Router] interface GigabitEthernet 0/0/1
[Router-GigabitEthernet0/0/1] traffic-filter inbound acl 3001
[Router-GigabitEthernet0/0/1] quit

[Router] interface GigabitEthernet 0/0/2
[Router-GigabitEthernet0/0/2] traffic-filter inbound acl 3002
[Router-GigabitEthernet0/0/2] quit
```


HUAWEI

- 配置步骤：
  4. 由于研发部和市场部互访的流量分别从接口GE0/0/1和GE0/0/2进入Router，所以在接口GE0/0/1和GE0/0/2的入方向配置流量过滤。
- 注：
  - **traffic-filter**命令，用来在接口上配置基于ACL对报文进行过滤。
  - 命令格式：**traffic-filter** { **inbound** | **outbound** } **acl** { *acl-number* | **name** *acl-name* }
    - **inbound**：指定在接口入方向上配置报文过滤。
    - **outbound**：指定在接口出方向上配置报文过滤。
    - **acl**：指定基于IPv4 ACL对报文进行过滤。

思考题

1. （单选）下列选项中，哪一项才是一条合法的基本ACL的规则？（　）

   A. rule permit ip

   B. rule deny ip

   C. rule permit source any

   D. rule deny tcp source any

2. 高级ACL可以基于哪些条件来定义规则？

1. C

2. 高级ACL可以基于源/目的IP地址，源/目的端口号，协议类型以及TCP标记值（SYN|ACK|FIN等）等参数来定义规则。

本章总结

- ACL是一种应用非常广泛的网络技术。它的基本原理是：配置了ACL的网络设备根据事先设定好的报文匹配规则对经过该设备的报文进行匹配，然后对匹配上的报文执行事先设定好的处理动作。这些匹配规则及相应的处理动作是根据具体的网络需求而设定的。处理动作的不同以及匹配规则的多样性，使得ACL可以发挥出各种各样的功效。
- ACL技术总是与防火墙、路由策略、QoS、流量过滤等其他技术结合使用。
- 在本章节中，主要介绍了ACL的相关技术知识，包括：ACL的作用，ACL的组成、匹配和分类、通配符的使用方法，以及ACL的基本配置及应用。


谢 谢
www.huawei.com

AAA原理与配置

前言

- 对于任何网络，用户管理都是最基本的安全管理要求之一。
- AAA（Authentication，Authorization，and Accounting）是一种管理框架，它提供了授权部分用户访问指定资源和记录这些用户操作行为的安全机制。因其具有良好的可扩展性，并且容易实现用户信息的集中管理而被广泛使用。AAA可以通过多种协议来实现，在实际应用中，最常使用RADIUS（Remote Authentication Dial-In User Service）协议。
- 本章将介绍AAA基本概念、AAA的实现方式、AAA的基本配置以及常见AAA应用场景。

目标

- 学完本课程后，您将能够：
  - 掌握AAA的基本原理
  - 描述AAA的应用场景
  - 描述RADIUS的基本原理
  - 掌握AAA的基本配置

目录

AAA基本概念

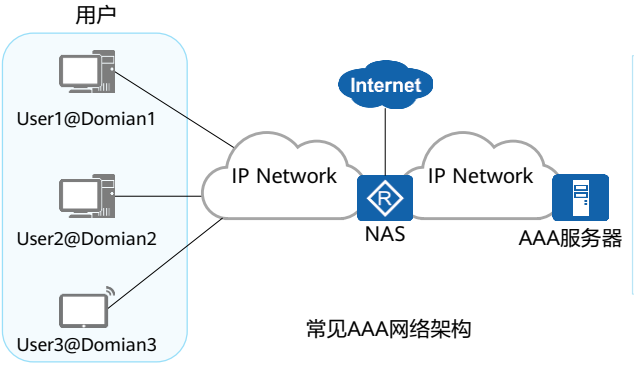
- AAA是Authentication（认证）、Authorization（授权）和Accounting（计费）的简称，是网络安全的一种管理机制，提供了认证、授权、计费三种安全功能。

HUAWEI

- 认证（Authentication）：验证用户是否可以获得访问权，确定哪些用户可以访问网络。
- 授权（Authorization）：授权用户可以使用哪些服务。
- 计费（Accounting）：记录用户使用网络资源的情况。
- 网络运营商（ISP）需要验证家庭宽带用户的账号密码之后才允许其上网，并记录用户的上网时长或上网流量等内容，这就是AAA技术最常见的应用场景。

AAA常见架构

- AAA常见网络架构中包括用户、NAS（Network Access Server）、AAA服务器（AAA Server）。



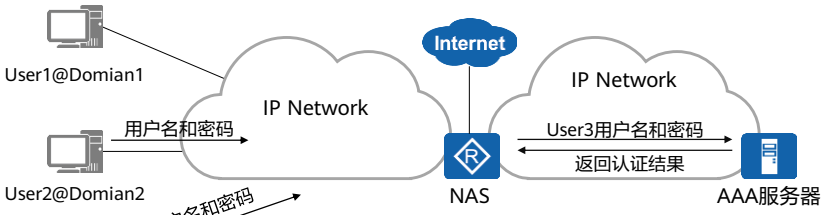
常见AAA网络架构

- NAS负责集中收集和管理用户的访问请求。
- 在NAS上会创建多个域来管理用户。不同的域可以关联不同的AAA方案。AAA方案包含认证方案，授权方案，计费方案。
- 当收到用户接入网络的请求时，NAS会根据用户名来判断用户所在的域，根据该域对应的AAA方案对用户进行管控。

- NAS基于域来对用户进行管理，每个域都可以配置不同的认证、授权和计费方案，用于对该域下的用户进行认证、授权和计费。
- 每个用户都属于某一个域。用户属于哪个域是由用户名中的域名分隔符@后的字符串决定。例如，如果用户名是user1@domain1，则用户属于domain1域。如果用户名后不带有@，则用户属于系统缺省域。

认证（Authentication）

- AAA支持的认证方式有：不认证，本地认证，远端认证。

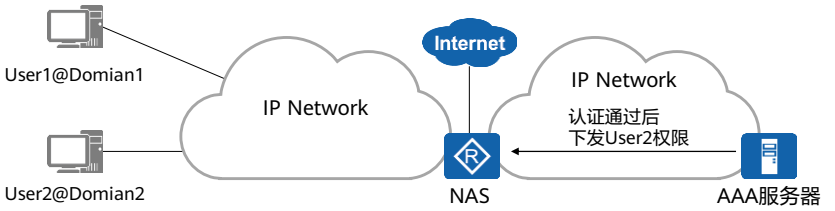


用户	域	认证方式
User1@Domian1	Domain1	不认证
User2@Domian2	Domain2	本地认证
User3@Domian3	Domain3	远端认证

- AAA支持三种认证方式:
  - 不认证：完全信任用户，不对用户身份进行合法性检查。鉴于安全考虑，这种认证方式很少被采用。
  - 本地认证：将本地用户信息（包括用户名、密码和各种属性）配置在NAS上，此时NAS就是AAA Server。本地认证的优点是处理速度快、运营成本低；缺点是存储信息量受设备硬件条件限制。这种认证方式常用于对用户登录设备进行管理，如Telnet，FTP用户等。
  - 远端认证：将用户信息（包括用户名、密码和各种属性）配置在认证服务器上。支持通过RADIUS协议或HWTACACS协议进行远端认证。NAS作为客户端，与RADIUS服务器或HWTACACS服务器进行通信。

授权（Authorization）

- AAA支持的授权方式有：不授权，本地授权，远端授权。
- 授权信息包括：所属用户组、所属VLAN、ACL编号等。

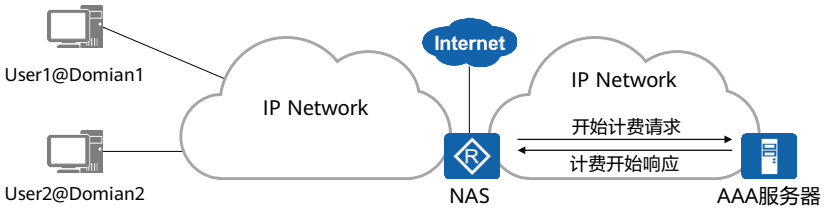


用户	域	授权方式	授权内容
User1@Domian1	Domain1	不授权	无
User2@Domian2	Domain2	本地授权	可以访问Internet
User3@Domian3	Domain3	远端授权	由远端服务器授权

- AAA授权功能赋予用户访问的特定网络或设备的权限。AAA支持以下授权方式：
  - 不授权：不对用户进行授权处理。
  - 本地授权：根据NAS上对应域下的配置进行授权。
  - 远端授权：支持由RADIUS服务器授权或HWTACACS服务器授权。
    - HWTACACS授权，使用HWTACACS服务器对所有用户授权。
    - RADIUS授权，只支持对通过RADIUS服务器认证的用户授权。RADIUS协议的认证和授权是绑定在一起的，不能单独使用RADIUS进行授权。
- 当采用远端授权时，用户可以同时从授权服务器和NAS获取授权信息。NAS配置的授权信息优先级比授权服务器下发的授权信息低。

计费（Accouting）

- 计费功能用于监控授权用户的网络行为和网络资源的使用情况。
- AAA支持的计费方式有：不计费，远端计费。

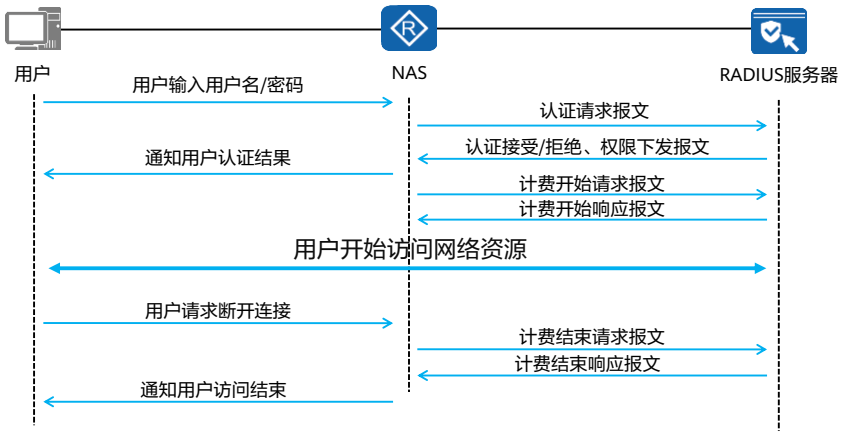


用户	域	计费方式
User1@Domian1	Domain1	不计费
User2@Domian2	Domain2	不计费
User3@Domian3	Domain3	远端计费

- AAA支持以下两种计费方式：
  - 不计费：为用户提供免费上网服务，不产生相关活动日志。
  - 远端计费：支持通过RADIUS服务器或HWTACACS服务器进行远端计费。

AAA实现协议 - RADIUS

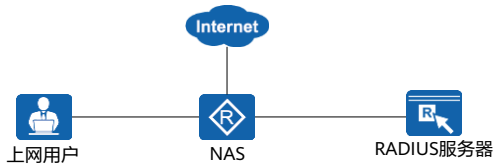
- AAA可以用多种协议来实现，最常用的是RADIUS协议。



- AAA可以用多种协议来实现，最常用的是RADIUS协议。RADIUS是一种分布式的、客户端/服务器结构的信息交互协议，可以实现对用户的认证、计费和授权功能。
- 通常由NAS作为RADIUS客户端，负责传输用户信息到指定的RADIUS服务器，然后根据从服务器返回的信息进行相应处理（如接受/拒绝用户接入）。
- RADIUS服务器一般运行在中心计算机或工作站上，维护相关的用户认证和网络服务访问信息，负责接收用户连接请求并认证用户，然后给客户端返回所有需要的信息（如接受/拒绝认证请求）。RADIUS使用UDP（User Datagram Protocol）作为传输协议，并规定UDP端口1812、1813分别作为认证、计费端口，具有良好的实时性；同时也支持重传机制和备用服务器机制，从而具有较好的可靠性。
- RADIUS客户端与服务器间的消息流程如下：
    1. 当用户接入网络时，用户发起连接请求，向RADIUS客户端（即NAS）发送用户名和密码。
    2. RADIUS客户端向RADIUS服务器发送包含用户名和密码信息的认证请求报文。
    3. RADIUS服务器接收到合法的请求后，完成认证，并把所需的用户授权信息返回给客户端；对于非法的请求，RADIUS服务器返回认证失败的信息给客户端。
    4. RADIUS客户端通知用户认证是否成功。
    5. RADIUS客户端根据接收到的认证结果接入/拒绝用户。如果允许用户接入，则RADIUS客户端向RADIUS服务器发送计费开始请求报文。
    6. RADIUS服务器返回计费开始响应报文，并开始计费。
    7. 用户开始访问网络资源。
    8. 当用户不再想要访问网络资源时，用户发起下线请求，请求停止访问网络资源。
    9. RADIUS客户端向RADIUS服务器提交计费结束请求报文。
    10. RADIUS服务器返回计费结束响应报文，并停止计费。
    11. RADIUS客户端通知用户访问结束，用户结束访问网络资源。

AAA常见应用场景

通过RADIUS提供上网用户的AAA



- 通过在NAS上配置AAA方案，实现NAS与RADIUS服务器的对接。
- 用户在客户端上输入用户名和密码后，NAS可以将这些信息发送至RADIUS服务器进行认证。
- 如果认证通过，则授予用户访问Internet的权限。
- 在用户访问过程中，RADIUS服务器还可以记录用户使用网络资源的情况。

对管理用户进行本地认证和授权

- 在Router上配置本地AAA方案后，当网络管理员登录Router时，Router将网络管理员的的用户名密码等信息，与本地配置的用户名信息进行比对认证。
- 认证通过后，Router将授予网络管理员一定的管理员权限。

目录

HUAWEI

AAA配置（1）

1. 进入AAA视图

```
[Huawei] aaa
```

从系统视图进入AAA视图进行配置

2. 创建认证方案

```
[Huawei-aaa] authentication-scheme authentication-scheme-name
```

创建认证方案并进入相应的认证方案视图

```
[Huawei-aaa-authentication-scheme-name] authentication-mode { hwtacacs | local | radius }
```

配置认证方式，local指定认证方式为本地认证。缺省情况下，认证方式为本地认证。

- **authorization-scheme** *authorization-scheme-name*命令用来配置域的授权方案。缺省情况下，域下没有绑定授权方案。
- **authentication-mode** { **hwtacacs** | **local** | **radius** }命令用来配置当前认证方案使用的认证方式。缺省情况下，认证模式为本地认证方式。

AAA配置（2）

3. 创建domain并绑定认证方案

```
[Huawei-aaa] domain domain-name
```

创建domain并进入相应的domain视图

```
[Huawei-aaa-domain-name] authentication-scheme authentication-scheme-name
```

在相应的domain视图下绑定认证方案

4. 创建用户

```
[Huawei-aaa] local-user user-name  password cipher password
```

创建本地用户，并配置本地用户的密码：

- 如果用户名中带域名分隔符，如@，则认为@前面的部分是用户名，后面部分是域名
- 如果没有@，则整个字符串为用户名，域为默认域

AAA配置（3）

5. 配置用户接入类型

   [Huawei-aaa] **local-user user-name service-type { { terminal | telnet | ftp | ssh | snmp | http } | ppp | none }**

   设置本地用户的接入类型。缺省情况下，本地用户关闭所有的接入类型。

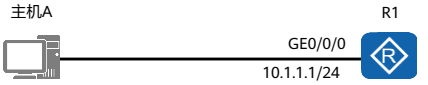
6. 配置用户级别

   [Huawei-aaa] **local-user** *user-name* **privilege level** *level*

   指定本地用户的权限级别。

AAA配置案例

- 在设备R1上配置用户密码和级别，使主机A可以通过配置的用户名和密码远程登录到设备。




```
[R1]aaa
[R1-aaa]local-user huawei password cipher huawei123
[R1-aaa]local-user huawei service-type telnet
[R1-aaa]local-user huawei privilege level 0
[R1]user-interface vty 0 4
[R1-ui-vty0-4]authentication-mode aaa
```

配置验证（1）

- AAA中，每个域都会与相应的认证授权和计费方案相关联，当前为默认域。

```
[R1]display domain name default_admin
  Domain-name:                  default_admin
  Domain-state:                 Active
  Authentication-scheme-name:   default
  Accounting-scheme-name:       default
  Authorization-scheme-name:    -
  Service-scheme-name:          -
  RADIUS-server-template:       -
  HWTACACS-server-template:     -
  User-group:                   -
```


- **display domain** [**name** *domain-name*]命令用来查看域的配置信息。
- **Domain-state**为Active表示激活状态。
- 如果用户名后不带有@，则用户属于系统缺省域，华为设备支持两种缺省域：
  - default域为普通用户的缺省域。
  - default_admin域为管理用户的缺省域。

配置验证 (2)

- 用户正常登录并且下线之后可以看到用户的记录信息。

```
[R1]display aaa offline-record all
  ------------------------------------------------------------------------------
  User name:              huawei
  Domain name:            default_admin
  User MAC:               00e0-fc12-3456
  User access type:       telnet
  User IP address:        10.1.1.2
  User ID:                1
  User login time:        2019/12/28 17:59:10
  User offline time:      2019/12/28 18:00:04
  User offline reason:    user request to offline
```


- **display aaa offline-record**命令用来查看系统中用户下线的记录。

思考题

1. AAA支持的认证、授权和计费方式分别有哪几种?
2. 当创建本地认证的普通用户时，没有关联自定义的域，则该用户属于哪个域?

1. AAA支持的认证方式有：不认证，本地认证，远端认证。AAA支持的授权方式有：不授权，本地授权，远端授权。AAA支持的计费方式有：不计费，远端计费。
2. 如果创建用户时未指定用户所属的域，用户会自动关联缺省域default（管理用户关联到default_admin域）。

本章总结

- AAA技术为了提高企业网络的安全性，防止非法用户登录，需要对企业内部员工，外部客户等进行身份的认证，可访问资源的授权和上网为行为的监控。
  - 认证（Authentication）：验证用户是否可以获得访问权，确定哪些用户可以访问网络。
  - 授权（Authorization）：授权用户可以使用哪些服务。
  - 计费（Accounting）：记录用户使用网络资源的情况。
- AAA技术可以本地实现，也可以通过远端服务器实现。
- AAA可以用多种协议来实现，最常用的是RADIUS协议。


谢 谢
www.huawei.com

网络地址转换

前言

- 随着Internet的发展和网络应用的增多，有限的IPv4公有地址已经成为制约网络发展的瓶颈。为解决这个问题，NAT（Network Address Translation，网络地址转换）技术应需而生。
- NAT技术主要用于实现内部网络的主机访问外部网络。一方面NAT缓解了IPv4地址短缺的问题，另一方面NAT技术让外网无法直接与使用私有地址的内网进行通信，提升了内网的安全性。
- 本章节我们将了解NAT的技术背景， 学习不同类型NAT的技术原理、使用场景。

目标

- 学完本课程后，您将能够：
  - 了解NAT的技术背景
  - 掌握NAT的分类和技术原理
  - 掌握不同场景下如何选用不同类型的NAT技术

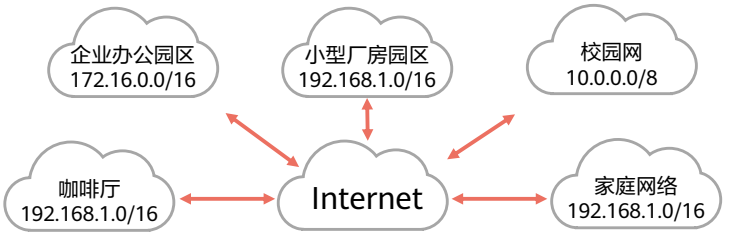
目录

NAT产生背景

- 随着互联网用户的增多，IPv4的公有地址资源显得越发短缺。
- 同时IPv4公有地址资源存在地址分配不均的问题，这导致部分地区的IPv4可用公有地址严重不足。
- 为解决该问题，使用过渡技术解决IPv4公有地址短缺就显得尤为必要。

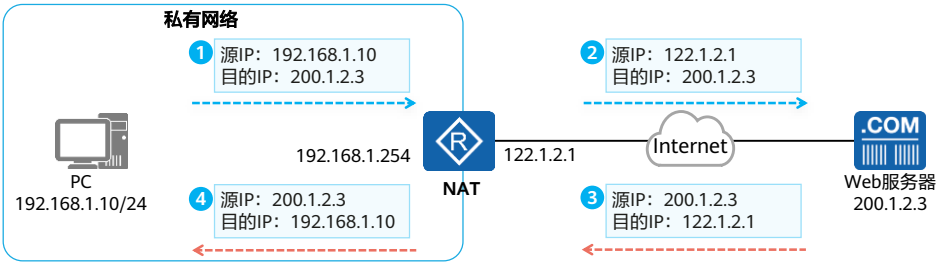
私网IP地址

- 公有地址：由专门的机构管理、分配，可以在Internet上直接通信的IP地址。
- 私有地址：组织和个人可以任意使用，无法在Internet上直接通信，只能在内网使用的IP地址。
- A、B、C类地址中各预留了一些地址专门作为私有IP地址：
  - A类：10.0.0.0 ~ 10.255.255.255
  - B类：172.16.0.0 ~ 172.31.255.255
  - C类：192.168.0.0 ~ 192.168.255.255



NAT技术原理

- NAT：对IP数据报文中的IP地址进行转换，是一种在现网中被广泛部署的技术，一般部署在网络出口设备，例如路由器或防火墙上。
- NAT的典型应用场景：在私有网络内部（园区、家庭）使用私有地址，出口设备部署NAT，对于“从内到外”的流量，网络设备通过NAT将数据包的源地址进行转换（转换成特定的公有地址），而对于“从外到内的”流量，则对数据包的目的地址进行转换。
- 通过私有地址的使用结合NAT技术，可以有效节约公网IPv4地址。



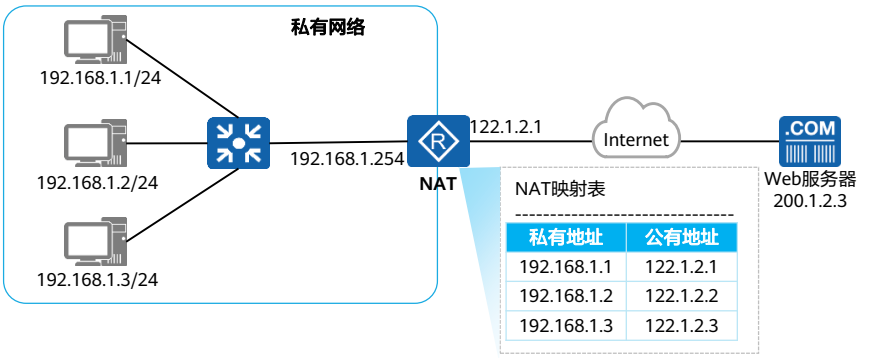
HUAWEI

- 由于私有地址无法在Internet上路由转发，访问Internet的IP数据包将缺乏路由无法到达私有网络出口设备。
- 如果使用了私有地址的私有网络需要访问Internet，必须在网络出口设备配置NAT，将访问Internet的IP数据报文中的私有网络源地址转换成公有网络源地址。

目录

静态NAT原理

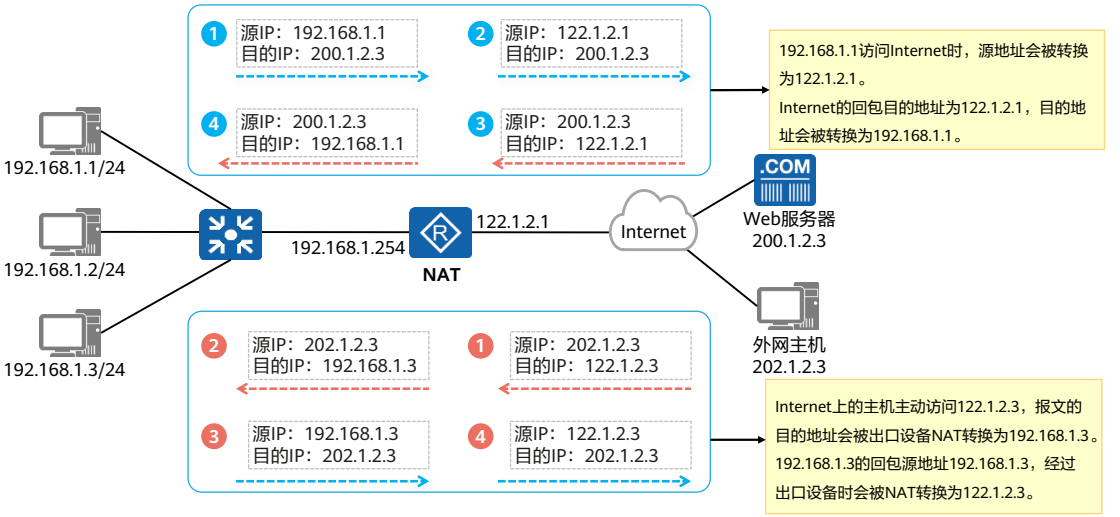
- 静态NAT：每个私有地址都有一个与之对应并且固定的公有地址，即私有地址和公有地址之间的关系是一对一映射。
- 支持双向互访：私有地址访问Internet经过出口设备NAT转换时，会被转换成对应的公有地址。同时，外部网络访问内部网络时，其报文中携带的公有地址（目的地址）也会被NAT设备转换成对应的私有地址。



NAT映射表

私有地址	公有地址
192.168.1.1	122.1.2.1
192.168.1.2	122.1.2.2
192.168.1.3	122.1.2.3

静态NAT转换示例



HUAWEI

静态NAT配置介绍

1. 方式一：接口视图下配置静态NAT

```
[Huawei-GigabitEthernet0/0/0] nat static  global { global-address} inside {host-address }
```

global参数用于配置外部公有地址，inside参数用于配置内部私有地址。

2. 方式二：系统视图下配置静态NAT

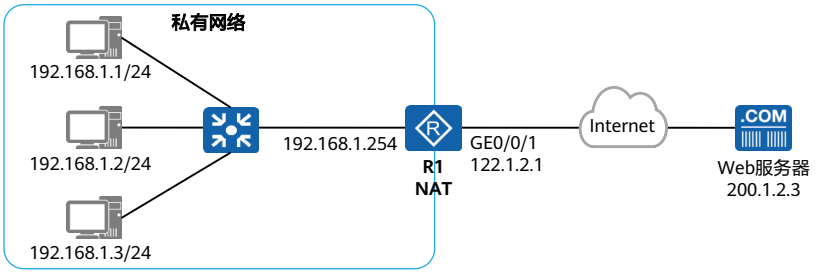
```
[Huawei] nat static  global { global-address} inside {host-address }
```

配置命令相同，视图为系统视图，之后在具体的接口下开启静态NAT。

```
[Huawei-GigabitEthernet0/0/0] nat static enable
```

在接口下使能nat static功能。

静态NAT配置示例



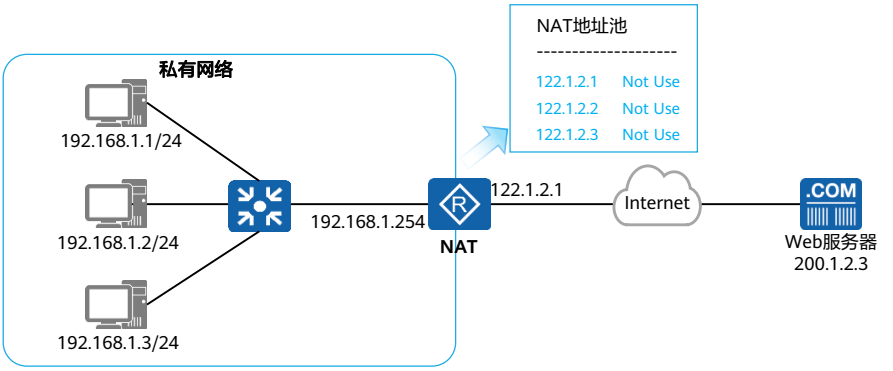
- 在R1上配置静态NAT将内网主机的私有地址一对一映射到公有地址。

```
[R1]interface GigabitEthernet0/0/1
[R1-GigabitEthernet0/0/1]ip address 122.1.2.1 24
[R1-GigabitEthernet0/0/1]nat static global 122.1.2.1 inside 192.168.1.1
[R1-GigabitEthernet0/0/1]nat static global 122.1.2.2 inside 192.168.1.2
[R1-GigabitEthernet0/0/1]nat static global 122.1.2.3 inside 192.168.1.3
```

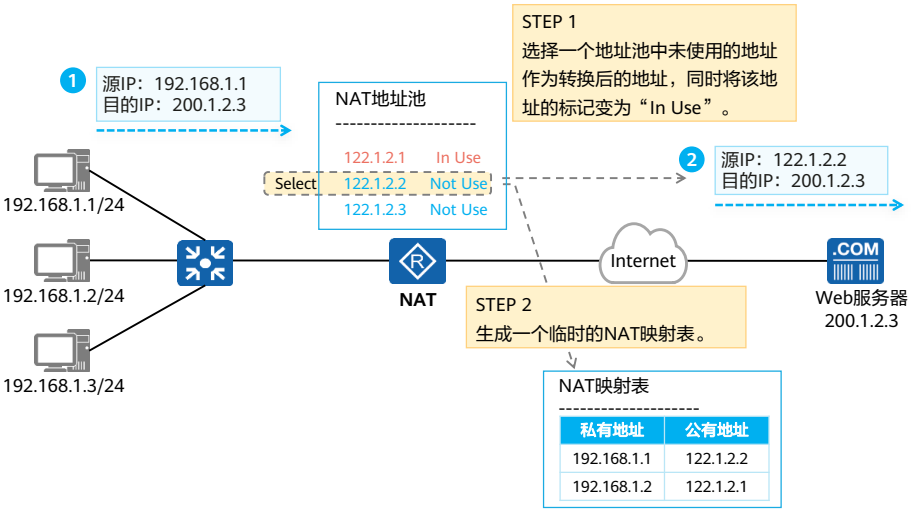
目录

动态NAT原理

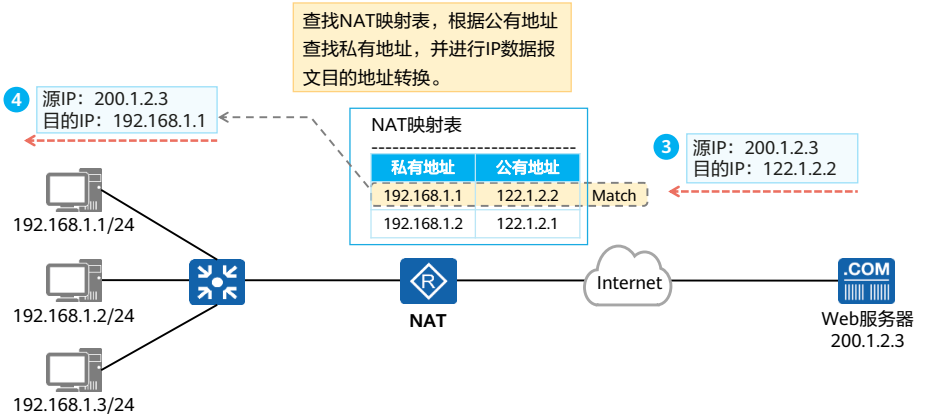
- 动态NAT：静态NAT严格地一对一进行地址映射，这就导致即便内网主机长时间离线或者不发送数据时，与之对应的公有地址也处于使用状态。为了避免地址浪费，动态NAT提出了地址池的概念：所有可用的公有地址组成地址池。
- 当内部主机访问外部网络时临时分配一个地址池中未使用的地址，并将该地址标记为“In Use”。当该主机不再访问外部网络时回收分配的地址，重新标记为“Not Use”。



动态NAT转换示例 (1)



动态NAT转换示例 (2)



HUAWEI

动态NAT配置介绍

1. 创建地址池

   [Huawei] **nat address-group** *group-index start-address end-address*

   配置公有地址范围，其中group-index为地址池编号，start-address、end-address分别为地址池起始地址、结束地址。

2. 配置地址转换的ACL规则

   [Huawei] **acl** *number*
   [Huawei-acl-basic-*number*] **rule permit source** *source-address source-wildcard*

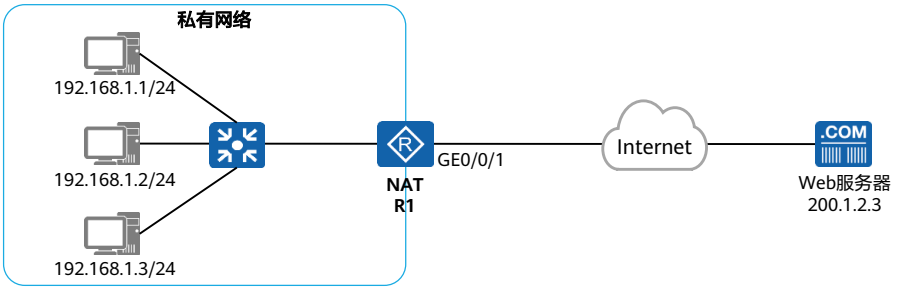
   配置基础ACL，匹配需要进行动态转换的源地址范围。

3. 接口视图下配置带地址池的NAT Outbound

   [Huawei-GigabitEthernet0/0/0] **nat outbound** *acl-number* **address-group** *group-index* [**no-pat**]

   接口下关联ACL与地址池进行动态地址转换，no-pat参数指定不进行端口转换。

动态NAT配置示例



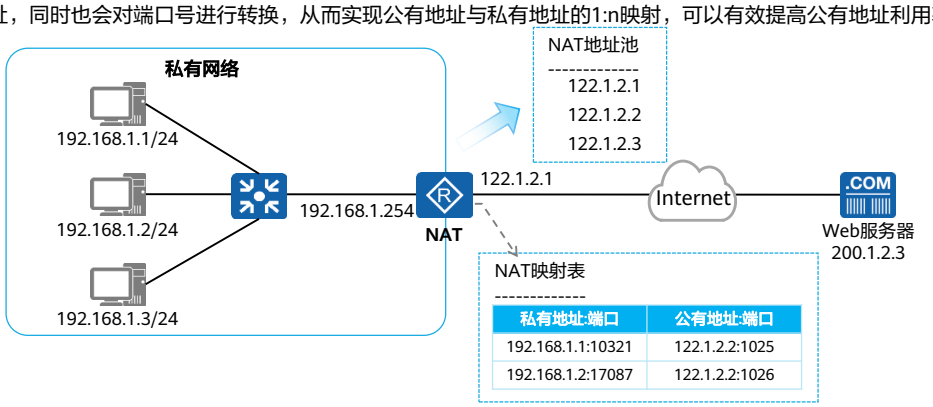
- 在R1上配置动态NAT将内网主机的私有地址动态映射到公有地址。

```
[R1]nat address-group 1 122.1.2.1 122.1.2.3
[R1]acl 2000
[R1-acl-basic-2000]rule 5 permit source 192.168.1.0 0.0.0.255
[R1-acl-basic-2000]quit
[R1]interface GigabitEthernet0/0/1
[R1-GigabitEthernet0/0/1]nat outbound 2000 address-group 1 no-pat
```

目录

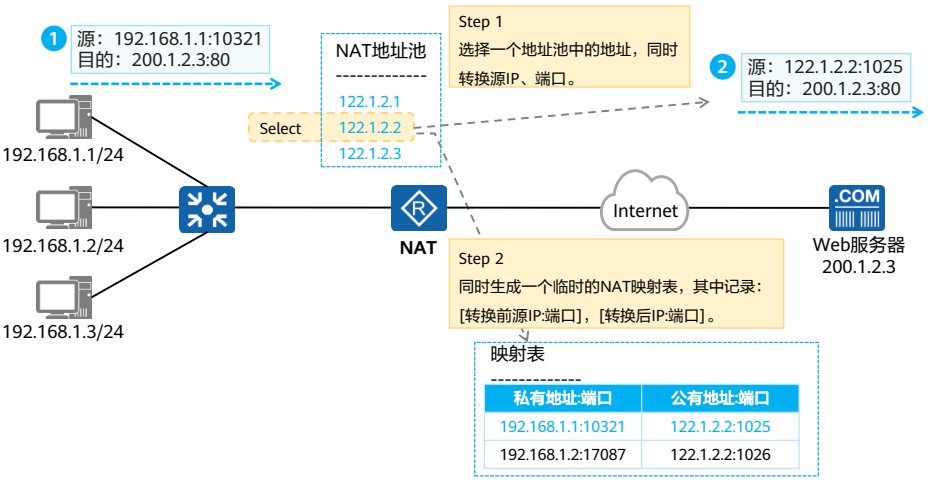
NAPT原理

- 动态NAT选择地址池中的地址进行地址转换时不会转换端口号，即No-PAT（No-Port Address Translation，非端口地址转换），公有地址与私有地址还是1:1的映射关系，无法提高公有地址利用率。
- NAPT（Network Address and Port Translation，网络地址端口转换）：从地址池中选择地址进行地址转换时不仅转换IP地址，同时也会对端口号进行转换，从而实现公有地址与私有地址的1:n映射，可以有效提高公有地址利用率。

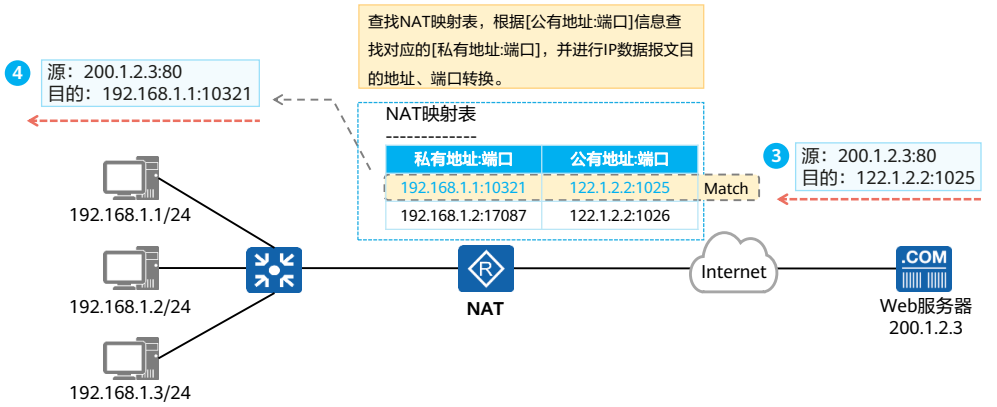


- NAPT借助端口可以实现一个公有地址同时对应多个私有地址。该模式同时对IP地址和传输层端口进行转换，实现不同私有地址（不同的私有地址，不同的源端口）映射到同一个公有地址（相同的公有地址，不同的源端口）。

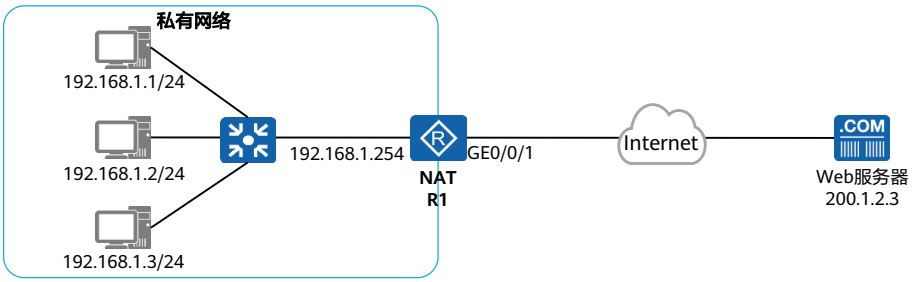
NAPT转换示例（1）



NAPT转换示例 (2)



NAPT配置示例

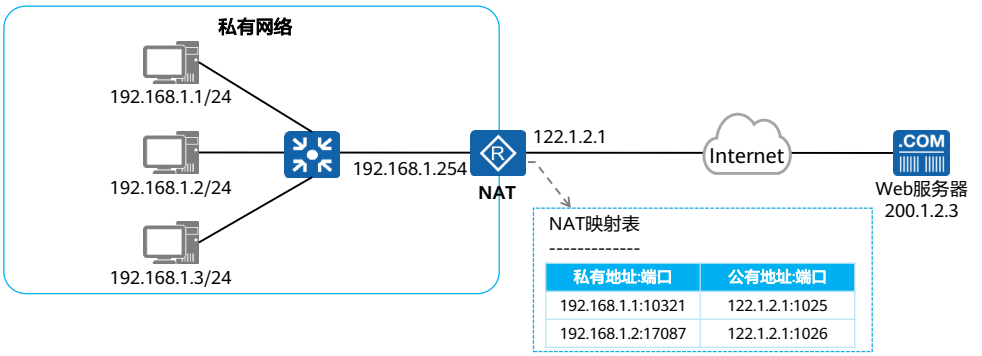


- 在R1上配置NAPT让内网所有私有地址通过122.1.2.1访问公网。

```
[R1]nat address-group 1 122.1.2.1 122.1.2.1
[R1]acl 2000
[R1-acl-basic-2000]rule 5 permit source 192.168.1.0 0.0.0.255
[R1-acl-basic-2000]quit
[R1]interface GigabitEthernet0/0/1
[R1-GigabitEthernet0/0/1]nat outbound 2000 address-group 1
```

Easy IP

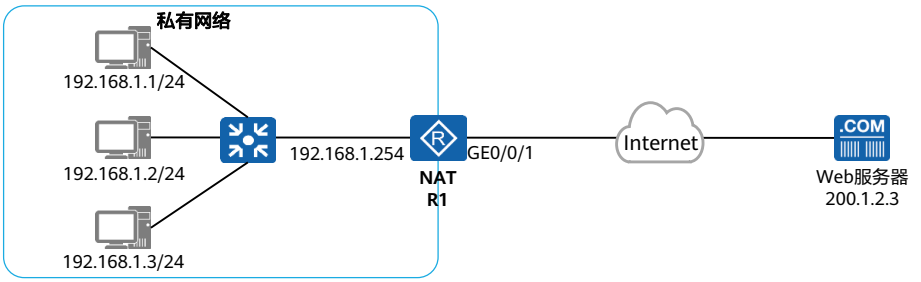
- Easy IP：实现原理和NAPT相同，同时转换IP地址、传输层端口，区别在于Easy IP没有地址池的概念，使用接口地址作为NAT转换的公有地址。
- Easy IP适用于不具备固定公网IP地址的场景：如通过DHCP、PPPoE拨号获取地址的私有网络出口，可以直接使用获取到的动态地址进行转换。



HUAWEI

- DHCP：Dynamic Host Configuration Protocol ，动态主机配置协议
- PPPoE：Point-to-Point Protocol over Ethernet ，以太网承载PPP协议

Easy IP配置示例



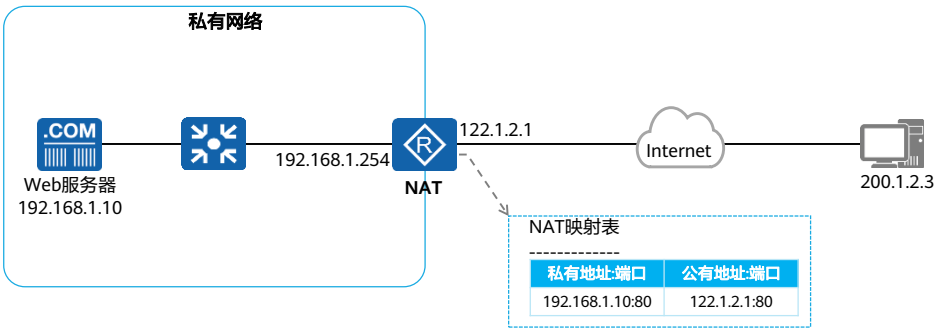
- 在R1上配置Easy-IP让内网所有私有地址通过122.1.2.1访问公网。

```
[R1]acl 2000
[R1-acl-basic-2000]rule 5 permit source 192.168.1.0 0.0.0.255
[R1-acl-basic-2000]quit
[R1]interface GigabitEthernet0/0/1
[R1-GigabitEthernet0/0/1]nat outbound 2000
```

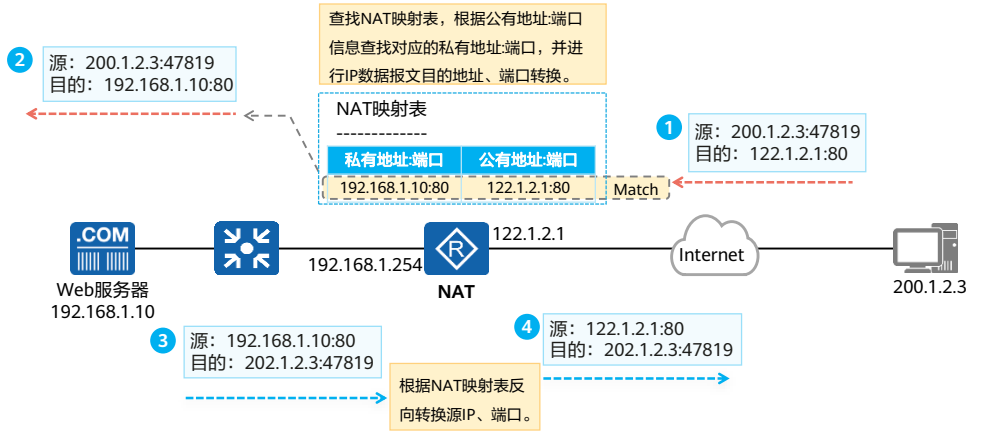
目录

NAT Server使用场景

- NAT Server：指定[**公有地址:端口**]与[**私有地址:端口**]的一对一映射关系，将内网服务器映射到公网，当私有网络中的服务器需要对公网提供服务时使用。
- 外网主机主动访问[**公有地址:端口**]实现对内网服务器的访问。

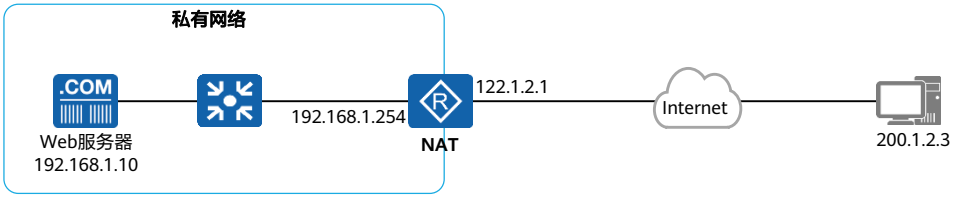


NAT Server转换示例



HUAWEI

NAT Server配置示例



- 在R1上配置NAT Server将内网服务器192.168.1.10的8080端口映射到公有地址122.1.2.1的80端口。

```
[R1]interface GigabitEthernet0/0/1
[R1-GigabitEthernet0/0/1]ip address 122.1.2.1 24
[R1-GigabitEthernet0/0/1]nat server protocol tcp global 122.1.2.1 www inside 192.168.1.10 8080
```

思考题

1. 何种NAT转换可以让外部网络主动访问内网服务器?
2. NAPT相比较于No-PAT有哪些优点?

1. 静态NAT、NAT server都可以。静态NAT实现了双向互访，所以自然容许外部网络对内网服务器的访问。NAT Server的场景本身就是让外部网络主动访问内部服务器。
2. NAPT支持多个私有地址转换为一个共同的公有地址，公有地址利用率更高。

本章总结

- 在私有网络内使用私有地址，并在网络出口使用NAT技术，可以有效减少网络所需的IPv4公有地址数目，NAT技术有效地缓解了IPv4公有地址短缺的问题。
- 动态NAT、NAPT、Easy IP为私网主机访问公网提供源地址转换。
- NAT Server实现了内网主机对公网提供服务。
- 静态NAT提供了一对一映射，支持双向互访。


谢 谢
www.huawei.com

网络服务与应用

前言

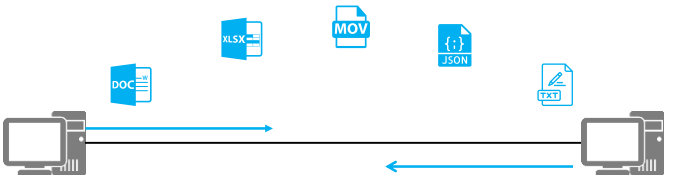
- 网络已经成为当今人们生活中的一部分：传输文件、发送邮件、在线视频、浏览网页、联网游戏。因为网络分层模型的存在，普通用户无需关注通信实现原理等技术细节就可以直接使用由应用层提供的各种服务。
- 之前的课程中我们已经学习了数据链路层、网络层、传输层相关的技术，本章让我们一起了解FTP、DHCP、HTTP等常见的网络服务与应用。

目标

- 学完本课程后，您将能够：
  - 掌握FTP的工作原理
  - 掌握TFTP的工作原理
  - 掌握DHCP的工作原理
  - 掌握Telnet的工作原理
  - 掌握HTTP的工作原理
  - 掌握DNS的基本概念
  - 掌握NTP的基本概念

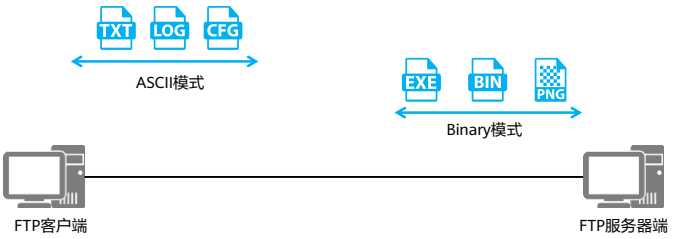
目录

文件传输协议



- 主机之间传输文件是IP网络的一个重要功能，如今人们可以方便地使用网页、邮箱进行文件传输。
- 然而在互联网早期，Web（World Wide Web，万维网）还未出现，操作系统使用命令行的时代，用户使用命令行工具进行文件传输。其中最通用的方式就是使用FTP（File Transfer Protocol，文件传输协议）以及TFTP（Trivial File Transfer Protocol，简单文件传输协议）。

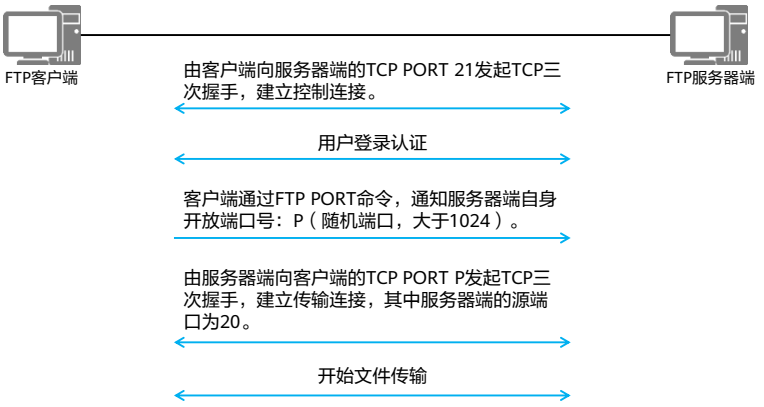
FTP基本概念



- FTP采用典型的C/S架构（即服务器端与客户端模型），客户端与服务器端建立TCP连接之后即可实现文件的上传、下载。
- 针对传输的文件类型不同，FTP可以采用不同的传输模式：
  - ASCII模式：传输文本文件（TXT、LOG、CFG ）时会对文本内容进行编码方式转换，提高传输效率。当传输网络设备的配置文件、日志文件时推荐使用该模式。
  - Binary（二进制）模式：非文本文件（cc、BIN、EXE、PNG），如图片、可执行程序等，以二进制直接传输原始文件内容。当传输网络设备的版本文件时推荐使用该模式。

- FTP传输数据时支持两种传输模式：ASCII模式和Binary模式。
- ASCII模式用于传输文本文件。发送端的字符在发送前被转换成ASCII码格式之后进行传输，接收端收到之后再将其转换成字符。二进制模式常用于发送图片文件和程序文件，发送端在发送这些文件时无需转换格式即可传输。
- cc：VRP版本文件。

FTP传输过程 - 主动模式

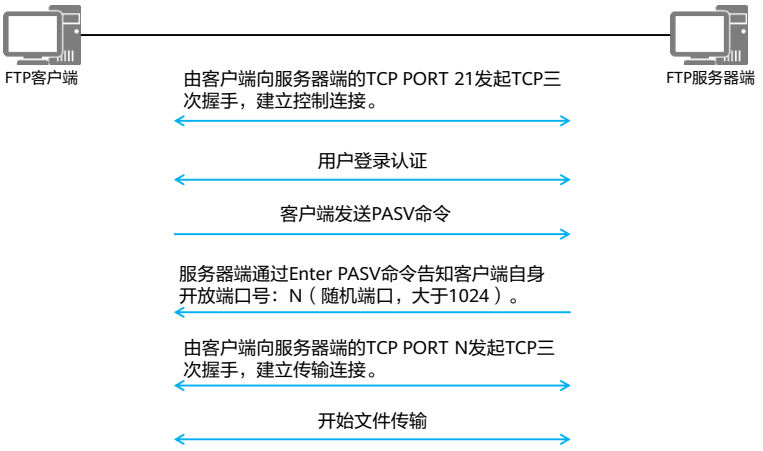


- FTP存在两种工作方式：主动模式（PORT）和被动模式（PASV）。

HUAWEI

- 使用主动模式时，FTP客户端使用一个随机端口（一般大于1024）向FTP服务器端的端口21发送连接请求；FTP服务器端接受请求，建立一条控制连接来传输控制消息。同时FTP客户端开始监听另一随机端口P（一般大于1024），并使用PORT命令通知FTP服务器端。当需要传输数据时，FTP服务器端从端口20向FTP客户端的端口P发送连接请求，建立一条传输连接来传输数据。

FTP传输过程 - 被动模式



HUAWEI

- 当使用被动模式时，FTP客户端使用一个随机端口（一般大于1024）向FTP服务器端的端口21发送连接请求，FTP服务器端接受请求，建立一条控制连接来传输控制消息。同时FTP客户端开始监听另一随机端口P（一般大于1024），并使用PASV命令通知FTP服务器端，FTP服务器端接到PASV 命令后，开启一个随机端口N（一般大于1024），并使用Enter PASV命令告知客户端自身开放端口号。当需要传输数据时，FTP客户端从端口P向FTP服务器端N端口发送连接请求，建立一条传输连接来传输数据。
- 主动模式和被动模式建立数据连接方式完全不同，在实际使用中各有利弊：
  - 使用主动模式传输数据时，如果FTP客户端在私有网络中并且FTP客户端和FTP服务器端之间存在NAT设备，那么FTP服务器端收到的PORT报文中携带的端口号、IP地址并不是FTP客户端经过NAT转换之后的地址、端口号，因此服务器端无法向PORT报文中携带的私网地址发起TCP连接（此时，客户端的私网地址在公有网络中路由不可达）。
  - 使用被动模式传输数据时，FTP客户端主动向服务器端的一个开放端口发起连接，如果FTP服务器端在防火墙内部区域中，并且没有放通客户端所在区域到服务器端所在区域的主动访问，那么这个连接将无法建立成功，从而导致FTP无法正常传输。

配置命令介绍 - 设备作为服务器端

用户通过FTP访问设备

1. 开启FTP服务器端功能

```
[Huawei]ftp [ ipv6 ] server enable
```

缺省情况下，设备的FTP服务器端功能是关闭的。

2. 配置FTP本地用户

```
[Huawei]aaa
[Huawei]local-user user-name password irreversible-cipher password
[Huawei]local-user user-name privilege level level
[Huawei]local-user user-name service-type ftp
[Huawei]local-user user-name ftp-directory directory
```

必须将用户级别配置在3级或者3级以上，否则FTP连接将无法成功。

配置命令介绍 - 设备作为客户端

1. VRP作为FTP客户端访问FTP服务器端

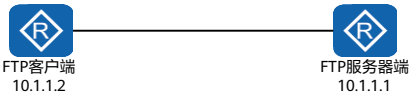
```
<FTP Client>ftp 10.1.1.1
Trying 10.1.1.1 ...
Press CTRL+K to abort
Connected to 10.1.1.1.
220 FTP service ready.
User(10.1.1.1:(none)):ftp
331 Password required for ftp.
Enter password:
230 User logged in.
```

2. VRP作为FTP客户端的常用命令

```
ascii      Set the file transfer type to ASCII, and it is the default type
binary     Set the file transfer type to support the binary image
ls         List the contents of the current or remote directory
passive    Set the toggle passive mode, the default is on
get        Download the remote file to the local host
put        Upload a local file to the remote host
```


- VRP（Versatile Routing Platform，通用路由平台）
  - ascii　Set the file transfer type to ASCII, and it is the default type：切换到ASCII传输模式。
  - binary　Set the file transfer type to support the binary image：切换到Binary传输模式。
  - ls　List the contents of the current or remote directory：查看FTP服务器端上的文件列表，也可以使用dir 。
  - passive　Set the toggle passive mode, the default is on：使用被动模式，undo passive使用主动模式。
  - get　Download the remote file to the local host：下载FTP服务器端的文件到本地。
  - put　Upload a local file to the remote host：上传本地文件到FTP服务器端 。

配置示例



- 上述两台路由器，一台作为FTP服务器端，一台作为FTP客户端。
- 首先通过配置，在FTP服务器端上开启FTP服务，创建一个账号作为FTP登录使用账号。然后FTP客户端登录FTP服务器端并使用get命令下载一个文件。

FTP服务器端配置如下：

```
<Huawei> system-view
[Huawei] sysname FTP_Server
[FTP_Server] ftp server enable
[FTP_Server] aaa
[FTP_Server-aaa] local-user admin1234 password irreversible-cipher
Helloworld@6789
[FTP_Server-aaa] local-user admin1234 privilege level 15
[FTP_Server-aaa] local-user admin1234 service-type ftp
[FTP_Server-aaa] local-user admin1234 ftp-directory flash:
```

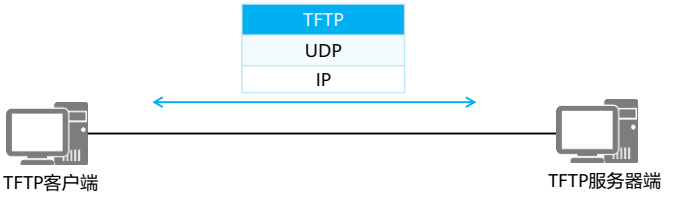
FTP客户端操作示例：

```
<FTP Client>ftp 10.1.1.1
[FTP Client-ftp]get sslvpn.zip
200 Port command okay.
FTP: 828482 byte(s) received in 2.990 second(s) 277.08Kbyte(s)/sec.
```

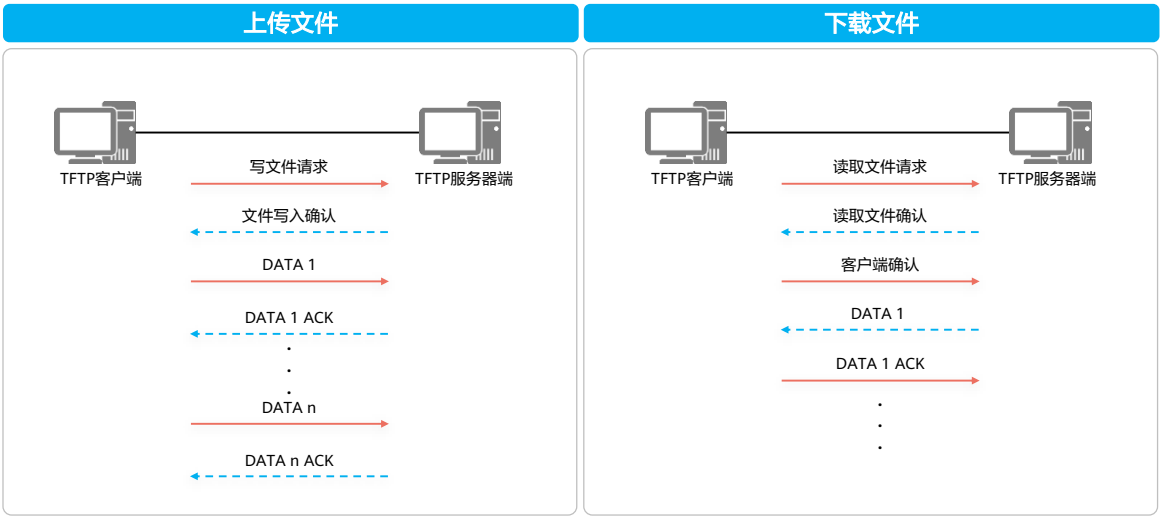
目录

TFTP基础

- 相较于FTP，TFTP的设计就是以传输小文件为目标，协议实现就简单很多：
  - 使用UDP进行传输（端口号69）
  - 无需认证
  - 只能直接向服务器端请求某个文件或者上传某个文件，无法查看服务器端的文件目录。



TFTP传输示例



HUAWEI

- TFTP存在5种报文格式：
  - RRQ：读请求包。
  - WRQ：写请求包。
  - DATA：数据传输报文。
  - ACK：应答包，用于确认收到对端的报文。
  - ERROR：差错控制报文。

配置命令介绍 - 设备作为客户端

1. VRP作为TFTP客户端下载文件

```
<Huawei> tftp TFTP_Server-IP-address get filename
```

TFTP无需登录，直接输入服务器端IP地址以及操作命令即可。

2. VRP作为TFTP客户端上传文件

```
<Huawei> tftp TFTP_Server-IP-address put filename
```

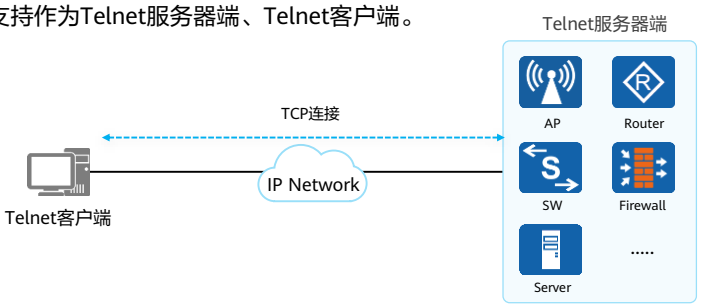
TFTP无需登录，直接输入服务器端IP地址以及操作命令即可。

目前VRP设备只支持作为TFTP客户端。

目录

Telnet应用场景

- 为方便通过命令行管理设备，可以使用Telnet协议对设备进行管理。
- Telnet协议与使用Console接口管理设备不同，无需专用线缆直连设备的Console接口，只要IP地址可达、能够和设备的TCP 23端口通信即可。
- 支持通过Telnet协议进行管理的设备被称为Telnet服务器端，而对应的终端则被称为Telnet客户端。很多网络设备同时支持作为Telnet服务器端、Telnet客户端。

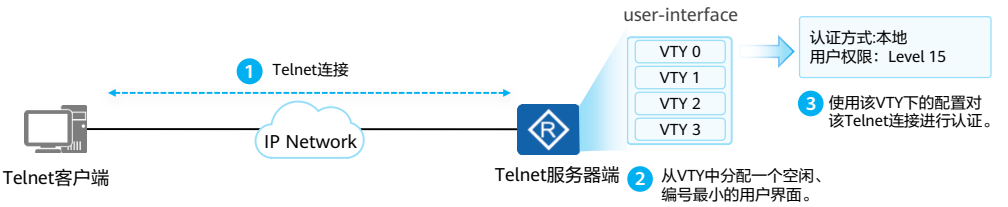


HUAWEI

- 目前主流的网络设备，如：AC（Access Controller，无线控制器）、AP（Access Point，接入点）、Firewall、Router、Switch、Server等都支持作为Telnet服务器端，同时也基本都支持作为Telnet客户端。

虚拟用户界面

- 当用户使用Console接口、Telnet等方式登录设备的时候，系统会分配一个用户界面（user-interface）来管理、监控设备与用户间的当前会话，每个用户界面视图可以配置一系列参数用于指定用户的认证方式、登录后的权限级别，当用户登录设备后将会受这些参数限制。
- Telnet所对应的用户界面类型为VTY（Virtual Type Terminal，虚拟类型终端）。



HUAWEI

配置命令介绍 (1)

1. 开启Telnet服务器端功能

   [Huawei] **telnet server enable**

   使能设备的Telnet服务器端功能。缺省情况下，设备的Telnet服务器端功能处于去使能状态，undo telnet server enable即可重新关闭Telnet服务器端功能。

2. 进入用户视图

   [Huawei] **user-interface vty** *first-ui-number [last-ui-number]*

   进入VTY用户界面视图。 不同设备型号的VTY接口可能并不一致。

3. 配置VTY用户界面支持的协议

   [Huawei-ui-vty0-4]] **protocol inbound** { **all** | **telnet** }

   缺省情况下，VTY用户界面支持的协议是SSH（Secure Shell Protocol ，安全外壳协议）和Telnet。

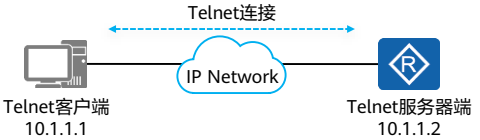
配置命令介绍 (2)

4. 配置认证方式以及密码认证方式下的认证密码

```
[Huawei-ui-vty0-4] authentication-mode {aaa | none | password}
[Huawei-ui-vty0-4] set authentication password cipher
```

缺省情况下，无默认认证方式，需要进行手动配置。
不同VRP版本执行set authentication password cipher命令有差异：某些版本需要回车后输入密码，某些版本可直接在命令后输入密码。

配置示例（1）

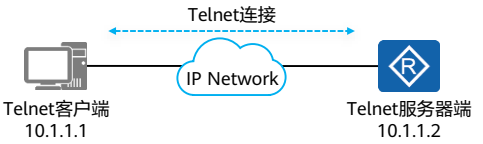


- 配置10.1.1.2作为Telnet服务器端，认证方式为AAA本地认证。在本地创建名为huawei的账号，密码为Huawei@123，权限为15级。
- 用户通过Telnet客户端软件登录并管理Telnet服务器端。

Telnet服务器端配置如下：

```
<Huawei> system-view
[Huawei] telnet server enable
[Huawei] aaa
[Huawei-aaa] local-user huawei password irreversible-cipher
Huawei@123
[Huawei-aaa] local-user huawei privilege level 15
[Huawei-aaa] local-user huawei service-type telnet
[Huawei-aaa] quit
[Huawei] user-interface vty 0 4
[Huawei-ui-vty0-4] authentication-mode aaa
```

配置示例 (2)



- 配置10.1.1.2作为Telnet服务器端，认证方式为AAA本地认证。在本地创建名为huawei的账号，密码为Huawei@123，权限为15级。
- 用户通过Telnet客户端软件登录并管理Telnet服务器端。

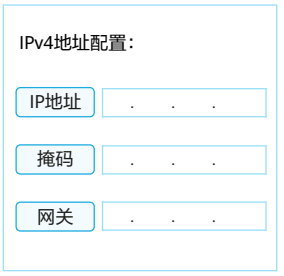
Telnet 客户端操作：

```
<Host>telnet 10.1.1.2
Login authentication

Username:huawei
Password:
Info: The max number of VTY users is 5, and the number
      of current VTY users on line is 1.
      The current login time is 2020-01-08 15:37:25.
<Huawei>
```

目录

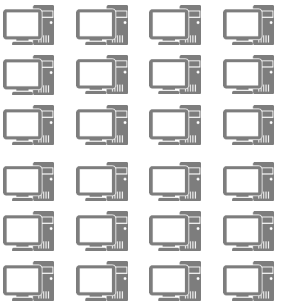
手动配置网络参数的问题（1）

参数多、理解难



- 普通用户对于网络参数不了解，经常配置错误，导致无法正常访问网络。随意配置IP地址导致地址冲突更是时常发生。

工作量大



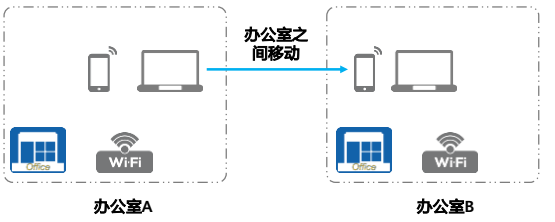
- 交由网络管理员统一配置，工作量巨大，属于重复性劳动。
- 网络管理员需要提前对IP地址进行规划、分配到个人。

手动配置网络参数的问题 (2)

利用率低

- 企业网中每个人固定使用一个IP地址，IP地址利用率低，有些地址可能长期处于未使用状态。

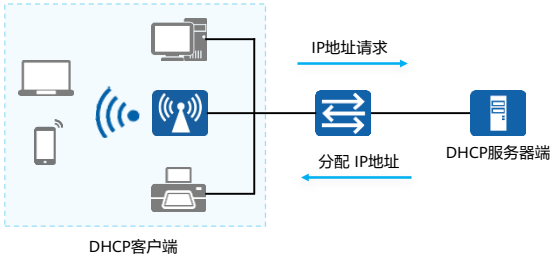
灵活性差



- WLAN（Wireless Local Area Network，无线局域网）的出现使终端位置不再固定，当无线终端移动到另外一个无线覆盖区域时，可能需要再次配置IP地址。

DHCP基本概念

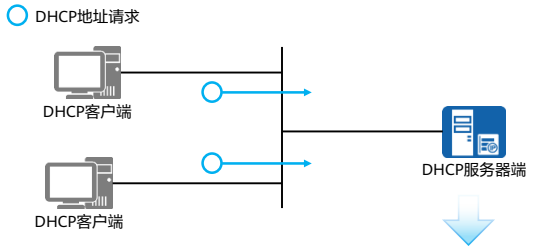
DHCP工作示意图



- 为解决传统的静态手工配置方式的不足，DHCP（Dynamic Host Configuration Protocol，动态主机配置协议）应运而生，其可以实现网络动态合理地分配IP地址给主机使用。
- DHCP采用C/S构架，主机无需配置，从服务器端获取地址，可实现接入网络后即插即用。

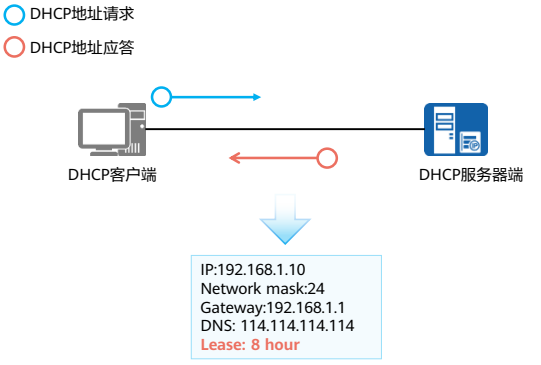
DHCP优点

统一管理



- IP地址由从服务器端的地址池中获取，服务器端会记录维护IP地址的使用状态，做到IP地址统一分配、管理。

地址租期

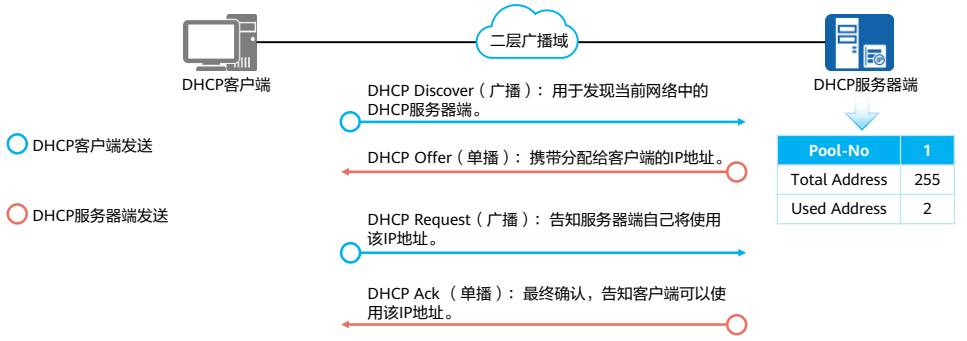


- DHCP提出了租期的概念，可有效提高地址利用率。

HUAWEI

- 对于已分配的IP地址，若终端超过租期仍未续租，服务器端判断该终端不再需要使用该IP地址，将IP地址回收，可继续分配给其他终端。

DHCP工作原理

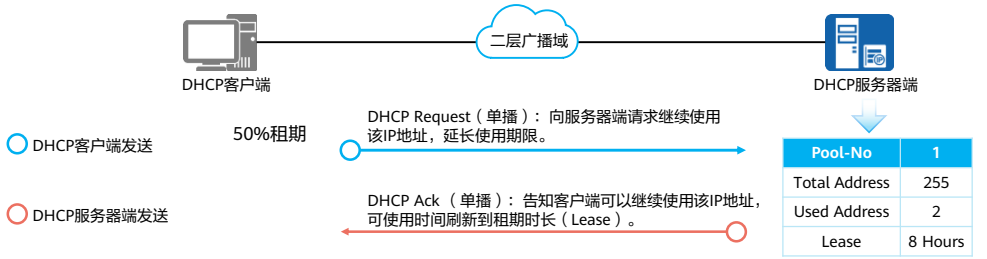


- 思考：为什么DHCP客户端收到Offer之后不直接使用该IP地址，还需要发送一个Request告知服务器端?

HUAWEI

- 广播的Request报文让网络中其他DHCP服务器端得知客户端已选择了某个服务器端分配的IP地址，保证其他服务器端可以释放通过单播Offer分配给该客户端的IP地址。

DHCP租期更新



- 如果在50%租期时客户端未得到原服务器端的回应，则客户端在87.5%租期时会广播发送DHCP Request，任意一台DHCP服务器端都可回应，该过程称为重绑定。

配置命令介绍 (1)

1. 开启DHCP功能

   [Huawei] **dhcp enable**

2. 开启接口采用接口地址池的DHCP服务器端功能

   [Huawei-Gigabitthernet0/0/0]**dhcp select interface**

3. 指定接口地址池下的DNS服务器地址

   [Huawei-Gigabitthernet0/0/0]**dhcp server dns-list** *ip-address*

4. 配置接口地址池中不参与自动分配的IP地址范围

   [Huawei-Gigabitthernet0/0/0]**dhcp server excluded-ip-address** *start-ip-address [end-ip-address]*

5. 配置DHCP服务器接口地址池中IP地址的租用有效期限功能

   [Huawei-Gigabitthernet0/0/0]**dhcp server lease** { **day** *day* [**hour** *hour* [**minute** *minute*]] | **unlimited** }

   缺省情况下，IP地址的租期为1天。

配置命令介绍 (2)

6. 创建全局地址池

   [Huawei]**ip pool** *ip-pool-name*

7. 配置全局地址池可动态分配的IP地址范围

   [Huawei-ip-pool-2]**network ip-address [mask {** *mask* | *mask-length* **}]**

8. 配置DHCP客户端的网关地址

   [Huawei-ip-pool-2]**gateway-list** *ip-address*

9. 配置DHCP客户端使用的DNS服务器的IP地址

   [Huawei-ip-pool-2]**dns-list** *ip-address*

10. 配置IP地址租期

    [Huawei-ip-pool-2] **lease { day** *day* **[hour** *hour* **[minute** *minute* **]] | unlimited }**

11. 使能接口的DHCP服务器功能

    [Huawei-Gigabitthernet0/0/0]**dhcp select global**

DHCP接口地址池配置

需求描述：

- 配置一台路由器作为DHCP服务器端，使用接口GE0/0/0所属的网段作为DHCP客户端的地址池，同时将接口地址设为DNS Server地址，租期设置为3天。

DHCP服务器端配置如下：

```
[Huawei]dhcp enable
[Huawei]interface GigabitEthernet0/0/0
[Huawei-GigabitEthernet0/0/0]dhcp select interface
[Huawei-GigabitEthernet0/0/0]dhcp server dns-list 10.1.1.2
[Huawei-GigabitEthernet0/0/0]dhcp server excluded-ip-address 10.1.1.2
[Huawei-GigabitEthernet0/0/0]dhcp server lease day 3
```

全局使能DHCP服务，进入接口视图下，关联当前接口到DHCP地址池，在接口视图下配置DNS地址、排除地址（将接口自身地址排除在外），同时配置给客户端分配IP地址的租期。

DHCP全局地址池配置

需求描述：

- 配置一台路由器作为DHCP服务器端，配置全局地址池ip pool 2为DHCP客户端分配IP地址；分配地址为1.1.1.0/24网段，网关地址1.1.1.1，DNS地址同样也是1.1.1.1，租期10天，在GE0/0/0接口下调用全局地址池。

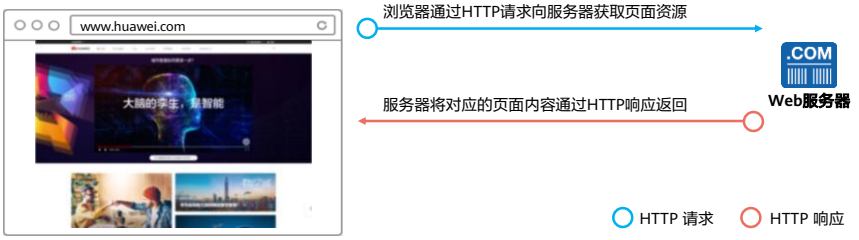
DHCP服务器端配置如下：

```
[Huawei]dhcp enable
[Huawei]ip pool pool2
Info: It's successful to create an IP address pool.
[Huawei-ip-pool-pool2]network 1.1.1.0 mask 24
[Huawei-ip-pool-pool2]gateway-list 1.1.1.1
[Huawei-ip-pool-pool2]dns-list 1.1.1.1
[Huawei-ip-pool-pool2]lease day 10
[Huawei-ip-pool-pool2]quit
[Huawei]interface GigabitEthernet0/0/0
[Huawei-GigabitEthernet0/0/0]dhcp select global
```

- 全局使能DHCP服务，配置全局地址池pool2。在pool2中配置地址池范围、网关地址、DNS地址、租期。
- 最后在具体的接口中配置选择全局地址池。当GE0/0/0收到DHCP请求就会从全局地址池中进行IP地址分配。

目录

使用浏览器访问网页

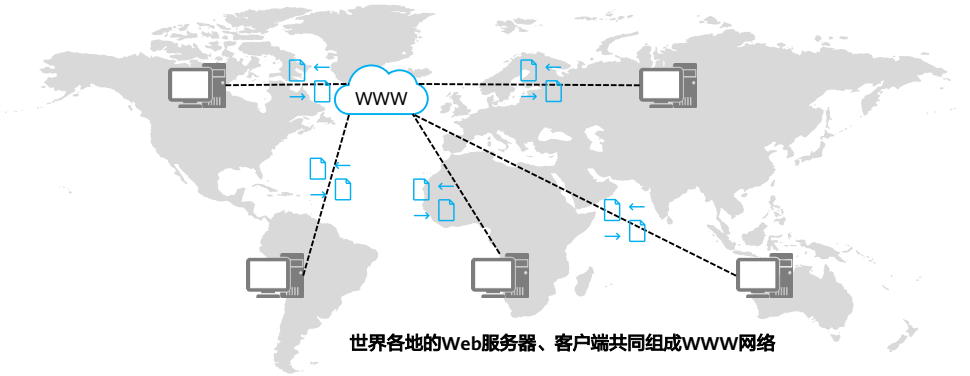


- 当我们在浏览器中输入URL（Uniform Resource Locator，统一资源定位符）时，浏览器就可以从某处获取内容，并将页面内容显示在浏览器中。
- HTTP（Hypertext Transfer Protocol，超文本传输协议）：客户端浏览器或其他程序与Web服务器之间的应用层通信协议。
- HTTP是典型的C/S构架应用，作为应用层协议使用TCP进行传输。

HUAWEI

- URL：唯一标识Internet上网页和其他资源位置的地址，可以包括如超文本网页（扩展名通常为.html或.htm）名称之类的详细信息。

诞生背景

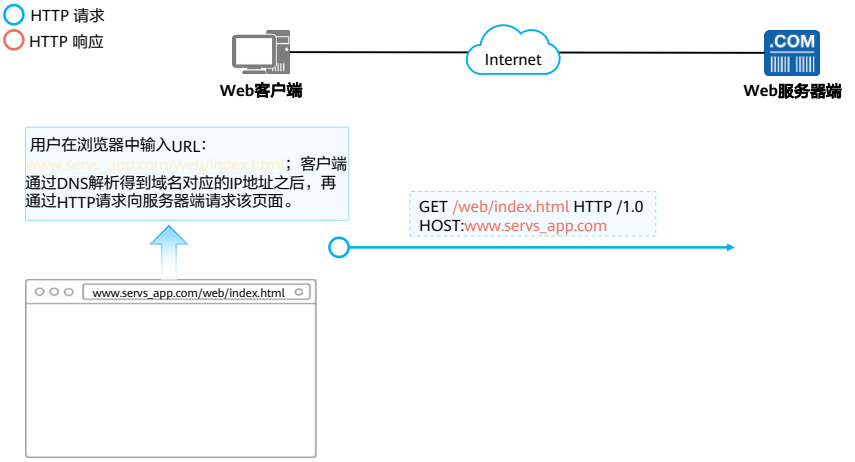


世界各地的Web服务器、客户端共同组成WWW网络

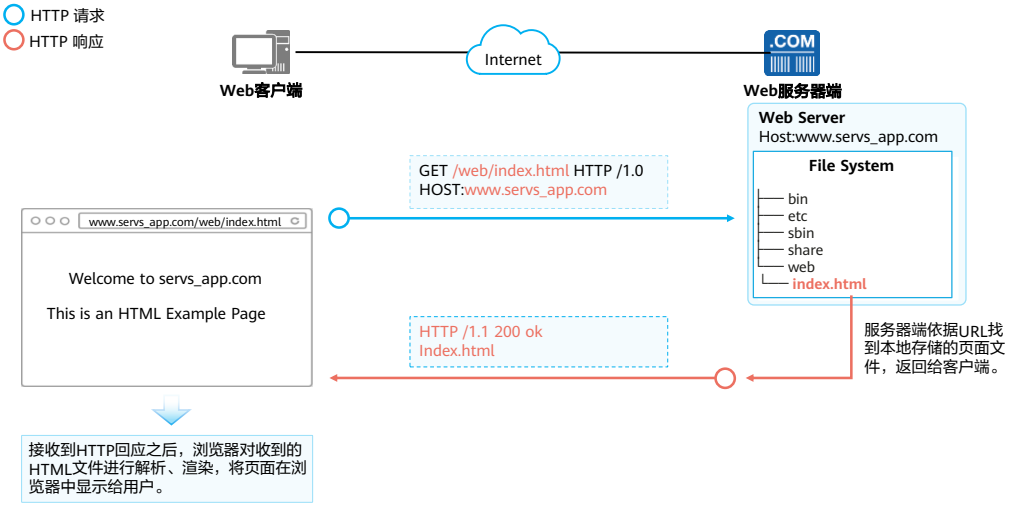
- 在互联网早期，为了进行文档之间的共享，人们提出了WWW（World Wide Web，万维网）。
- WWW由三部分组成：在浏览器中显示文档内容的页面标记语言HTML（Hypertext Markup Language，超文本标记语言）、在网络上传输文档的协议HTTP、在网络上表明文档位置的URL。
- WWW早期其实是浏览HTML的客户端应用程序的名称，现在则代表三项技术的合集，也可以简称为Web。

传输示例（1）

- HTTP 请求
- HTTP 响应



传输示例 (2)

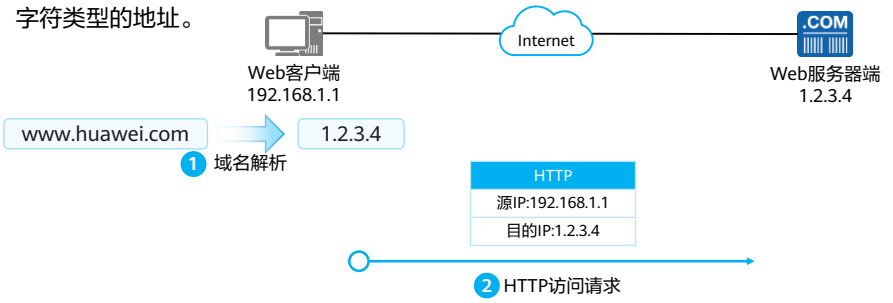


HUAWEI

目录

DNS的诞生

- 当我们在浏览器中输入一个域名访问某个网站时，这个域名最终会被解析为一个IP地址，我们的浏览器实际是在和这个IP地址进行通信。
- 负责将域名解析到IP地址的协议为DNS（Domain Name System，域名解析系统）。
- 网络中每个节点都有自己唯一的IP地址，通过IP地址可以实现节点之间的相互访问，但是如果和所有的节点进行通信都使用IP地址的方式，人们很难记住这么多IP地址，为此提出了DNS，将难以记忆的IP地址映射为字符类型的地址。

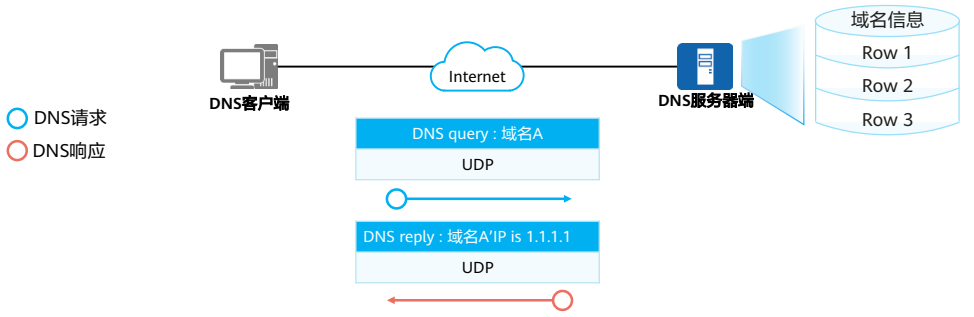


HUAWEI

- Internet的前身ARPAnet时就已经存在主机名称和IP地址的对应关系，只是当时主机数目很少，只需要一个HOSTS.txt文件就可以维护对应关系，HOSTS.txt由NIC（network information center）维护，改动自己主机名的使用者通过电子邮件将自身改动发送给NIC，NIC定期更新HOSTS.txt，这一切在主机数目很少时都没有什么问题。但当ARPAnet使用TCP/IP协议之后，网络用户数量出现了激增，手动维护HOSTS.txt似乎变得困难起来：
  - 名称冲突：NIC虽然可以保证管理的主机名称一致性，但很难保证主机不会随机修改名称和别人正在使用的一致。
  - 一致性：随着网络规模扩大，用户的HOSTS.txt很难保持一致性，很可能主机的HOSTS.txt文件还未更新，其余主机的名称已经变动了数次。
- 于是接替者DNS由此诞生。

域名系统组成

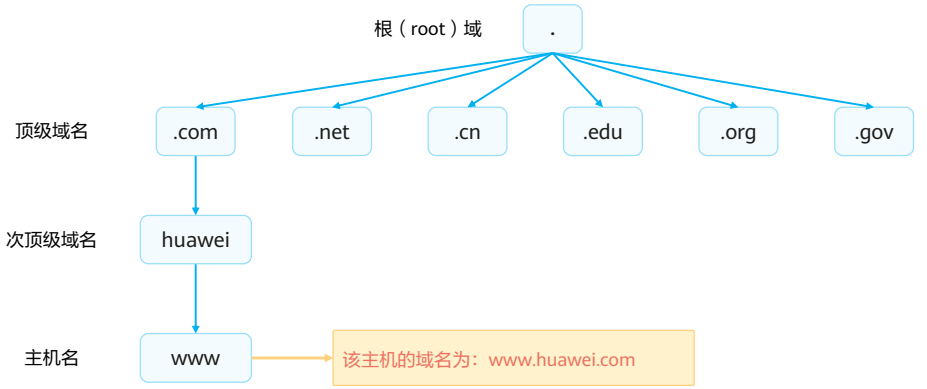
- 域名：主机的字符标识方式。大部分情况下，我们访问网站时在浏览器内输入的URL就是该网站的域名。
- 域名解析服务器（DNS Server）：负责维护域名与IP地址对应关系的数据库，并对解析者的请求进行响应。



- 域名系统是一个分布式的结构，每个服务器上的数据库只保存了部分域名与IP的对应关系。

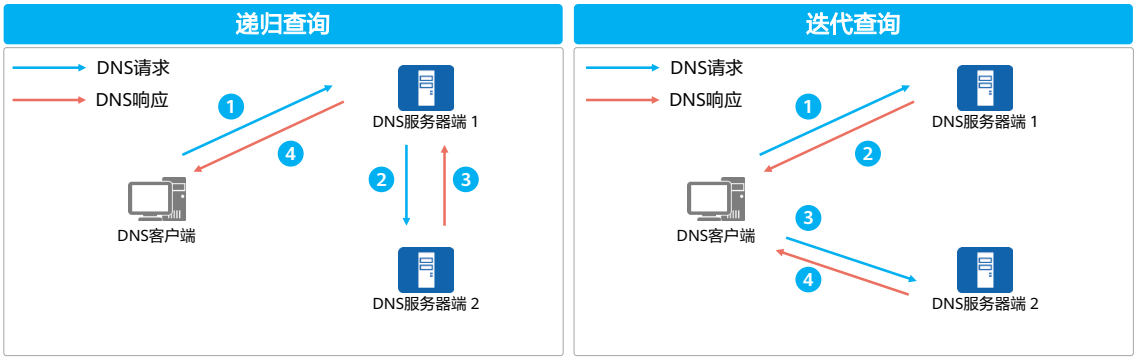
域名的表示方法

- 域名的表示方法为：主机名.次顶级域名.顶级域名.根域，根域为“.”，一般最后的根域不表示。



DNS查询方式

- DNS是一个分布式系统，绝大多数的DNS服务器端的数据库不会拥有所有的域名记录，当客户端向一个DNS服务器端查询域名但该DNS服务器端上却没有该域名的记录时，此时会有两种继续查询的方式：
  - 递归查询：由DNS服务器向其他DNS服务器进行查询，将最终查询结果返回给DNS客户端
  - 迭代查询：DNS服务器告知DNS客户端其他DNS服务器地址，客户端自行向其他DNS服务器进行查询。



HUAWEI

- 迭代查询不同于递归查询，DNS服务器1返回的DNS响应里的内容是另外一个DNS服务器地址。

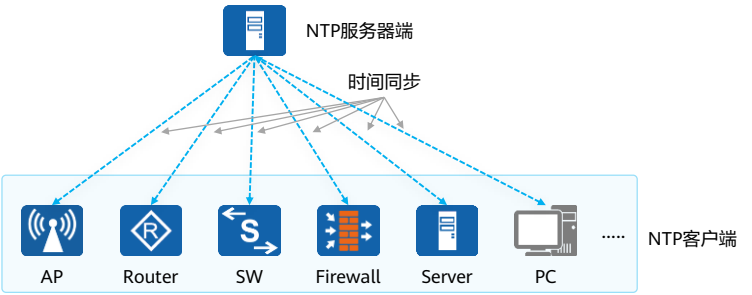
目录

时间同步需求

- 当今企业园区网络中很多场景都需要所有设备保持时钟一致：
  - 网络管理：对从不同路由器采集来的日志信息、调试信息进行分析时，需要以时间作为参照依据。
  - 计费系统：要求所有设备的时钟保持一致。
  - 多个系统协同处理同一个复杂事件：为保证正确的执行顺序，多个系统必须参考同一时钟。
  - 备份服务器和客户机之间进行增量备份：要求备份服务器和所有客户机之间的时钟同步。
  - 系统时间：某些应用程序需要知道用户登录系统的时间以及文件修改的时间。

NTP简介

- 如果采用管理员手工输入命令修改系统时间来进行时间同步，不但工作量巨大，而且也不能保证适中的精确性。为此可以使用NTP(Network Time Protocol)技术来同步设备的时钟。
- 网络时间协议NTP（Network Time Protocol）是TCP/IP协议族里面的一个应用层协议。NTP用于在一系列分布式时间服务器与客户端之间同步时钟。NTP的实现基于IP和UDP。NTP报文通过UDP传输，端口号是123。

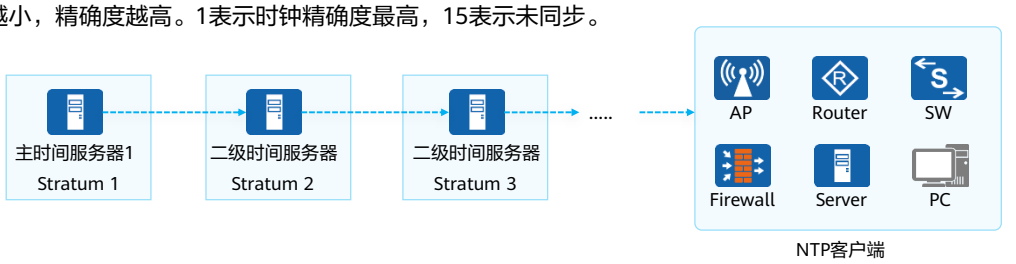


HUAWEI

- 目前主流的网络设备，如：AC（Access Controller，无线控制器）、AP（Access Point，接入点）、Firewall、Router、Switch、Server等基本都作为NTP的客户端，同时其中部分还可以作为NTP服务器端。

NTP网络结构

- 主时间服务器：通过线缆或无线电直接同步到标准参考时钟，标准参考时钟通常是Radio Clock或卫星定位系统等。
- 二级时间服务器：通过网络中的主时间服务器或者其他二级服务器取得同步。二级时间服务器通过NTP将时间信息传送到局域网内部的其它主机。
- 层数（stratum）：层数是对时钟同步情况的一个分级标准，代表了一个时钟的精确度，取值范围1～15，数值越小，精确度越高。1表示时钟精确度最高，15表示未同步。



思考题

1. 传输网络设备上的日志文件、配置文件，推荐使用FTP的哪种模式？为什么？
2. 为什么DHCP客户端收到Offer之后不直接使用该IP地址，还需要发送一个Request告知服务器端？
3. HTML、URL、HTTP的作用分别是什么？

1. ASCII模式；Binary模式更加适用于传输无法进行编码转换的非文本文件，如：EXE、BIN、cc（VRP版本文件）等。
2. 广播的Request报文让网络中其他DHCP服务器端得知客户端已选择了某个服务器端分配的IP地址，保证其他服务器端可以释放通过单播Offer分配给该客户端的IP地址。
3. HTML用于显示页面内容，URL用于定位文档文件的网络位置，HTTP用于请求、传输文档。

本章总结

- FTP用于文件传输，传输不同的文件推荐使用不同的模式；由于其基于TCP，因此使用FTP传输文件可以保障传输的可靠性、高效性。
- 为解决手动分配IP地址产生的问题，使用DHCP进行动态IP地址分配可以有效减少管理员的工作量，避免用户手动配置网络参数造成的IP地址冲突。
- 作为WWW的文档传输协议，HTTP在当今的网络中有着广泛的应用。


谢 谢
www.huawei.com

WLAN概述

前言

- 以有线电缆或光纤作为传输介质的有线局域网应用广泛，但有线传输介质的铺设成本高，位置固定，移动性差。随着人们对网络的便携性和移动性的要求日益增强，传统的有线网络已经无法满足需求，WLAN（Wireless Local Area Network，无线局域网)技术应运而生。
- 目前，WLAN已经成为一种经济、高效的网络接入方式。
- 本课程介绍了WLAN在不同阶段的发展历程，其次介绍了WLAN技术相关的概念以及常见组网架构的工作原理，最后介绍WLAN常见组网架构的基本配置和WLAN技术的未来发展趋势。

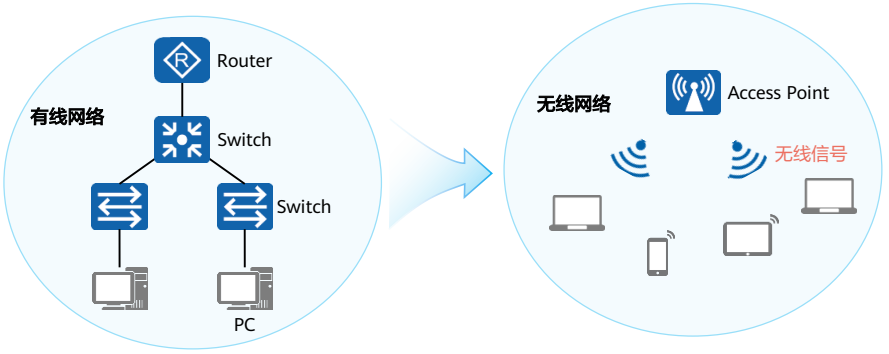
目标

- 学完本课程后，您将能够：
  - 了解WLAN的基本概念和802.11协议族的发展历史
  - 区分WLAN的不同设备
  - 区分WLAN的不同组网方式
  - 掌握WLAN工作流程
  - 完成WLAN的基本配置

目录

什么是WLAN

- WLAN即Wireless LAN（无线局域网），是指通过无线技术构建的无线局域网络。WLAN广义上是指以无线电波、激光、红外线等无线信号来代替有线局域网中的部分或全部传输介质所构成的网络。
- 通过WLAN技术，用户可以方便地接入到无线网络，并在无线网络覆盖区域内自由移动，彻底摆脱有线网络的束缚。

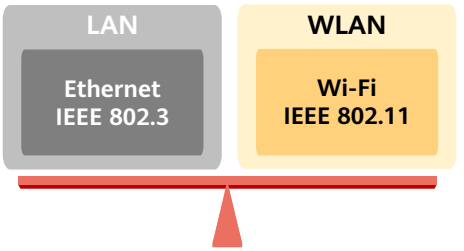


HUAWEI

- WLAN即Wireless LAN（无线局域网），是指通过无线技术构建的无线局域网络。
  - 这里指的无线技术不仅仅包含Wi-Fi，还有红外、蓝牙、ZigBee等等。
  - 通过WLAN技术，用户可以方便地接入到无线网络，并在无线网络覆盖区域内自由移动，摆脱有线网络的束缚。
- 无线网络根据应用范围可分为WPAN、WLAN、WMAN、WWAN。
  - WPAN (Wireless Personal Area Network)：个人无线网络，常用技术有：Bluetooth、ZigBee、NFC、HomeRF、UWB。
  - WLAN (Wireless Local Area Network)：无线局域网，常用技术有：Wi-Fi。(WLAN中也会使用WPAN的相关技术。)
  - WMAN (Wireless Metropolitan Area Network)：无线城域网，常用技术有：WiMax。
  - WWAN (Wireless Wide Area Network)：无线广域网，常用技术有：GSM、CDMA、WCDMA、TD-SCDMA、LTE、5G。
- WLAN的优点：
  - 网络使用自由：凡是自由空间均可连接网络，不受限于线缆和端口位置。在办公大楼、机场候机厅、度假村、商务酒店、体育场馆、咖啡店等场所尤为适用。
  - 网络部署灵活：对于地铁、公路交通监控等难于布线的场所，采用WLAN进行无线网络覆盖，免去或减少了繁杂的网络布线，实施简单，成本低，扩展性好。
- 本课程介绍的WLAN特指通过Wi-Fi技术基于802.11标准系列，利用高频信号（例如2.4GHz或5GHz）作为传输介质的无线局域网。

IEEE 802.11、WLAN与Wi-Fi

- IEEE 802.11是现今无线局域网通用的标准。它是由国际电机电子工程学会(IEEE)定义的无线网络通信的标准。
- Wi-Fi联盟制造商的商标，并做为产品的品牌认证，是一种创建于IEEE 802.11标准上的无线局域网技术。在大多数场景下，Wi-Fi可等同于802.11。

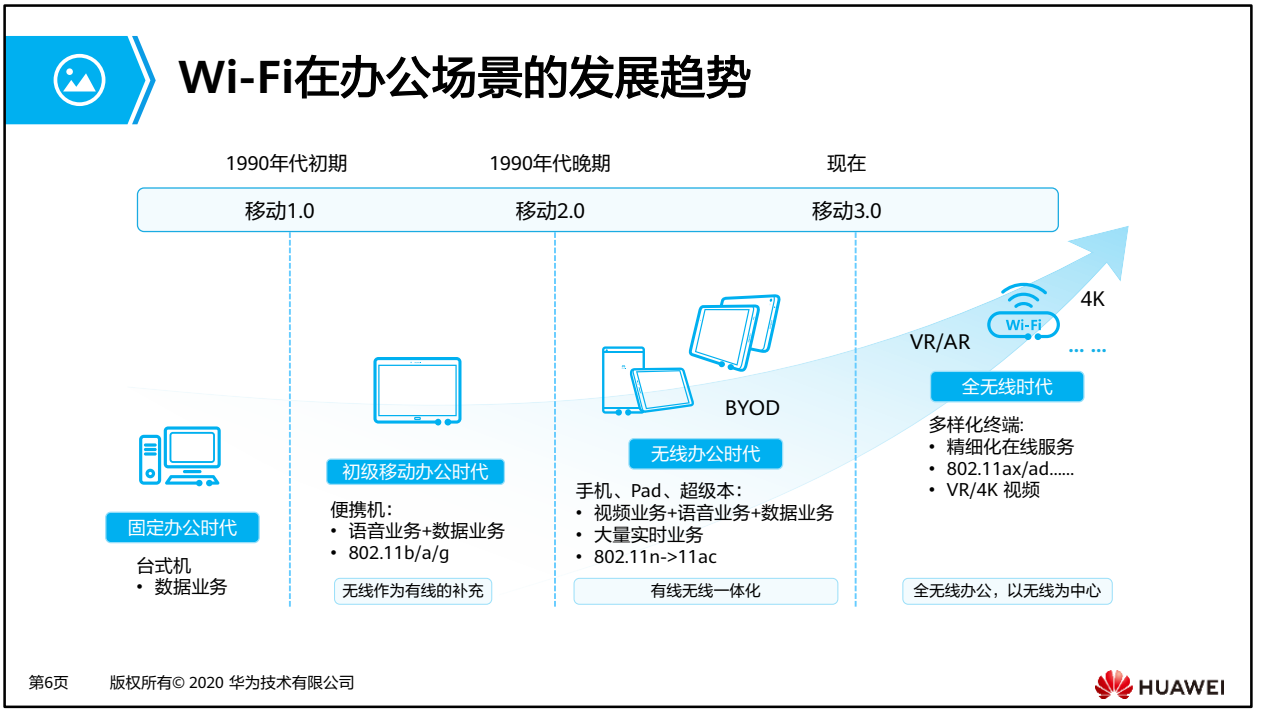


- **IEEE 802.11标准与Wi-Fi的世代**

频率	2.4GHz	2.4GHz	2.4GHz、5GHz	2.4GHz & 5GHz	5GHz	5GHz	2.4GHz & 5GHz
速率	2Mbit/s	11Mbit/s	54Mbit/s	300Mbit/s	1300Mbit/s	6.9Gbit/s	9.6Gbit/s
协议	802.11	802.11b	802.11a、802.11g	802.11n	802.11ac wave1	802.11ac wave2	802.11ax
Wi-Fi	Wi-Fi 1	Wi-Fi 2	Wi-Fi 3	Wi-Fi 4	Wi-Fi 5		Wi-Fi 6
时间	1997	1999	2003	2009	2013	2015	2018

HUAWEI

- IEEE 802.11标准聚焦在TCP/IP对等模型的下两层：
  - 数据链路层：主要负责信道接入、寻址、数据帧校验、错误检测、安全机制等内容。
  - 物理层：主要负责在空口（空中接口）中传输比特流，例如规定所使用的频段等。
- Wi-Fi联盟成立于1999年，当时的名称叫做Wireless Ethernet Compatibility Alliance（WECA）。在2002年10月，正式改名为Wi-Fi Alliance。
- IEEE 802.11第一个版本发表于1997年。此后，更多的基于IEEE 802.11的补充标准逐渐被定义，最为熟知的是影响Wi-Fi代际演进的标准：802.11b、802.11a、802.11g、802.11n、802.11ac等。
- 在IEEE 802.11ax标准推出之际，Wi-Fi联盟将新Wi-Fi规格的名字简化为Wi-Fi 6，主流的IEEE 802.11ac改称Wi-Fi 5、IEEE 802.11n改称Wi-Fi 4，其他世代以此类推。

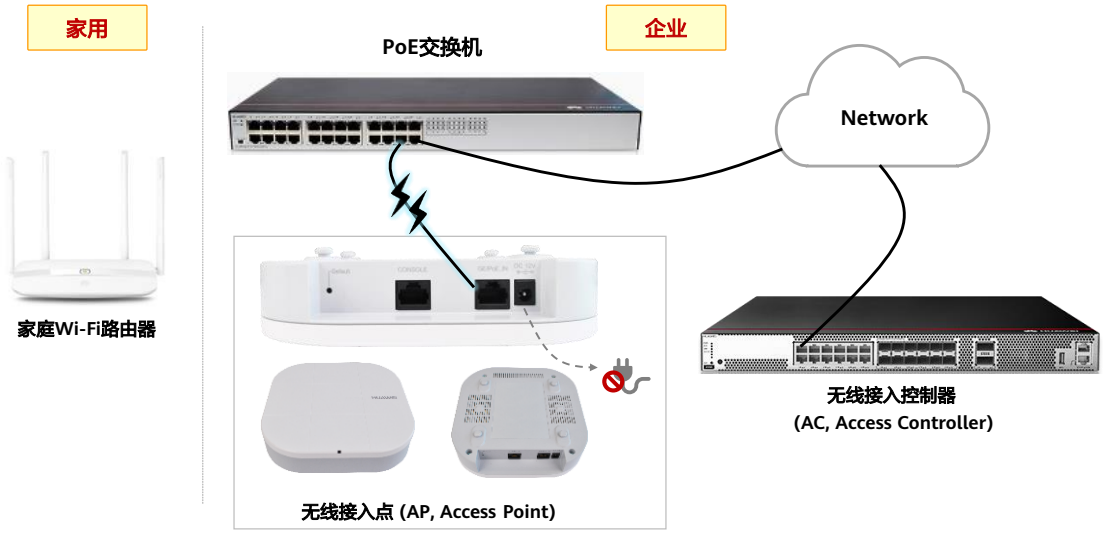


- 第一阶段：初级移动办公时代，无线作为有线的补充
  - WaveLAN技术的应用可以被认为是最早的企业WLAN雏形。早期的Wi-Fi技术主要应用在类似“无线收音机”这样的物联设备上，但是随着802.11a/b/g标准的推出，无线连接的优势越来越明显。企业和消费者开始认识到Wi-Fi技术的应用潜力，无线热点开始出现在咖啡店、机场和酒店。
  - Wi-Fi也在这一时期诞生，它是Wi-Fi联盟的商标，该联盟最初的目的是为了推动802.11b标准的制定，并在全球范围内推行Wi-Fi产品的兼容认证。随着标准的演进和遵从标准产品的普及，人们往往将Wi-Fi等同于802.11标准。
  - 802.11标准是众多WLAN技术中的一种，只是802.11标准已成为业界的主流技术，所以人们提到WLAN时，通常是指使用Wi-Fi技术的WLAN。
  - 这是WLAN应用的第一阶段，主要是解决“无线接入”的问题，核心价值是摆脱有线的束缚，设备在一定范围内可以自用移动，用无线网络延伸了有线网络。但是这一阶段的WLAN对安全、容量和漫游等方面没有明确的诉求，接入点（Access Point，AP）的形态还是单个接入点，用于单点组网覆盖。通常称单个接入点架构的AP为FAT AP。

- 第二阶段：无线办公时代，有线无线一体化
    - 随着无线设备的进一步普及，WLAN从起初仅仅作为有线网络的补充，发展到和有线网络一样不可或缺，由此进入第二阶段。
    - 在这个阶段，WLAN作为网络的一部分，还需要为企业访客提供网络接入。
    - 在办公场景下，存在大量视频、语音等大带宽业务，对WLAN的带宽有更大的需求。从2012年开始，802.11ac标准趋于成熟，对工作频段、信道带宽、调制与编码方式等做出了诸多改进，与以往的Wi-Fi标准相比，其具有更高的吞吐率、更少的干扰，能够允许更多的用户接入。
- 第三阶段：全无线办公时代，以无线为中心
    - 目前，WLAN已经进入第三阶段，在办公环境中，使用无线网络彻底替代有线网络。办公区采用全Wi-Fi覆盖，办公位不再提供有线网口，办公环境更为开放和智能。
    - 未来，云桌面办公、智真会议、4K视频等大带宽业务将从有线网络迁移至无线网络，而VR/AR等新技术将直接基于无线网络部署。新的应用场景对WLAN的设计与规划提出更高的要求。
    - 2018年，新一代Wi-Fi标准Wi-Fi 6 (IEEE命名为802.11ax，Wi-Fi 6是Wi-Fi联盟的命名)发布，这是Wi-Fi发展史上的又一重大里程碑，Wi-Fi 6的核心价值是容量的进一步提升，引领无线通信进入10Gbit/s时代；多用户并发性能提升4倍，让网络在高密度接入、业务重载的情况下，依然保持优秀的服务能力。

目录

WLAN设备介绍

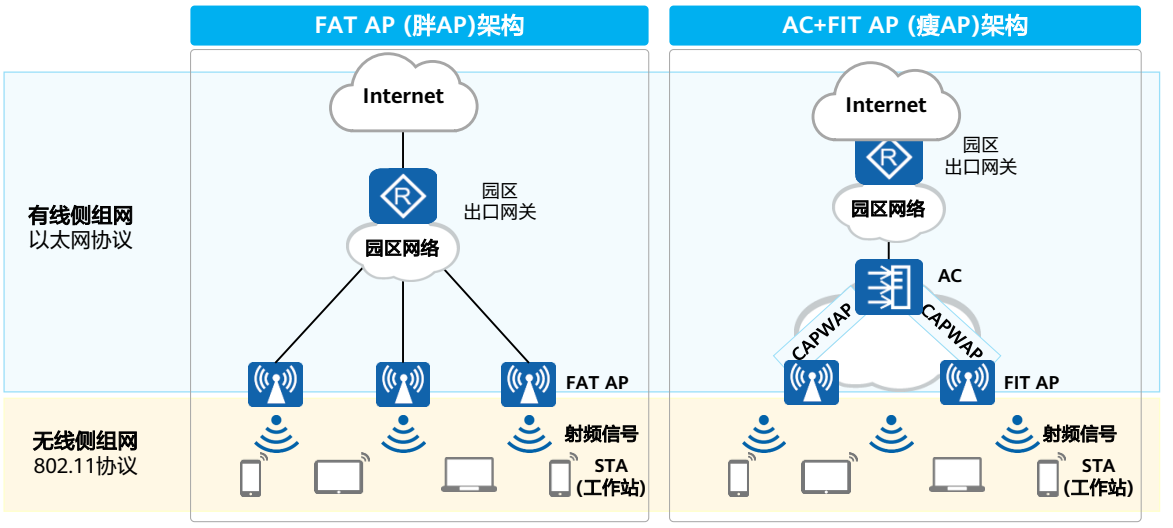


- 华为无线局域网产品形态丰富，覆盖室内室外、家庭企业等各种应用场景，提供高速、安全和可靠的无线网络连接。
- 家庭WLAN产品：
  - 家庭Wi-Fi路由器：通过把有线网络信号转换成无线信号，供家庭电脑、手机等设备接收，实现无线上网功能。
- 企业WLAN产品：
  - 无线接入点 (AP, Access Point)
    - 一般支持FAT AP（胖AP）、FIT AP（瘦AP）和云管理AP三种工作模式，根据网络规划的需求，可以灵活地在多种模式下切换。
    - FAT AP：适用于家庭，独立工作，需单独配置，功能较为单一，成本低。独立完成用户接入、认证、数据安全、业务转发和QoS等功能。
    - FIT AP：适用于大中型企业，需要配合AC使用，由AC统一管理和配置，功能丰富，对网络维护人员的技能要求高。用户接入、AP上线、认证、路由、AP管理、安全协议、QoS等功能需要同AC配合完成。
    - 云管理：适用于中小型企业，需要配合云管理平台使用，由云管理平台统一管理和配置，功能丰富，即插即用，对网络维护人员的技能要求低。
  - 无线接入控制器 (AC, Access Controller)
    - 一般位于整个网络的汇聚层，提供高速、安全、可靠的WLAN业务。
    - 提供大容量、高性能、高可靠性、易安装、易维护的无线数据控制业务，具有

组网灵活、绿色节能等优势。

- PoE交换机
  - PoE（Power over Ethernet，以太网供电）是指通过以太网网络进行供电，也被称为基于局域网的供电系统PoL（Power over LAN）或有源以太网（Active Ethernet）。
  - PoE允许电功率通过传输数据的线路或空闲线路传输到终端设备。
  - 在WLAN网络中，可以通过PoE交换机对AP设备进行供电。

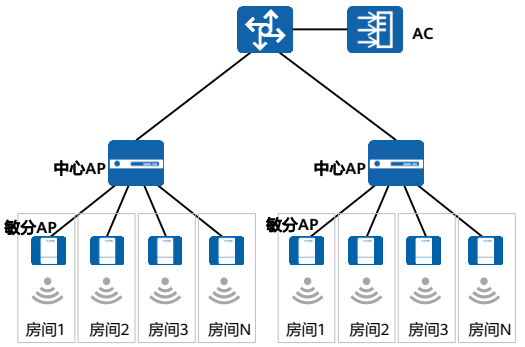
基本的WLAN组网架构



- WLAN网络架构分有线侧和无线侧两部分，有线侧是指AP上行到Internet的网络使用以太网协议，无线侧是指STA到AP之间的网络使用802.11协议。
- 无线侧接入的WLAN网络架构为集中式架构。从最初的FAT AP架构，演进为AC+FIT AP架构。
  - FAT AP (胖AP)架构
    - 这种架构不需要专门的设备集中控制就可以完成无线用户的接入、业务数据的加密和业务数据报文的转发等功能，因此又称为自治式网络架构。
    - 适用范围：家庭
    - 特点：AP独立工作，需要单独配置，功能较为单一，成本低。
    - 缺点：随着WLAN覆盖面积增大，接入用户增多，需要部署的FAT AP数量也会增多，但FAT AP是独立工作的，缺少统一的控制设备，因此管理维护这些FAT AP就十分麻烦。
  - AC+FIT AP (瘦AP)架构
    - 这种架构中，AC负责WLAN的接入控制、转发和统计、AP的配置监控、漫游管理、AP的网管代理、安全控制；FIT AP负责802.11报文的加解密、802.11的物理层功能、接受AC的管理等简单功能。
    - 适用范围：大中型企业
    - 特点：AP需要配合AC使用，由AC统一管理和配置，功能丰富，对网络运维人员的技能要求高。
- 注：在本课程中，我们主要以AC+FIT AP架构为例进行课程的讲解。

- WLAN基本概念：
  - 工作站STA (Station)：
    - 支持802.11标准的终端设备。例如带无线网卡的电脑、支持WLAN的手机等。
  - 无线接入控制器AC (Access Controller)：
    - 在AC+FIT AP网络架构中，AC对无线局域网中的所有FIT AP进行控制和管理。例如，AC可以通过与认证服务器交互信息来为WLAN用户提供认证服务。
  - 无线接入点AP (Access Point)：
    - 为STA提供基于802.11标准的无线接入服务，起到有线网络和无线网络的桥接作用。
  - 无线接入点控制与规范CAPWAP (Control And Provisioning of Wireless Access Points)：
    - 由RFC5415协议定义的，实现AP和AC之间的互通的一个通用封装和传输机制。
  - 射频信号 (无线电磁波)：
    - 提供基于802.11标准的WLAN技术的传输介质，是具有远距离传输能力的高频电磁波。本课程指的射频信号是2.4G或5G频段的无线电磁波。

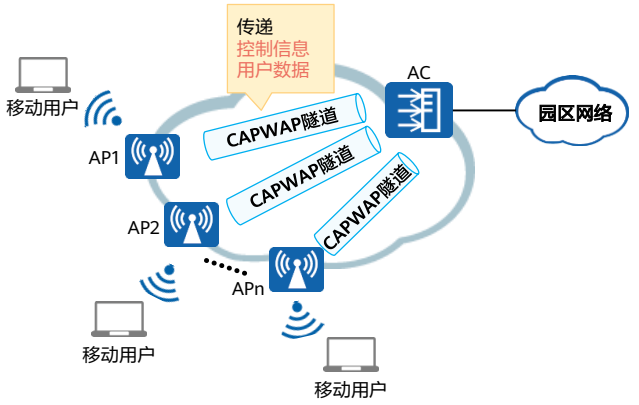
敏捷分布式AP架构



架构特点

- AP的一种特殊架构，将AP拆分为中心AP和敏分AP两部分，中心AP可管理多台敏分AP，在适用的场景下，成本低，覆盖好。敏捷分布式AP可以用于FAT AP、AC+FIT AP、云管理架构。
- 适用范围：房间分布密集的场景。

有线侧组网概念：CAPWAP协议



什么是CAPWAP隧道

- **CAPWAP（Control And Provisioning of Wireless Access Points Protocol，无线接入点控制和配置协议）**：该协议定义了如何对AP进行管理、业务配置，即AC通过CAPWAP隧道来实现对AP的集中管理和控制。

CAPWAP隧道的功能

- AP与AC间的状态维护。
- AC通过CAPWAP隧道对AP进行管理、业务配置下发。
- 当采用隧道转发模式时，AP将STA发出的数据通过CAPWAP隧道实现与AC之间的交互。

HUAWEI

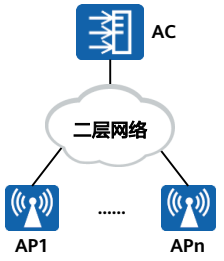
- 为满足大规模组网的要求，需要对网络中的多个AP进行统一管理，IETF成立了CAPWAP工作组，最终制定CAPWAP协议。该协议定义了AC如何对AP进行管理、业务配置，即AC与AP间首先会建立CAPWAP隧道，然后AC通过CAPWAP隧道来实现对AP的集中管理和控制。
- CAPWAP是基于UDP进行传输的应用层协议。
  - CAPWAP协议在传输层运输两种类型的消息：
    - 业务数据流量，封装转发无线数据帧 。——通过CAPWAP数据隧道。
    - 管理流量，管理AP和AC之间交换的管理消息 。——通过CAPWAP控制隧道。
  - CAPWAP数据和控制报文基于不同的UDP端口发送：
    - 管理流量端口为UDP端口5246。
    - 业务数据流量端口为UDP端口5247。

- 注：国际互联网工程任务组（The Internet Engineering Task Force，简称 IETF）。

有线侧组网概念：AP-AC组网方式

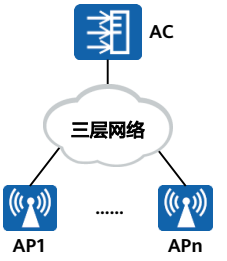
- AP和AC间的组网分为：二层组网和三层组网。

二层组网



- 二层组网：AP与AC之间的网络为直连或者二层网络。
- 由于二层组网比较简单，适用于简单临时的组网，能够进行比较快速的组网配置，但不适用于大型组网架构。

三层组网



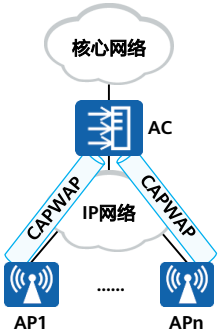
- 三层组网：AP与AC之间的网络为三层网络。
- 在实际组网中，一台AC可以连接几十甚至几百台AP，组网一般比较复杂，在大型组网中一般采用三层组网。

- AP-AC组网：二层是指AP和AC之间是二层组网，三层是指AC和AP之间是三层组网；二层组网AP可以通过二层广播，或者DHCP过程，即插即用上线；三层网络下，AP无法直接发现AC，需要通过DHCP或DNS方式动态发现，或者配置静态IP。
- 在实际组网中，一台AC可以连接几十甚至几百台AP，组网一般比较复杂。比如在企业网络中，AP可以布放在办公室，会议室，会客间等场所，而AC可以安放在公司机房。这样，AP和AC之间的网络就是比较复杂的三层网络。因此，在大型组网中一般采用三层组网。

有线侧组网概念：AC连接方式

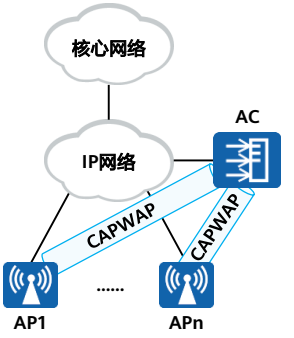
- AC的连接方式分为：直连式组网和旁挂式组网。

直连式组网



- 直连式组网可以认为AP、AC与上层网络串联在一起，所有数据必须通过AC到达上层网络。
- 直连式组网中AC同时扮演AC和汇聚交换机的功能，AP的数据业务和管理业务都由AC集中转发和处理。

旁挂式组网

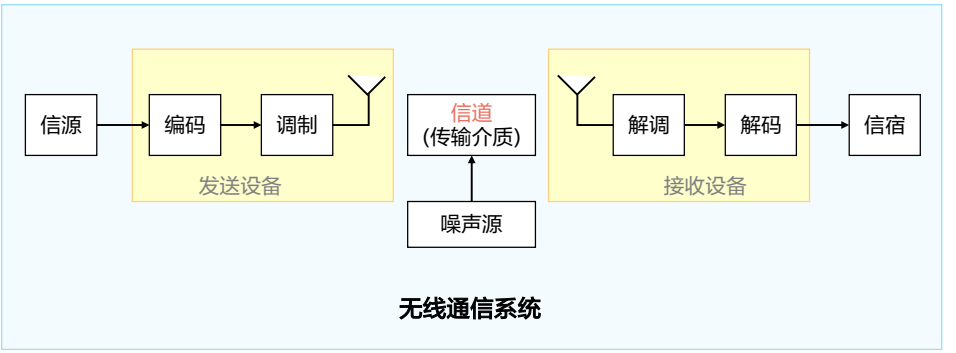


- 旁挂式组网，AC旁挂在AP与上行网络的直连网络中，不再直接连接AP。
- 旁挂式组网，AC旁挂在AP与上行网络的直连网络上，AP的业务数据可以不经AC而直接到达上行网络 。

- AC连接方式：直连模式下AC部署在用户的转发路径上，旁挂则相反；直连模式用户流量要经过AC，会消耗AC转发能力，旁挂一般流量不会经过AC。
- 直连式组网：
  - 采用这种组网方式，对AC的吞吐量以及处理数据能力要求比较高，否则AC会是整个无线网络带宽的瓶颈。
  - 但用此种组网，组网架构清晰，组网实施起来简单。
- 旁挂式组网：
  - 由于实际组网中，大部分不是早期就规划好无线网络，无线网络的覆盖架设大部分是后期在现有网络中扩展而来。而采用旁挂式组网就比较容易进行扩展，只需将AC旁挂在现有网络中，比如旁挂在汇聚交换机上，就可以对终端AP进行管理。所以此种组网方式使用率比较高。
  - 在旁挂式组网中，AC只承载对AP的管理功能，管理流封装在CAPWAP隧道中传输。数据业务流可以通过CAPWAP数据隧道经AC转发，也可以不经过AC转发直接转发，后者无线用户业务流经汇聚交换机由汇聚交换机传输至上层网络。

无线侧组网概念：无线通信系统

- 无线通信系统中，信息可以是图像、文字、声音等。信息需要先经过信源编码转换为方便于电路计算和处理的数字信号，再经过信道编码和调制，转换为无线电波发射出去。

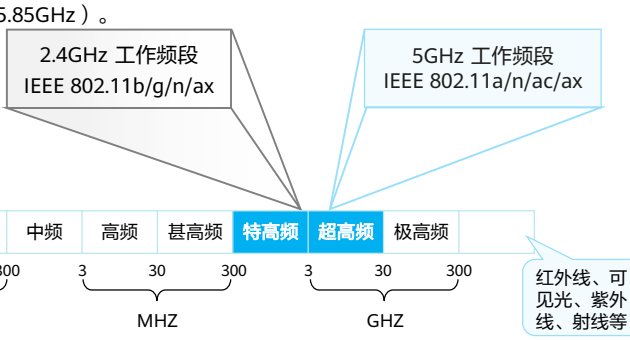


无线通信系统

- 编码：
  - 信源编码：将最原始的信息，经过对应的编码，转换为数字信号的过程。
  - 信道编码：是一种对信息纠错、检错的技术，可以提升信道传输的可靠性。信息在无线传输过程中容易受到噪声的干扰，导致接收信息出错，引入信道编码能够在接收设备上最大程度地恢复信息，降低误码率。
- 调制：将数字信号叠加到高频振荡电路产生的高频信号上，才能通过天线转换成无线电波发射出去。叠加动作就是调制的过程。
- 信道：传输信息的通道，无线信道就是空间中的无线电波。
- 空中接口：简称空口，无线信道使用的接口。发送设备和接收设备使用接口和信道连接，对于无线通信，接口是不可见的，连接着不可见的空间。

无线侧组网概念：无线电磁波

- 无线电磁波是频率介于3赫兹和约300G赫兹之间的电磁波，也叫作射频电波，或简称射频、射电。无线电技术将声音讯号或其他信号经过转换，利用无线电磁波传播。
- WLAN技术就是通过无线电磁波在空间中传输信息。当前我们使用的频段是：
  - 2.4GHz频段 (2.4GHz~2.4835GHz)；
  - 5GHz频段（5.15GHz~5.35GHz，5.725GHz~5.85GHz ）。



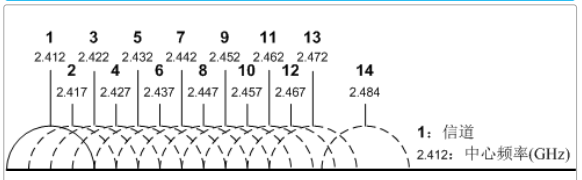
- **无线电磁波频谱：**

- 极低频 (3Hz–30Hz)：潜艇通讯或直接转换成声音。
- 超低频 (30Hz–300Hz)： 直接转换成声音或交流输电系统（ 50-60赫兹)。
- 特低频 (300Hz–3KHz)： 矿场通讯或直接转换成声音。
- 甚低频 (3KHz–30KHz)： 直接转换成声音、超声、地球物理学研究。
- 低频 (30KHz–300KHz)： 国际广播。
- 中频 (300KHz–3MHz)： 调幅(AM)广播、海事及航空通讯。
- 高频 (3MHz–30MHz)： 短波、民用电台。
- 甚高频 (30MHz–300MHz)： 调频(FM)广播、电视广播、航空通讯。
- 特高频 (300MHz–3GHz)： 电视广播、无线电话通讯、无线网络、微波炉。
- 超高频 (3GHz–30GHz)： 无线网络、雷达、人造卫星接收。
- 极高频 (30GHz–300GHz)： 射电天文学、遥感、人体扫描安检仪。
- 300GHz以上：红外线、可见光、紫外线、射线等。

无线侧组网概念：无线信道

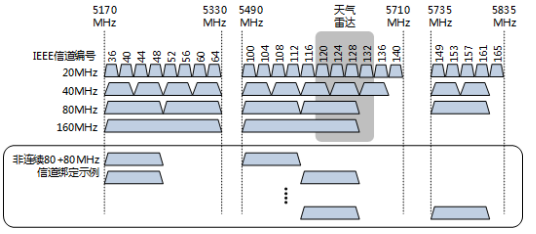
- 信道是传输信息的通道，无线信道就是空间中的无线电磁波。无线电磁波无处不在，如果随意使用频谱资源，那将带来无穷无尽的干扰问题，所以无线通信协议除了要定义出允许使用的频段，还要精确划分出频率范围，每个频率范围就是信道。

2.4GHz频段



- 2.4GHz频段被划分为14个有重叠的、频率宽度是20MHz的信道。其中包含重叠信道和非重叠信道。
  - 重叠信道：在一个空间内同时存在重叠信道，则会产生干扰问题。
  - 非重叠信道：在一个空间内可以同时存在非重叠信道，不会产生干扰问题。

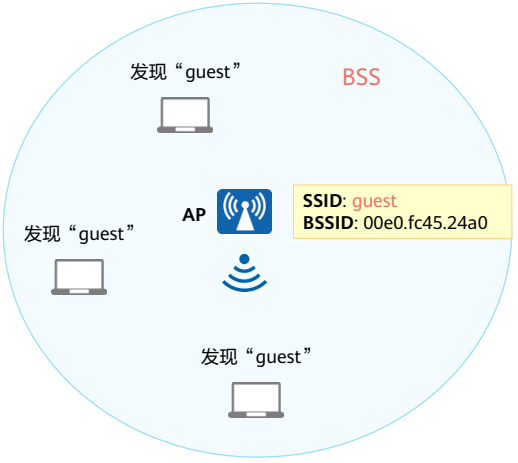
5GHz频段



- 对于5G频段，支持更宽频谱，频率资源更为丰富，AP不仅支持20MHz带宽的信道，还支持40MHz、80MHz及更大带宽的信道。

- WLAN中，AP的工作状态会受到周围环境的影响。例如，当相邻AP的工作信道存在重叠频段时，某个AP的功率过大会对相邻AP造成信号干扰。
- 通过射频调优功能，动态调整AP的信道和功率，可以使同一AC管理的各AP的信道和功率保持相对平衡，保证AP工作在最佳状态。

无线侧组网概念：BSS/SSID/BSSID

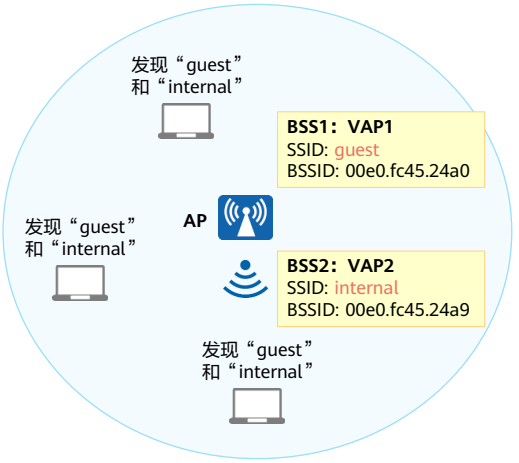


- 基本服务集BSS (Basic Service Set)：
  - 一个AP所覆盖的范围。
  - 在一个BSS的服务区域内，STA可以相互通信。
- 基本服务集标识符BSSID (Basic Service Set Identifier)：
  - 是无线网络的一个身份标识，用AP的MAC地址表示。
- 服务集标识符SSID (Service Set Identifier)：
  - 是无线网络的一个身份标识，用字符串表示。
  - 为了便于用户辨识不同的无线网络，用SSID代替BSSID。

HUAWEI

- BSS (Basic Service Set)：
  - 无线网络的基本服务单元，通常由一个AP和若干STA组成，BSS是802.11网络的基本结构。由于无线介质共享性，BSS中报文收发需携带BSSID（MAC地址）。
- 基本服务集标识符BSSID（Basic Service Set Identifier）：
  - AP上的数据链路层MAC地址。
  - 终端要发现和找到AP，需要通过AP的一个身份标识，这个身份标识就是BSSID。
  - 为了区分BSS，要求每个BSS都有唯一的BSSID，因此使用AP的MAC地址来保证其唯一性。
- 服务集标识符SSID（Service Set Identifier）：
  - 表示无线网络的标识，用来区分不同的无线网络。例如，当我们在笔记本电脑上搜索可接入无线网络时，显示出来的网络名称就是SSID。
  - 如果一个空间部署了多个BSS，终端就会发现多个BSSID，只要选择加入的BSSID就行。但是做选择的是用户，为了使得AP的身份更容易辨识，则用一个字符串来作为AP的名字。这个字符串就是SSID。

无线侧组网概念：VAP

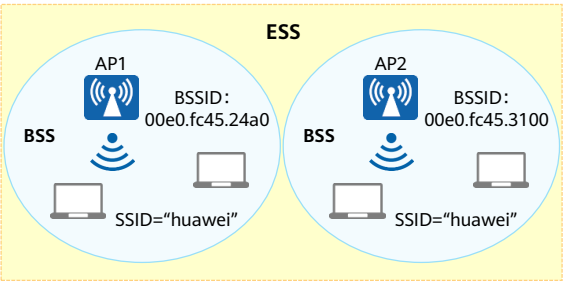


- 早期的AP只支持1个BSS，如果要在同一空间内部署多个BSS，则需要安放多个AP，这不但增加了成本，还占用了信道资源。为了改善这种状况，现在的AP通常支持创建出多个虚拟AP (Virtual Access Point, VAP)。
- 虚拟接入点VAP：
  - VAP就是在一个物理实体AP上虚拟出的多个AP。每一个被虚拟出的AP就是一个VAP。每个VAP提供和物理实体AP一样的功能。
  - 每个VAP对应1个BSS。这样1个AP，就可以提供多个BSS，可以再为这些BSS，设置不同的SSID。

  HUAWEI

- 虚拟接入点VAP（Virtual Access Point）：
  - 是AP设备上虚拟出来的业务功能实体。用户可以在一个AP上创建不同的VAP来为不同的用户群体提供无线接入服务。
- VAP简化了WLAN的部署，但不意味VAP越多越好，要根据实际需求进行规划。一味增加VAP的数量，不仅要让用户花费更多的时间找到SSID，还会增加AP配置的复杂度。而且VAP并不等同于真正的AP，所有的VAP都共享这个AP的软件和硬件资源，所有VAP的用户都共享相同的信道资源，所以AP的容量是不变的，并不会随着VAP数目的增加而成倍的增加。

无线侧组网概念：ESS



- 为了满足实际业务的需求，需要对BSS的覆盖范围进行扩展。同时用户从一个BSS移动到另一个BSS时，不能感知到SSID的变化，则可以通过扩展服务集ESS实现。
- 扩展服务集ESS (Extend Service Set)：
  - 由多个使用相同SSID的BSS组成，是采用相同的SSID的多个BSS组成的更大规模的虚拟BSS。

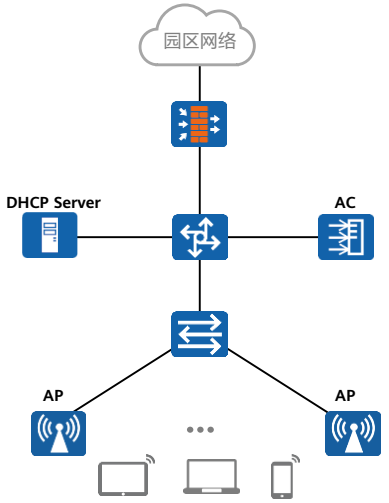
  HUAWEI

- ESS (Extend Service Set)：
  - 采用相同的SSID的多个BSS组成的更大规模的虚拟BSS。
  - 用户可以带着终端在ESS内自由移动和漫游，不管用户移动到哪里，都可以认为使用的同一个WLAN。
- WLAN漫游：
  - 指STA在同属一个ESS的不同AP的覆盖范围之间移动且保持用户业务不中断的行为。
  - WLAN网络的最大优势就是STA不受物理介质的影响，可以在WLAN覆盖范围内四处移动并且能够保持业务不中断。同一个ESS内包含多个AP设备，当STA从一个AP覆盖区域移动到另外一个AP覆盖区域时，利用WLAN漫游技术可以实现STA用户业务的平滑切换。

目录

AP上线 配置下发 STA接入 数据转发

WLAN工作流程概述

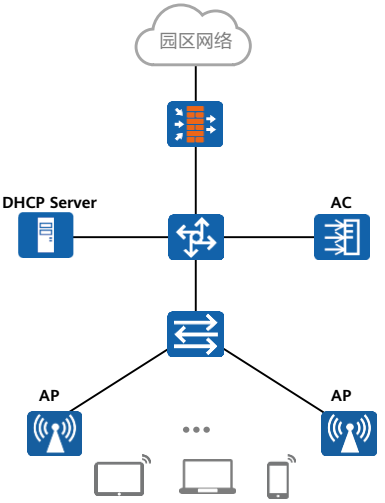


WLAN工作流程

1. AP上线
   AP获取IP地址并发现AC，与AC建立连接
2. WLAN业务配置下发
   AC将WLAN业务配置下发到AP生效
3. STA接入
   STA搜索到AP发射的SSID并连接、上线，接入网络
4. WLAN业务数据转发
   WLAN网络开始转发业务数据

- AC+FIT AP组网架构中，是通过AC对AP进行统一的管理，因此所有的配置都是在AC上进行的。

WLAN工作流程：步骤1



WLAN工作流程

① AP上线

FIT AP需完成上线过程，AC才能实现对AP的集中管理和控制，以及业务下发。AP的上线过程包括如下步骤：

1. AP获取IP地址；
2. AP发现AC并与之建立CAPWAP隧道；
3. AP接入控制；
4. AP版本升级；
5. CAPWAP隧道维持。

② WLAN业务配置下发

③ STA接入

④ WLAN业务数据转发

AP获取IP地址

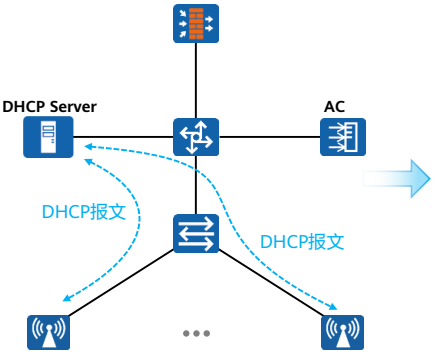
- AP必须获得IP地址才能够与AC通信，WLAN网络才能够正常工作。

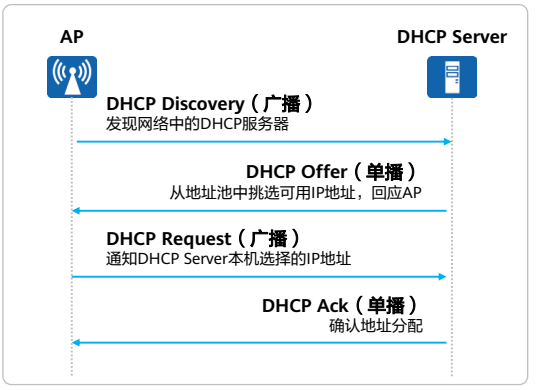
AP获取IP地址

- AP获取IP地址的方式包括以下：
  - 静态方式：登录到AP设备上手工配置IP地址。
  - DHCP方式：通过配置DHCP服务器，使AP作为DHCP客户端向DHCP服务器请求IP地址。
- 典型方案：
  - 部署专门的DHCP Server为AP分配IP地址。
  - 使用AC的DHCP服务为AP分配IP地址。
  - 使用网络中的设备，例如核心交换机为AP分配IP地址。

AP获取IP地址：DHCP方式

- **AP获取IP地址**
- CAPWAP 隧道建立
- AP接入控制
- AP的版本升级（可选）
- CAPWAP 隧道维持





HUAWEI

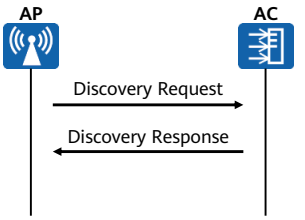
CAPWAP隧道建立

- AC通过CAPWAP隧道来实现对AP的集中管理和控制。

AP获取IP地址
CAPWAP隧道建立
AP接入控制
AP的版本升级（可选）
CAPWAP隧道维持

Step 1：Discovery阶段（AP发现AC阶段）

- AP通过发送Discovery Request报文，找到可用的AC。
- AP发现AC有两种方式：
  - 静态方式：AP上预先配置AC的静态IP地址列表。
  - 动态方式：DHCP方式、DNS方式和广播方式。

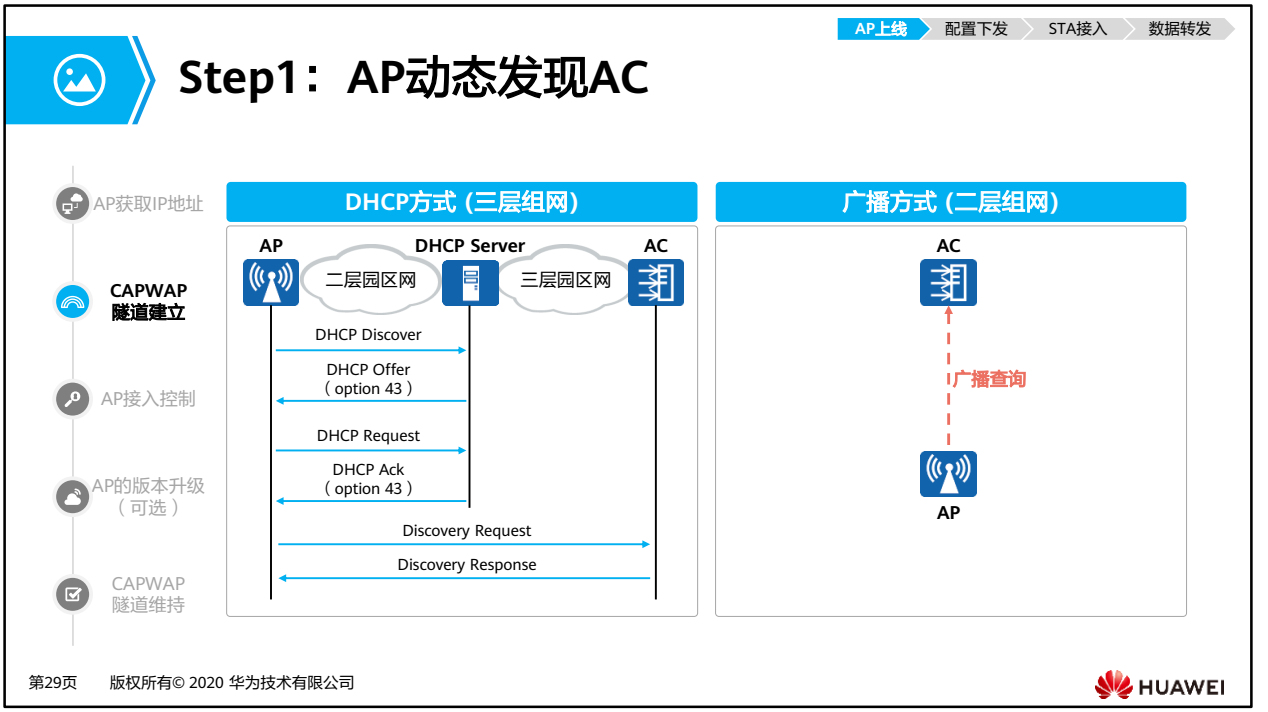


Step 2：建立CAPWAP隧道阶段

- AP与AC关联，完成CAPWAP隧道建立。包括数据隧道和控制隧道：
  - 数据隧道：AP接收的业务数据报文经过CAPWAP数据隧道集中到AC上转发。
  - 控制隧道：通过CAPWAP控制隧道实现AP与AC之间的管理报文的交互。

HUAWEI

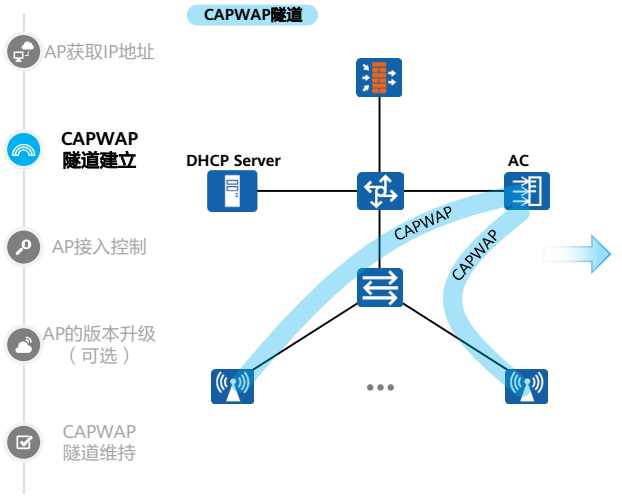
- CAPWAP隧道可以实现：
  - AP与AC间的状态维护；
  - AC对AP进行管理和业务配置下发；
  - 业务数据经过CAPWAP隧道集中到AC上转发。
- AP发现AC阶段：
  - 静态方式：AP上预先配置了AC的静态IP地址列表。AP上线时，AP分别发送Discovery Request单播报文到所有预配置列表对应IP地址的AC。然后AP通过接收到AC返回的Discovery Response报文，选择一个AC开始建立CAPWAP隧道。
  - 动态方式：分为DHCP方式、DNS方式和广播方式，其中本章主要介绍DHCP方式和广播方式。



- DHCP方式：
  - 通过DHCP的四步交互过程，获取AC的IP地址：
    - 在没有预配置AC的IP列表时，则启动AP动态AC发现机制。通过DHCP获取IP地址，并通过DHCP协议中的Option返回AC地址列表（在DHCP服务器上配置DHCP响应报文中携带Option 43，且Option 43携带AC的IP地址列表）。
    - 首先是AP发送DHCP Discover广播报文，请求DHCP Server响应，在DHCP服务器侦听到DHCP Discover报文后，它会从没有租约的地址范围中，选择最前面的闲置IP，连同其他TCP/IP设定，响应AP一个DHCP Offer报文，该报文中会包含一个租约期限的信息。
    - 由于DHCP Offer报文既可以是单播报文，也可以是广播报文，当AP端收到多台DHCP Server的响应时，只会挑选其中一个Offer(通常是最先抵达的那个)，然后向网络中发送一个DHCP Request广播报文，告诉所有的Offer，并重新发送DHCP，将指定接收哪一台服务器提供的IP地址。
    - 当DHCP Server接收到AP的Request报文之后，会向AP发送一个DHCP Ack响应，该报文中携带的信息包括了AP的IP地址，租约期限，网关信息，以及DNS Server IP等，以此确定租约的正式生效，就此完成DHCP的四步交互工作。

    - 通过AC发现机制，与AC关联：
        - AP通过DHCP服务获取AC的IP地址后，使用AC发现机制来获知哪些AC是可用的，决定与最佳AC来建立CAPWAP的连接。
        - AP启动CAPWAP协议的发现机制，以单播或广播的形式发送发现请求报文试图关联AC，AC收到AP的Discovery Request以后，会发送一个单播Discovery Response 给AP，AP可以通过Discover Response中所带的AC优先级或者AC上当前AP的个数等，确定与哪个AC建立会话。
- 广播方式：
    - 当AP启动后，如果DHCP方式和DNS方式均未获得AC的IP或AP发出发现请求报文后未收到响应，则AP启动广播发现流程，以广播包方式发出发现请求报文。
    - 接收到发现请求报文的AC检查该AP是否有接入本机的权限（已经授权的MAC地址或者序列号），如果有则发回响应。如果该AP没有接入权限，AC则拒绝请求。
    - 广播发现方式只适用于AC/AP间为二层可达的网络场景。

Step2：建立CAPWAP隧道



Step 2：建立CAPWAP隧道阶段

- AP与AC关联，完成CAPWAP隧道建立。包括数据隧道和控制隧道：
  - 数据隧道：AP接收的业务数据报文经过CAPWAP数据隧道集中到AC上转发。同时还可以选择对数据隧道进行数据传输层安全DTLS（Datagram Transport Layer Security）加密，使能DTLS加密功能后，CAPWAP数据报文都会经过DTLS加解密。
  - 控制隧道：通过CAPWAP控制隧道实现AP与AC之间的管理报文的交互。同时还可以选择对控制隧道进行数据传输层安全DTLS加密，使能DTLS加密功能后，CAPWAP控制报文都会经过DTLS加解密。

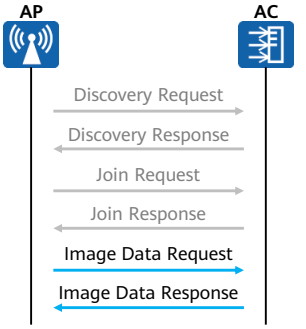
HUAWEI

- 在收到AP发送的Join Request报文之后，AC会进行AP合法性的认证，认证通过则添加相应的AP设备。
- AC上支持三种对AP的认证方式：
  - MAC认证
  - 序列号（SN）认证
  - 不认证
- AC上添加AP的方式有三种：
  - 离线导入AP：预先配置AP的MAC地址和SN，当AP与AC连接时，如果AC发现AP和预先增加的AP的MAC地址和SN匹配，则AC开始与AP建立连接。
  - 自动发现AP：当配置AP的认证模式为不认证或配置AP的认证模式为MAC或SN认证且将AP加入AP白名单中，则当AP与AC连接时，AP将被AC自动发现并正常上线。
  - 手工确认未认证列表中的AP：当配置AP的认证模式为MAC或SN认证，但AP没有离线导入且不在已设置的AP白名单中，则该AP会被记录到未授权的AP列表中。需要用户手工确认后，此AP才能正常上线。

AP的版本升级

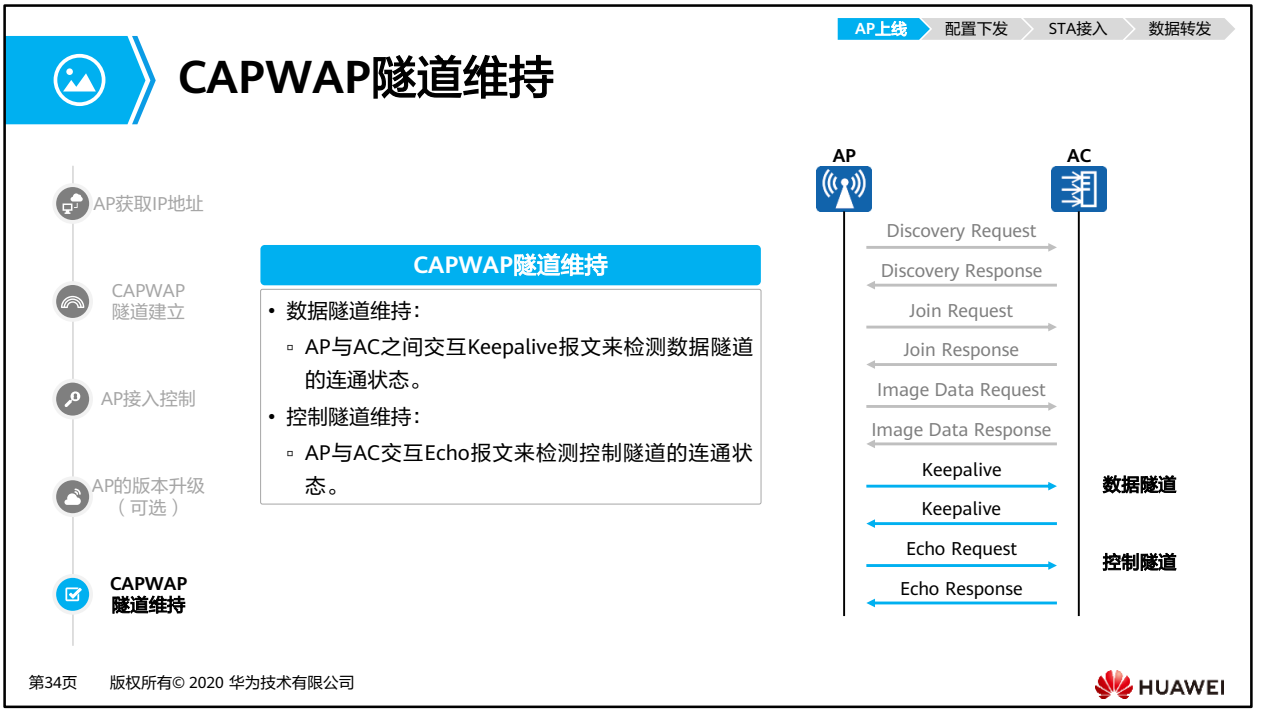
AP的版本升级

- AP根据收到的Join Response报文中的参数判断当前的系统软件版本是否与AC上指定的一致。如果不一致，则AP通过发送Image Data Request报文请求软件版本，然后进行版本升级，升级方式包括AC模式、FTP模式和SFTP模式。
- AP在软件版本更新完成后重新启动，重复进行前面三个步骤。

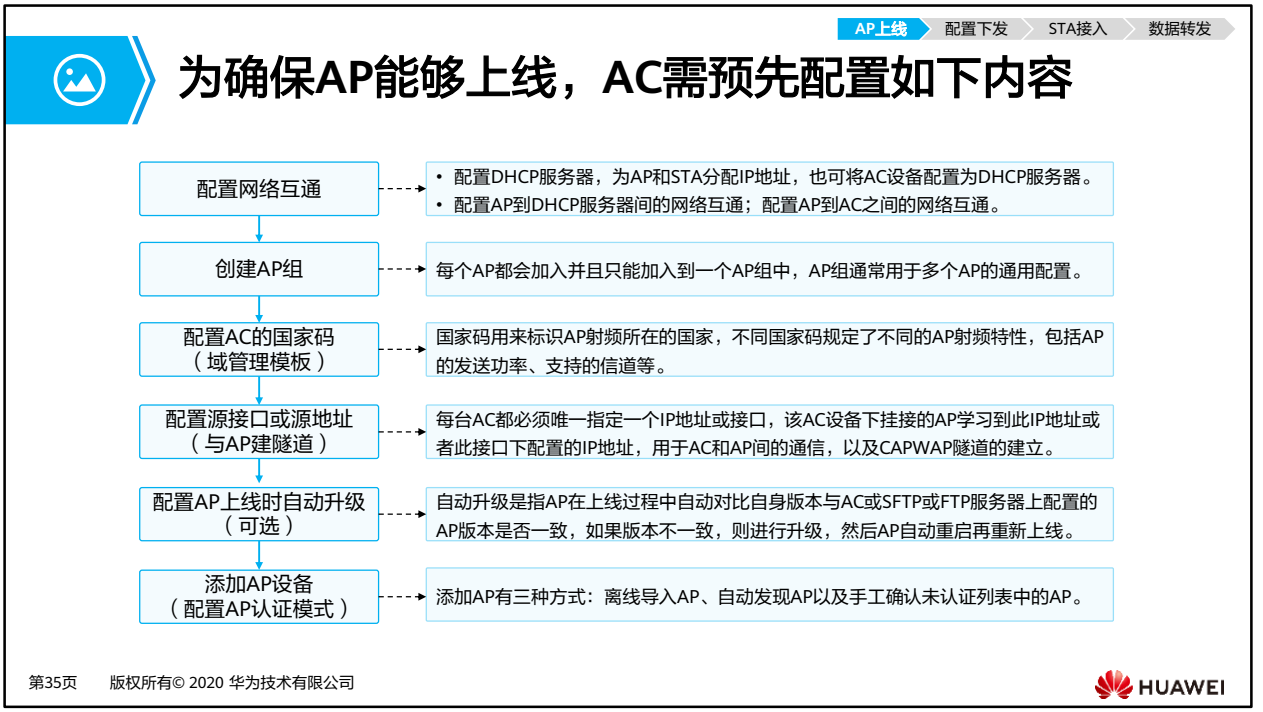


  HUAWEI

- 在AC上给AP升级方式：
  - 自动升级：主要用于AP还未在AC中上线的场景。通常先配置好AP上线时的自动升级参数，然后再配置AP接入。AP在之后的上线过程中会自动完成升级。如果AP已经上线，配置完自动升级参数后，任意方式触发AP重启，AP也会进行自动升级。但相比于自动升级，使用在线升级方式升级能够减少业务中断的时间。
    - AC模式：AP升级时从AC上下载升级版本，适用于AP数量较少时的场景。
    - FTP模式：AP升级时从FTP服务器上下载升级版本，适用于网络安全性要求不是很高的文件传输场景中，采用明文传输数据，存在安全隐患。
    - SFTP模式：AP升级时从SFTP服务器上下载升级版本，适用于网络安全性要求高的场景，对传输数据进行了严格加密和完整性保护。
  - 在线升级：主要用于AP已经在AC中上线并已承载了WLAN业务的场景。
  - 定时升级：主要用于AP已经在AC中上线并已承载了WLAN业务的场景。通常指定在网络访问量少的时间段升级。

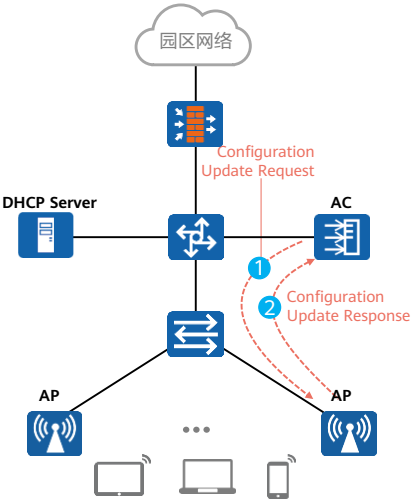


- 数据隧道维持：
  - AP与AC之间交互Keepalive (UDP端口号为5247)报文来检测数据隧道的连通状态。
- 控制隧道维持：
  - AP与AC交互Echo (UDP端口号为5246)报文来检测控制隧道的连通状态。



- 域管理模板：
  - 域管理模板提供对AP的国家码、调优信道集合和调优带宽等的配置。
  - 国家码用来标识AP射频所在的国家，不同国家码规定了不同的AP射频特性，包括AP的发送功率、支持的信道等。配置国家码是为了使AP的射频特性符合不同国家或区域的法律法规要求。
- 配置AC的源接口或源地址：
  - 每台AC都必须唯一指定一个IP地址、VLANIF接口或者Loopback接口，该AC设备下挂接的AP学习到此IP地址或者此接口下配置的IP地址，用于AC和AP间的通信。此IP地址或者接口称为源地址或源接口。
  - 只有为每台AC指定唯一一个源接口或源地址，AP才能与AC建立CAPWAP隧道。
  - 设备支持使用VLANIF接口或Loopback接口作为源接口，支持使用VLANIF接口或Loopback接口下的IP地址作为源地址。
- 添加AP设备：即配置AP认证模式，AP上线。
  - 添加AP有三种方式：离线导入AP、自动发现AP以及手工确认未认证列表中的AP。

WLAN工作流程：步骤2



WLAN工作流程

① AP上线

② WLAN业务配置下发

AC向AP发送Configuration Update Request请求消息，AP回应Configuration Update Response消息，AC再将AP的业务配置信息下发给AP。

③ STA接入

④ WLAN业务数据转发

- AP上线后，会主动向AC发送Configuration Status Request报文，该信息中包含了现有AP的配置，为了做AP的现有配置和AC设定配置的匹配检查。当AP的当前配置与AC要求不符合时，AC会通过Configuration Status Response通知AP。
- 说明：AP上线后，首先会主动向AC获取当前配置，而后统一由AC对AP进行集中管理和业务配置下发。

配置模板

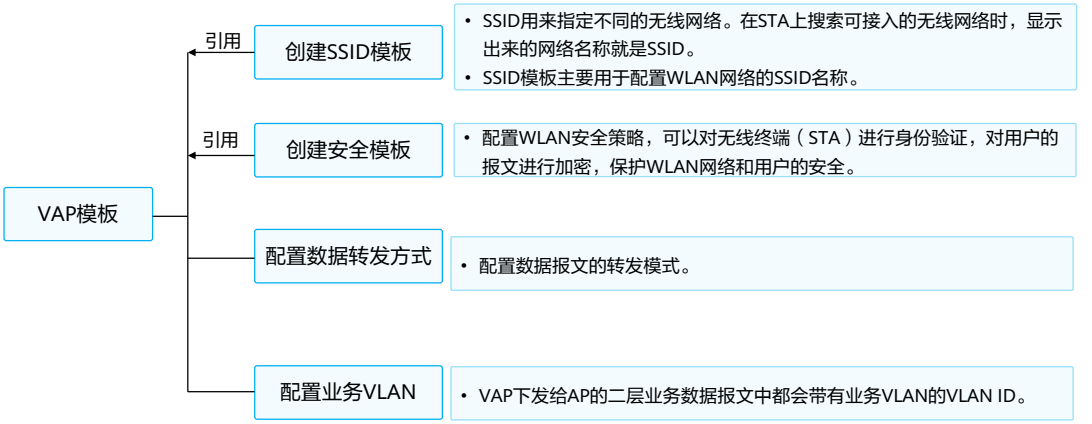
- 为了方便用户配置和维护WLAN的各个功能，针对WLAN的不同功能和特性设计了各种类型的模板，这些模板统称为WLAN模板。

  HUAWEI

- WLAN网络中存在着大量的AP，为了简化AP的配置操作步骤，可以将AP加入到AP组中，在AP组中统一对AP进行同样的配置。但是每个AP也有着不同于其它AP的参数配置，不便于通过AP组来进行统一配置，这类个性化的参数可以直接在每个AP下配置。每个AP在上线时都会加入并且只能加入到一个AP组中。当AP从AC上获取到AP组和AP个性化的配置后，会优先使用AP下的配置。
- AP组和AP下都能够引用如下模板：域管理模板、AP系统模板、、射频模板、VAP模板，部分模板例还能继续引用其它模板。
  - 域管理模板：
    - 国家码用来标识AP射频所在的国家，不同国家码规定了不同的AP射频特性，包括AP的发送功率、支持的信道等。配置国家码是为了使AP的射频特性符合不同国家或区域的法律法规要求。
    - 通过配置调优信道集合，可以在配置射频调优功能时指定AP信道动态调整的范围，同时避开雷达信道和终端不支持信道。
  - 射频模板：
    - 根据实际的网络环境对射频的各项参数进行调整和优化，使AP具备满足实际需求的射频能力，提高WLAN网络的信号质量。射频模板中各项参数下发到AP后，只有AP支持的参数才会在AP上生效。
    - 可配置的参数包括：射频的类型、射频的速率、射频的无线报文组播发送速率、AP发送Beacon帧的周期等。

- VAP模板：
  - 在VAP模板下配置各项参数，然后在AP组或AP中引用VAP模板，AP上就会生成VAP，VAP用来为STA提供无线接入服务。通过配置VAP模板下的参数，使AP实现为STA提供不同无线业务服务的能力。
  - VAP模板下还能继续引用SSID模板、安全模板、流量模板等。
- 其他模板：如WIDS模板、AP有线口模板、WDS模板、定位模板和Mesh模板等。
- 射频参数配置：
  - AP射频需要根据实际的WLAN网络环境来配置不同的基本射频参数，以使AP射频的性能达到更优。
  - WLAN网络中，相邻AP的工作信道存在重叠频段时，容易产生信号干扰，对AP的工作状态产生影响。为避免信号干扰，使AP工作在更佳状态，提高WLAN网络质量，可以手动配置相邻AP工作在非重叠信道上。
  - 根据实际网络环境的需求，配置射频的发射功率和天线增益，使射频信号强度满足实际网络需求，提高WLAN网络的信号质量。
  - 实际应用场景中，两个AP之间的距离可能为几十米到几十公里，因为AP间的距离不同，所以AP之间传输数据时等待ACK报文的时间也不相同。通过调整合适的超时时间参数，可以提高AP间的数据传输效率。

VAP模板

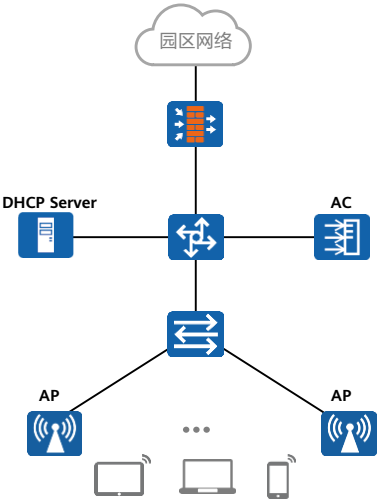


  HUAWEI

- SSID模板：主要用于配置WLAN网络的SSID名称，还可以配置其他功能，主要包括如下功能：
  - 隐藏SSID功能：用户在创建无线网络时，为了保护无线网络的安全，可以对无线网络名称进行隐藏设置。这样，只有知道网络名称的无线用户才能连接到这个无线网络中。
  - 单个VAP下能够关联成功的最大用户数：单个VAP下接入的用户数越多，每个用户能够使用的平均网络资源就越少，为了保证用户的上网体验，可以根据实际的网络状况配置合理的最大用户接入数。
  - 用户数达到最大时自动隐藏SSID的功能：使能用户数达到最大时自动隐藏SSID的功能后，当WLAN网络下接入的用户数达到最大时，SSID会被隐藏，新用户将无法搜索到SSID。
- 安全模板：配置WLAN安全策略，可以对无线终端进行身份验证，对用户的报文进行加密，保护WLAN网络和用户的安全。
  - WLAN安全策略支持开放认证、WEP、WPA/WPA2-PSK、WPA/WPA2-802.1X等，在安全模板中选择其中一种进行配置。

- 数据转发方式：
  - 控制报文是通过CAPWAP的控制隧道转发的，用户的数据报文分为隧道转发（又称为“集中转发”）方式、直接转发（又称为“本地转发”）方式。这部分内容在后面的课程中会详细介绍。
- 业务VLAN：
  - 由于WLAN无线网络灵活的接入方式，STA可能会在某个地点（例如办公区入口或体育场馆入口）集中接入到同一个WLAN无线网络中，然后漫游到其它AP覆盖的无线网络环境下。
    - 业务VLAN配置为单个VLAN时，在接入STA数众多的区域容易出现IP地址资源不足、而其它区域IP地址资源浪费的情况。
    - 业务VLAN配置为VLAN pool时，可以在VLAN pool中加入多个VLAN，然后通过将VLAN pool配置为VAP的业务VLAN，实现一个SSID能够同时支持多个业务VLAN。新接入的STA会被动态的分配到VLAN pool中的各个VLAN中，减少了单个VLAN下的STA数目，缩小了广播域；同时每个VLAN尽量均匀的分配IP地址，减少了IP地址的浪费。

WLAN工作流程：步骤3



WLAN工作流程

① AP上线

② WLAN业务配置下发

❸ STA接入

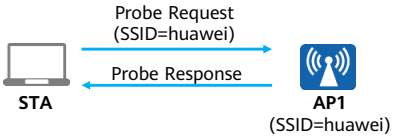
CAPWAP隧道建立完成后，用户就可以接入无线网络。
STA接入过程分为六个阶段：扫描阶段、链路认证阶段、关联阶段、接入认证阶段、DHCP、用户认证。

④ WLAN业务数据转发

扫描

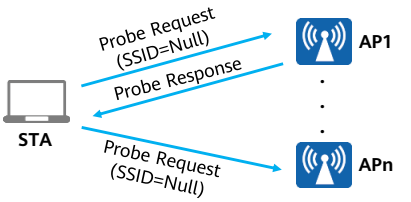
- STA可以通过主动扫描，定期搜索周围的无线网络，获取到周围的无线网络信息。
- 根据Probe Request帧（探测请求帧）是否携带SSID，可以将主动扫描分为两种：

携带有指定SSID的主动扫描方式



- 客户端发送携带有指定SSID的Probe Request：STA依次在每个信道发出Probe Request帧，寻找与STA有相同SSID的AP，只有能够提供指定SSID无线服务的AP接收到该探测请求后才回复探查响应。

携带空SSID的主动扫描方式



- 客户端发送广播Probe Request，客户端会定期地在其支持的信道列表中，发送Probe Request帧扫描无线网络。当AP收到Probe Request帧后，会回应Probe Response帧通告可以提供的无线网络信息。

- 主动扫描：
  - 携带有指定SSID的主动扫描方式：适用于STA通过主动扫描接入指定的无线网络。
  - 携带空SSID的主动扫描方式：适用于STA通过主动扫描可以获知是否存在可使用的无线服务。
- 被动扫描：
  - STA也支持被动扫描搜索无线网络。
  - 被动扫描是指客户端通过侦听AP定期发送的Beacon帧（信标帧，包含：SSID、支持速率等信息）发现周围的无线网络，缺省状态下AP发送Beacon帧的周期为100TUs（1TU=1024us）。

AP上线 配置下发 STA接入 数据转发

无线接入安全协议

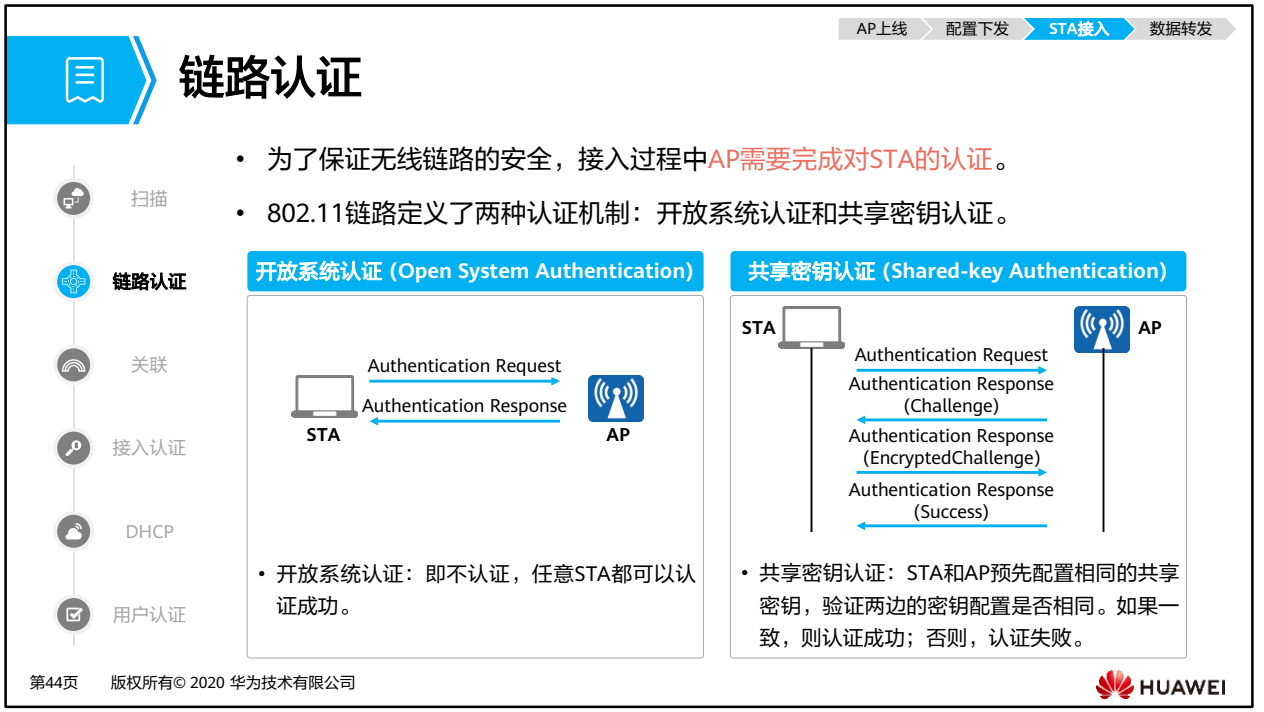
扫描
链路认证
关联
接入认证
DHCP
用户认证

- WLAN技术是以无线射频信号作为业务数据的传输介质，这种开放的信道使攻击者很容易对无线信道中传输的业务数据进行窃听和篡改。因此，安全性成为阻碍WLAN技术发展的最重要因素。
- 常用安全策略：

安全策略	链路认证方式	接入认证方式	数据加密方式	说明
WEP	Open	不涉及	不加密或WEP加密	不安全的安全策略
	Shared-key Authentication	不涉及	WEP加密	不安全的安全策略
WPA/ WPA2-802.1X	Open	802.1X（EAP）	TKIP或CCMP	安全性高的安全策略，适用于大型企业
WPA/ WPA2-PSK	Open	PSK	TKIP或CCMP	安全性高的安全策略，适用于中小型企业或家庭用户

HUAWEI

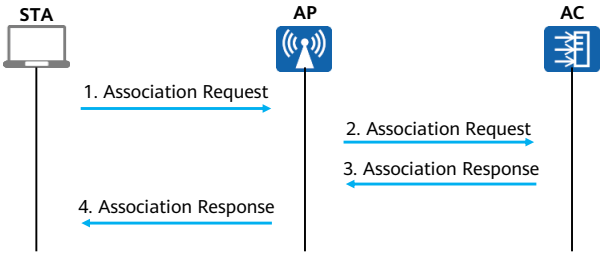
- WLAN安全提供了WEP、WPA、WPA2等安全策略机制。每种安全策略体现了一整套安全机制，包括无线链路建立时的链路认证方式，无线用户上线时的用户接入认证方式和无线用户传输数据业务时的数据加密方式。
- WEP（Wired Equivalent Privacy）
  - 有线等效加密WEP协议是由802.11标准定义的，用来保护无线局域网中的授权用户所传输的数据的安全性，防止这些数据被窃听。WEP的核心是采用RC4算法，加密密钥长度有64位、128位和152位，其中有24bit的IV（初始向量）是由系统产生的，所以WLAN服务端和WLAN客户端上配置的密钥长度是40位、104位或128位。WEP加密采用静态的密钥，接入同一SSID下的所有STA使用相同的密钥访问无线网络。
- WPA/WPA2 (Wi-Fi Protected Access)
  - 由于WEP共享密钥认证采用的是基于RC4对称流的加密算法，需要预先配置相同的静态密钥，无论从加密机制还是从加密算法本身，都很容易受到安全威胁。为了解决这个问题，在802.11i标准没有正式推出安全性更高的安全策略之前，Wi-Fi联盟推出了针对WEP改良的WPA。WPA的核心加密算法还是采用RC4，在WEP基础上提出了临时密钥完整性协议TKIP（Temporal Key Integrity Protocol）加密算法，采用了802.1X的身份验证框架，支持EAP-PEAP、EAP-TLS等认证方式。随后802.11i安全标准组织又推出WPA2，区别于WPA，WPA2采用安全性更高的区块密码锁链-信息真实性检查码协议CCMP（Counter Mode with CBC-MAC Protocol）加密算法。
  - 为了实现更好的兼容性，在目前的实现中，WPA和WPA2都可以使用802.1X的接入认证、TKIP或CCMP的加密算法，他们之间的不同主要表现在协议报文格式上，在安全性上几乎没有差别。
  - 综上所述，WPA/WPA2安全策略涉及了链路认证阶段、接入认证阶段、密钥协商和数据加密阶段。



- WLAN需要保障用户接入安全，即保障用户接入无线网络的合法性和安全性，STA接入WLAN网络前需要进行终端身份验证，即链路认证。链路认证通常被认为是终端连接AP并访问WLAN的起点。
- 共享密钥认证：
  - STA和AP预先配置相同的共享密钥，AP在链路认证过程验证两边的密钥配置是否相同。如果一致，则认证成功；否则，认证失败。
  - 认证过程：
    1. STA向AP发送认证请求（Authentication Request）。
    2. AP随即生成一个“挑战短语（Challenge）”发给STA。
    3. STA使用预先设置好的密钥加密“挑战短语”（EncryptedChallenge）并发给AP。
    4. AP接收到经过加密的“挑战短语”，用预先设置好的密钥解密该消息，然后将解密后的“挑战短语”与之前发送给STA的进行比较。如果相同，认证成功；否则，认证失败。

关联

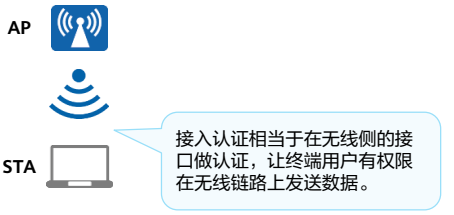
- 完成链路认证后，STA会继续发起链路服务协商，具体的协商通过Association报文实现。
- 终端关联过程实质上就是链路服务协商的过程，协商内容包括：支持的速率、信道等。



- 瘦接入点（FIT AP）架构中关联阶段处理过程：
  1. STA向AP发送Association Request请求，请求帧中会携带STA自身的各种参数以及根据服务配置选择的各种参数（主要包括支持的速率、支持的信道、支持的QoS的能力等）。
  2. AP收到Association Request请求帧后将其进行CAPWAP封装，并上报AC。
  3. AC收到关联请求后判断是否需要进行用户的接入认证，并回应Association Response。
  4. AP收到Association Response后将其进行CAPWAP解封装，并发给STA。

接入认证

- 接入认证即对用户进行区分，并在用户访问网络之前限制其访问权限。相对于链路认证，接入认证安全性更高。
- 主要包含：PSK认证和802.1X认证。

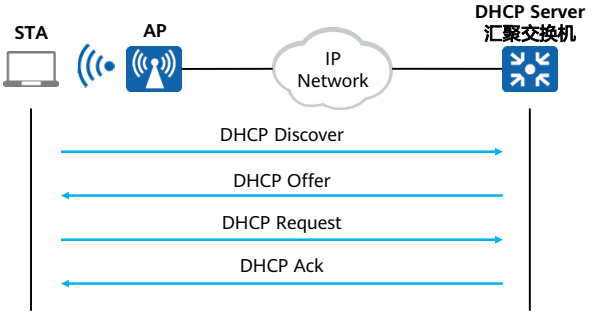


- 数据加密：
  - 除了用户接入认证外，对数据报文还需要使用加密的方式来保证数据安全，也是在接入认证阶段完成的。数据报文经过加密后，只有持有密钥的特定设备才可以对收到的报文进行解密，其他设备即使收到了报文，也因没有对应的密钥，无法对数据报文进行解密。

STA地址分配

扫描
链路认证
关联
接入认证
DHCP
用户认证

- STA获取到自身的IP地址，是STA正常上线的前提条件。
- 如果STA是通过DHCP方式获取IP地址，可以用AC设备或汇聚交换机作为DHCP服务器为STA分配IP地址。一般情况下使用汇聚交换机作为DHCP服务器。



用户认证

- 用户认证是一种“端到端”的安全结构，包括：802.1X认证、MAC认证和Portal认证。

扫描

链路认证

关联

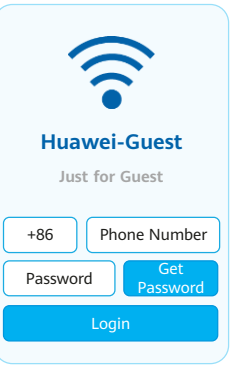
接入认证

DHCP

用户认证

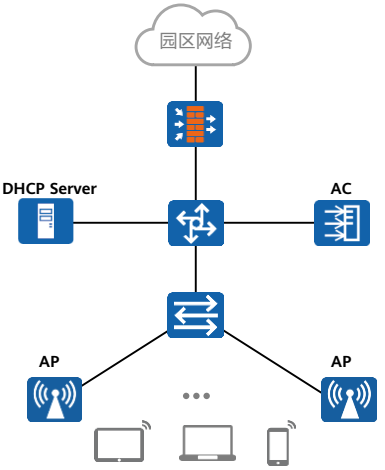
Portal认证

- 也称Web认证，一般将Portal认证网站称为门户网站。
- 用户上网时，必须在门户网站进行认证。只有认证通过后才可以使用网络资源。



- 随着企业网络的应用和发展，病毒、木马、间谍软件、网络攻击等各种信息安全威胁也在不断增加。在传统的企业网络建设思路中，一般认为企业内网是安全的，安全威胁主要来自外界。但是研究证明，80%的网络安全漏洞都存在于网络内部。它们对网络的破坏程度和范围持续扩大，经常引起系统崩溃、网络瘫痪。另外，内部员工在浏览某些网站时，一些间谍软件、木马程序等恶意软件也会不知不觉地被下载到电脑中，并且在企业内网传播，产生严重的安全隐患。
- 因此，随着安全挑战的不断升级，仅通过传统的安全措施已经远远不够。安全模型需要由被动模式向主动模式转变。从根源（终端）彻底解决网络安全问题，提高整个企业的信息安全水平。

WLAN工作流程：步骤4



WLAN工作流程

① AP上线

② WLAN业务配置下发

③ STA接入

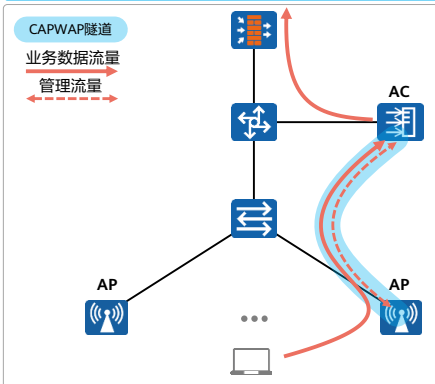
④ WLAN业务数据转发

CAPWAP中的数据包括控制报文（管理报文）和数据报文。
控制报文是通过CAPWAP的控制隧道转发的；
用户的数据报文分为隧道转发（又称为“集中转发”）方式和直接转发（又称为“本地转发”）方式。

HUAWEI

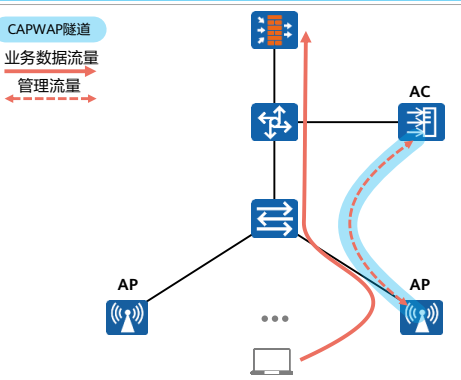
数据转发方式

隧道转发方式



- 隧道转发方式是指用户的数据报文到达AP后，需要经过CAPWAP数据隧道封装后发送给AC，然后由AC再转发到上层网络。

直接转发方式



- 直接转发方式是指用户的数据报文到达AP后，不经过CAPWAP的隧道封装而直接转发到上层网络。

HUAWEI

- 隧道转发方式：
  - 优点：AC集中转发数据报文，安全性好，方便集中管理和控制。
  - 缺点：业务数据必须经过AC转发，报文转发效率比直接转发方式低，AC所受压力大。
- 直接转发方式：
  - 优点：数据报文不需要经过AC转发，报文转发效率高，AC所受压力小。
  - 缺点：业务数据不便于集中管理和控制。

目录

WLAN的基础配置命令 - 配置AP上线 (1)

1. 配置AC作为DHCP服务器，配置Option 43字段

```
[AC-ip-pool-pool1] option code [ sub-option sub-code ] { ascii ascii-string | hex hex-string | cipher cipher-string | ip-address ip-address
```

配置DHCP服务器分配给DHCP客户端的自定义选项。

2. 创建域管理模板，并配置国家码

```
[AC] wlan
[AC-wlan-view]
```

进入WLAN视图。

```
[AC-wlan-view] regulatory-domain-profile name profile-name
[AC-wlan-regulate-domain-profile-name]
```

创建域管理模板，并进入模板视图，若模板已存在则直接进入模板视图。

```
[AC-wlan-regulate-domain-profile-name] country-code country-code
```

配置设备的国家码标识。

  HUAWEI

- 命令：**option** *code* [**sub-option** *sub-code*] { **ascii** ascii-string | **hex** hex-string | **cipher** *cipher-string* | **ip-address** *ip-address*
  - *code*：指定自定义选项Option的数值。整数形式，取值范围1～254，但1、3、6、15、44、46、50、51、52、53、54、55、57、58、59、61、82、121、184不能配置。
  - **sub-option** *sub-code*：指定自定义的Option子选项的数值。整数形式，取值范围是1～254。知名选项请参考RFC2132。
  - **ascii** | **hex** | **cipher**：指定自定义的选项码为ASCII字符串类型，或十六进制字符串类型，或密文字符串类型。
  - **ip-address** *ip-address*：指定自定义的选项码为IP地址类型。
- 命令：**regulatory-domain-profile name** *profile-name*
  - **name** *profile-name*：指定域管理模板的名称。字符串类型，不区分大小写，可输入的字符串长度为1～35个字符。可见字符，不能包含“?”和空格，双引号不能出现在字符串的首尾。
- 命令：**country-code** *country-code*
  - *country-code*：指定设备的国家码标识。字符串格式，枚举类型。
  - AC支持的国家码有很多，如：
    - CN 中国（缺省值）
    - AU 澳大利亚
    - CA 加拿大
    - DE 德国
    - FR 法国
    - US 美国
    - ……

WLAN的基础配置命令 - 配置AP上线 (2)

[AC-wlan-view] **ap-group name** *group-name*
[AC-wlan-ap-group-group-name]

创建AP组，并进入AP组视图，若AP组已存在则直接进入AP组视图。

[AC-wlan-ap-group-group-name] **regulatory-domain-profile** *profile-name*

将指定的域管理模板引用到AP或AP组。

3. 配置源接口或源地址

[AC] **capwap source interface** { **loopback** *loopback-number* | **vlanif** *vlan-id* }

配置AC与AP建立CAPWAP隧道的源接口。

[AC] **capwap source ip-address** *ip-address*

配置AC的源IP地址。

- 命令：**ap-group name** *group-name*
  - **name** *group-name*：指定AP组的名称。字符串类型，可输入的字符串长度为1～35个字符。可见字符，不能包含“?”、“/”和空格，双引号不能出现在字符串的首尾。

WLAN的基础配置命令 - 配置AP上线 (3)

4. 添加AP设备 – 离线导入AP

```
[AC-wlan-view] ap auth-mode { mac-auth | sn-auth }
```

配置AP认证模式为MAC地址认证，或SN认证，缺省为MAC地址认证。

```
[AC-wlan-view] ap-id ap-id [ [ type-id type-id | ap-type ap-type ] { ap-mac ap-mac | ap-sn ap-sn | ap-mac ap-mac ap-sn ap-sn } ]
[AC-wlan-ap-ap-id] ap-name ap-name
```

离线增加AP设备或进入AP视图，并配置单个AP的名称。

```
[AC-wlan-view] ap-id 0
[AC-wlan-ap-0] ap-group ap-group
```

配置AP所加入的组。

5. 检查AP上线结果

```
[AC] display ap { all | ap-group ap-group }}
```

查看AP信息。

- 命令：**ap-id** *ap-id* [[**type-id** *type-id* | **ap-type** *ap-type*] { **ap-mac** *ap-mac* | **ap-sn** *ap-sn* | **ap-mac** *ap-mac* **ap-sn** *ap-sn* }]
  - *ap-id*：AP设备索引。整数类型，取值范围：0～8191。
  - **type-id** *type-id*：AP设备类型索引。整数类型，取值范围：0～255。
  - **ap-type** *ap-type*：AP设备类型。字符串类型，取值范围为1～31个字符。
  - **ap-mac** *ap-mac*：AP的MAC地址。格式为H-H-H，其中H为4位的十六进制数。
  - **ap-sn** *ap-sn*：AP的序列号。字符串类型，取值范围为1～31个字符，只能包括字母和数字。

WLAN的基础配置命令 - 配置射频 (1)

1. 进入射频视图

```
[AC-wlan-view] ap-id 0
[AC-wlan-ap-0] radio radio-id
[AC-wlan-radio-0]
```

2. 配置指定射频的工作带宽和信道

```
[AC-wlan-radio-0/0] channel { 20mhz | 40mhz-minus | 40mhz-plus | 80mhz | 160mhz } channel
Warning: This action may cause service interruption. Continue?[Y/N]y
```

```
[AC-wlan-radio-0/0] channel 80+80mhz channel1 channel2
Warning: This action may cause service interruption. Continue?[Y/N]y
```

配置AP组中所有AP或单个AP指定射频的工作带宽和信道。

3. 配置天线的增益

```
[AC-wlan-radio-0/0] antenna-gain antenna-gain
```

配置AP组中所有AP或单个AP指定射频的天线增益。

HUAWEI

- 命令：**radio** *radio-id*
  - *radio-id*：射频ID。必须是已存在的射频ID。
- 命令：
  - **channel** { **20mhz** | **40mhz-minus** | **40mhz-plus** | **80mhz** | **160mhz** } *channel*
  - **channel 80+80mhz** *channel1 channel2*
  - 20mhz：指定射频的工作带宽为20MHz。
  - 40mhz-minus：指定射频的工作带宽为40MHz Minus。
  - 40mhz-plus：指定射频的工作带宽为40MHz Plus。
  - 80mhz：指定射频的工作带宽为80MHz。
  - 160mhz：指定射频的工作带宽为160MHz。
  - 80+80mhz：指定射频的工作带宽为80+80MHz。
  - *channel/channel1/channel2*：指定射频的工作信道，信道基于国家代码和射频模式来进行选择。枚举值类型，取值范围根据国家代码和射频模式来进行选择。
- 命令：**antenna-gain** *antenna-gain*
  - *antenna-gain*：天线增益。整数类型，取值范围：0～30，单位：dB。

WLAN的基础配置命令 - 配置射频 (2)

4. 配置射频的发射功率

```
[AC-wlan-radio-0/0] eirp eirp
```

配置AP组中所有AP或单个AP指定射频的发射功率。

5. 配置射频覆盖距离参数

```
[AC-wlan-radio-0/0] coverage distance distance
```

配置AP组中所有AP或单个AP指定射频的射频覆盖距离参数。

6. 配置射频工作的频段

```
[AC-wlan-radio-0/0] frequency { 2.4g | 5g }
```


- 命令：**eirp** *eirp*
  - *eirp*：发射功率值。整数形式，取值范围是1～127，单位：dBm。
- 命令：**coverage distance** *distance*
  - *distance*：射频覆盖距离参数。每个射频覆盖距离参数对应一组slottime、acktimeout和ctstimeout数值。根据AP间的实际距离配置射频覆盖距离参数后，AP设备根据此参数值调整对应的slottime、acktimeout和ctstimeout数值。整数类型，取值范围：1～400，单位为100m。
- 命令：**frequency { 2.4g | 5g }**
  - 缺省情况下，射频0工作在2.4GHz频段，射频2工作在5GHz频段。

WLAN的基础配置命令 - 配置射频 (3)

7. 创建射频模板

```
[AC-wlan-view] radio-2g-profile name profile-name
```

创建2G射频模板，并进入模板视图，若模板已存在则直接进入模板视图。

8. 引用射频模板

```
[AC-wlan-view] ap-group name group-name
[AC-wlan-ap-group-group-name] radio-2g-profile profile-name radio { radio-id | all }
```

在AP组中，将指定的2G射频模板引用到2G射频。

- 命令：**radio-2g-profile name** *profile-name*
  - **name** *profile-name*：指定2G射频模板的名称。字符串类型，不区分大小写，可输入的字符串长度为1～35个字符。可见字符，不能包含“?”和空格，双引号不能出现在字符串的首尾。
  - 缺省情况下，系统上存在名为default的2G射频模板。
- 命令：**radio-2g-profile** *profile-name* **radio** { *radio-id* | **all** }
  - *profile-name*：指定2G射频模板的名称。必须是已存在的2G射频模板名称。
  - **radio** *radio-id*：指定射频的ID。整数类型，取值范围：0和2。
  - **radio all**：指定所有的射频。

WLAN的基础配置命令 - 配置VAP (1)

1. 创建VAP模板

```
[AC-wlan-view] vap-profile name profile-name
[AC-wlan-vap-prof-profile-name]
```

创建VAP模板，并进入模板视图，若模板已存在则直接进入模板视图。

2. 配置数据转发方式

```
[AC-wlan-vap-prof-profile-name] forward-mode { direct-forward | tunnel }
```

配置VAP模板下的数据转发方式，可以是直接转发或隧道转发。

3. 配置业务VLAN

```
[AC-wlan-vap-prof-profile-name] service-vlan { vlan-id vlan-id | vlan-pool pool-name }
```

配置VAP的业务VLAN。

WLAN的基础配置命令 - 配置VAP (2)

4. 配置安全模板

```
[AC-wlan-view] security-profile name profile-name
[AC-wlan-sec-prof-profile-name]
```

创建安全模板或者进入安全模板视图。

缺省情况下，系统已经创建名称为default、default-wds和default-mesh的安全模板。

```
[AC-wlan-view] vap-profile name profile-name
[AC-wlan-vap-prof-profile-name] security-profile profile-name
```

在指定VAP模板中引用安全模板。

WLAN的基础配置命令 - 配置VAP (3)

5. 配置SSID模板

```
[AC-wlan-view] ssid-profile name profile-name
[AC-wlan-ssid-prof-profile-name]
```

创建SSID模板，并进入模板视图，若模板已存在则直接进入模板视图。
缺省情况下，系统上存在名为default的SSID模板。

```
[AC-wlan-ssid-prof-profile-name] ssid ssid
```

配置当前SSID模板中的服务组合识别码SSID（Service Set Identifier）。
缺省情况下，SSID模板中的SSID为HUAWEI-WLAN。

```
[AC-wlan-view] vap-profile name profile-name
[AC-wlan-vap-prof-profile-name] ssid-profile profile-name
```

在指定VAP模板中引用SSID模板。

- 命令：**ssid** *ssid*
  - *ssid*：指定SSID的名称。文本类型，区分大小写，可输入的字符串长度为1～32字符，支持中文字符，也支持中英文字符混合，不支持制表符。
  - 如果想设置SSID首字符为空格，则输入的SSID内容应该以“"”开头以“"”结束，如" hello"，其中前后的“"”占用两个字符。如果想设置SSID首字符为“"”，则需要在“"”前输入转义字符“\”，如\"hello，其中“\”占用一个字符。

WLAN的基础配置命令 - 配置VAP (4)

6. 引用VAP模板

```
[AC-wlan-view] ap-group name group-name
[AC-wlan-ap-group-group-name] vap-profile profile-name wlan wlan-id radio { radio-id | all } [ service-vlan
{ vlan-id vlan-id | vlan-pool pool-name } ]
```

在AP组中，将指定的VAP模板引用到射频。

7. 查看VAP信息

```
[AC] display vap { ap-group ap-group-name | { ap-name ap-name | ap-id ap-id } [ radio radio-id ] } [ ssid ssid ]
```

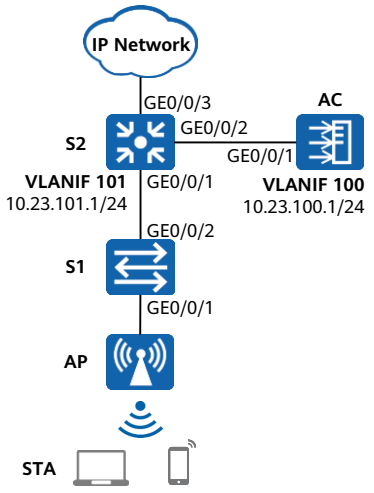
```
[AC] display vap { all | ssid ssid }
```

查看业务型VAP的相关信息。

  HUAWEI

- 命令：**display vap** { **ap-group** *ap-group-name* | { **ap-name** *ap-name* | **ap-id** *ap-id* } [**radio** *radio-id*] } [**ssid** *ssid*]
  - **ap-group** *ap-group-name*：查看指定AP组下的所有业务型VAP的相关信息。必须是已存在的AP组名称。
  - **ap-name** *ap-name*：查看指定名称的AP的业务型VAP的相关信息。必须是已存在的AP名称。
  - **ap-id** *ap-id*：查看指定ID的AP的业务型VAP的相关信息。必须是已存在的AP ID。
  - **radio** *radio-id*：查看指定射频的业务型VAP的相关信息。整数类型，取值范围：0～2。
  - **ssid** *ssid*：查看指定SSID的业务型VAP的相关信息。必须是已存在的SSID。
- 命令：**display vap** { **all** | **ssid** *ssid* }
  - **all**：查看所有业务型VAP的相关信息。

案例：旁挂二层组网隧道转发

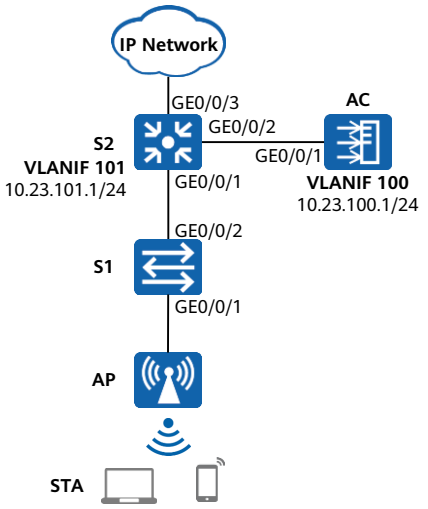


数据	配置
AP管理VLAN	VLAN100
STA业务VLAN	VLAN101
DHCP服务器	AC作为DHCP服务器为AP分配IP地址 汇聚交换机S2作为DHCP服务器为STA分配IP地址，STA的默认网关为10.23.101.1
AP的IP地址池	10.23.100.2～10.23.100.254/24
STA的IP地址池	10.23.101.2～10.23.101.254/24
AC的源接口IP地址	VLANIF100：10.23.100.1/24
AP组	名称：ap-group1；引用模板：VAP模板wlan-net、域管理模板
域管理模板	名称：default 国家码：中国
SSID模板	名称：wlan-net SSID名称：wlan-net
安全模板	名称：wlan-net 安全策略：WPA-WPA2+PSK+AES 密码：a1234567
VAP模板	名称：wlan-net 转发模式：隧道转发 业务VLAN：VLAN101 引用模板：SSID模板wlan-net、安全模板wlan-net

HUAWEI

- 业务需求
  - 企业用户通过WLAN接入网络，以满足移动办公的最基本需求。
- 组网需求
  - AC组网方式：旁挂二层组网。
  - DHCP部署方式：
    - AC作为DHCP服务器为AP分配IP地址。
    - 汇聚交换机S2作为DHCP服务器为STA分配IP地址。
  - 业务数据转发方式：隧道转发。
- 配置思路：
  - 配置AP、AC和周边网络设备之间实现网络互通。
  - 配置AP上线。
    - 创建AP组，用于将需要进行相同配置的AP都加入到AP组，实现统一配置。
    - 配置AC的系统参数，包括国家码、AC与AP之间通信的源接口。
    - 配置AP上线的认证方式并离线导入AP，实现AP正常上线。
  - 配置WLAN业务参数，实现STA访问WLAN网络功能。

配置网络互通



1、S1、S2、AC创建对应VLAN及接口。

2、配置DHCP服务器为STA和AP分配IP地址。

在AC上配置VLANIF100接口为AP提供IP地址。

```
[AC] dhcp enable
[AC] interface vlanif 100
[AC-Vlanif100] ip address 10.23.100.1 24
[AC-Vlanif100] dhcp select interface
```

在S2上配置VLANIF101接口为STA提供IP地址，并指定10.23.101.1作为STA的默认网关地址。

```
[S2] dhcp enable
[S2] interface vlanif 101
[S2-Vlanif101] ip address 10.23.101.1 24
[S2-Vlanif101] dhcp select interface
```


- 1、S1、S2、AC创建对应VLAN及接口。
  - S1配置：

```
[S1] vlan batch 100
[S1] interface gigabitethernet 0/0/1
[S1-GigabitEthernet0/0/1] port link-type trunk
[S1-GigabitEthernet0/0/1] port trunk pvid vlan 100
[S1-GigabitEthernet0/0/1] port trunk allow-pass vlan 100
[S1-GigabitEthernet0/0/1] quit
[S1] interface gigabitethernet 0/0/2
[S1-GigabitEthernet0/0/2] port link-type trunk
[S1-GigabitEthernet0/0/2] port trunk allow-pass vlan 100
[S1-GigabitEthernet0/0/2] quit
```

- S2配置：

  [S2] vlan batch 100 101

  [S2] interface gigabitethernet 0/0/1

  [S2-GigabitEthernet0/0/1] port link-type trunk

  [S2-GigabitEthernet0/0/1] port trunk allow-pass vlan 100

  [S2-GigabitEthernet0/0/1] quit

  [S2] interface gigabitethernet 0/0/2

  [S2-GigabitEthernet0/0/2] port link-type trunk

  [S2-GigabitEthernet0/0/2] port trunk allow-pass vlan 100 101

  [S2-GigabitEthernet0/0/2] quit

  [S2] interface gigabitethernet 0/0/3

  [S2-GigabitEthernet0/0/3] port link-type trunk

  [S2-GigabitEthernet0/0/3] port trunk allow-pass vlan 101

  [S2-GigabitEthernet0/0/3] quit

- AC配置

  [AC] vlan batch 100 101

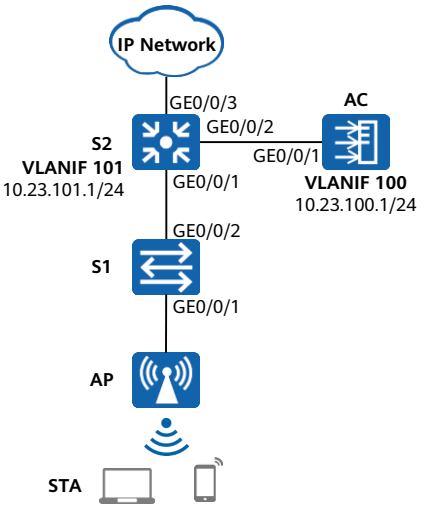
  [AC] interface gigabitethernet 0/0/1

  [AC-GigabitEthernet0/0/1] port link-type trunk

  [AC-GigabitEthernet0/0/1] port trunk allow-pass vlan 100 101

  [AC-GigabitEthernet0/0/1] quit

配置AP上线 (1)



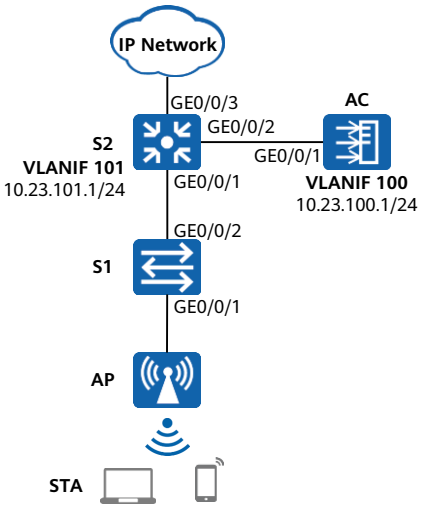
1、创建AP组。

```
[AC] wlan
[AC-wlan-view] ap-group name ap-group1
[AC-wlan-ap-group-ap-group1] quit
```

2、创建域管理模板，并配置AC的国家码。

```
AC-wlan-view] regulatory-domain-profile name default
[AC-wlan-regulate-domain-default] country-code cn
[AC-wlan-regulate-domain-default] quit
[AC-wlan-view] ap-group name ap-group1
[AC-wlan-ap-group-ap-group1] regulatory-domain-profile default
Warning: Modifying the country code will clear channel, power and
antenna gain configurations of the radio and reset the AP. Continu
e?[Y/N]:y
[AC-wlan-ap-group-ap-group1] quit
[AC-wlan-view] quit
```

配置AP上线 (2)



3、配置AC的源接口。

```
[AC] capwap source interface vlanif 100
```

4、在AC上离线导入AP。

```
[AC] wlan
[AC-wlan-view] ap auth-mode mac-auth
[AC-wlan-view] ap-id 0 ap-mac 60de-4476-e360
[AC-wlan-ap-0] ap-name area_1
Warning: This operation may cause AP reset. Continue? [Y/N]:y
[AC-wlan-ap-0] ap-group ap-group1
Warning: This operation may cause AP reset. If the country code
changes, it will clear channel, power and antenna gain
configurations of the radio, Whether to continue? [Y/N]:y
[AC-wlan-ap-0] quit
```


- 在AC上离线导入AP
  - 将AP加入AP组“ap-group1”中。假设AP的MAC地址为60de-4476-e360，并且根据AP的部署位置为AP配置名称，便于从名称上就能够了解AP的部署位置。例如MAC地址为60de-4476-e360的AP部署在1号区域，命名此AP为area_1。

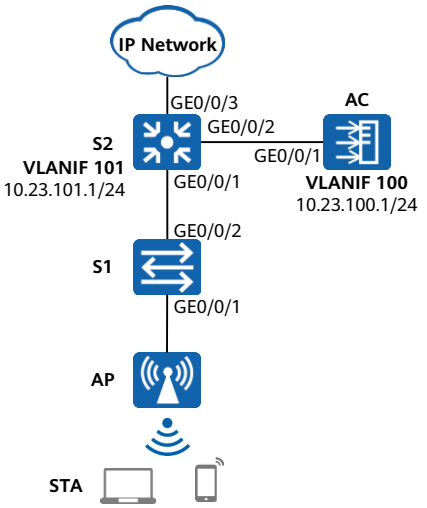
查看AP上线

- 将AP上电后，当执行命令display ap all查看到AP的“State”字段为“nor”时，表示AP正常上线。

```
[AC-wlan-view] display ap all
Total AP information:
nor  : normal          [1]
Extra information:
P  : insufficient power supply
--------------------------------------------------------------------------------
ID   MAC             Name      Group      IP              Type        State STA Uptime    ExtraInfo
--------------------------------------------------------------------------------
0    60de-4476-e360  area_1    ap-group1  10.23.100.254   AP5030DN    nor   0   10S       -
--------------------------------------------------------------------------------
Total: 1
```


- **display ap**命令输出信息描述：
  - ID：AP ID。
  - MAC：AP MAC地址。
  - Name：AP名称。
  - Group：AP所属的AP组名称。
  - IP：AP的IP地址。在NAT场景下，AP在私网侧，AC在公网侧，该值为AP私网侧的IP地址。可通过命令display ap run-info查看AP公网侧IP地址。
  - Type：AP类型。
  - State：AP状态。
    - normal：AP正常状态，指AP在AC上成功上线。
    - commit-failed：AP上线后WLAN业务配置下发失败状态。
    - download：AP正在升级状态。
    - fault：AP上线失败状态。
    - idle：AP和AC建链前的初始状态。
  - STA：AP上接入的终端用户数。
  - Uptime：AP已上线时长。
  - ExtraInfo：额外的信息。P表示设备供电不足。

配置WLAN业务参数 (1)



1、创建名为“wlan-net”的安全模板，并配置安全策略。

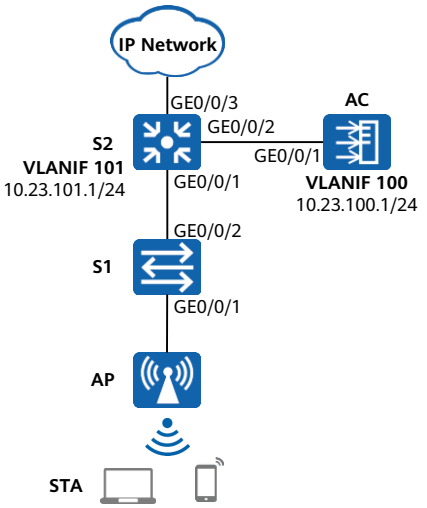
```
[AC-wlan-view] security-profile name wlan-net
[AC-wlan-sec-prof-wlan-net] security wpa-wpa2 psk pass-phrase a1234567 aes
[AC-wlan-sec-prof-wlan-net] quit
```

2、创建名为“wlan-net”的SSID模板，并配置SSID名称为“wlan-net”。

```
[AC-wlan-view] ssid-profile name wlan-net
[AC-wlan-ssid-prof-wlan-net] ssid wlan-net
[AC-wlan-ssid-prof-wlan-net] quit
```


HUAWEI

配置WLAN业务参数 (2)



3、创建名为“wlan-net”的VAP模板，配置业务数据转发模式、业务VLAN，并且引用安全模板和SSID模板。

```
[AC-wlan-view] vap-profile name wlan-net
[AC-wlan-vap-prof-wlan-net] forward-mode tunnel
[AC-wlan-vap-prof-wlan-net] service-vlan vlan-id 101
[AC-wlan-vap-prof-wlan-net] security-profile wlan-net
[AC-wlan-vap-prof-wlan-net] ssid-profile wlan-net
[AC-wlan-vap-prof-wlan-net] quit
```

4、配置AP组引用VAP模板，AP上射频0和射频1都使用VAP模板“wlan-net”的配置。

```
[AC-wlan-view] ap-group name ap-group1
[AC-wlan-ap-group-ap-group1] vap-profile wlan-net wlan 1 radio 0
[AC-wlan-ap-group-ap-group1] vap-profile wlan-net wlan 1 radio 1
[AC-wlan-ap-group-ap-group1] quit
```

查看VAP模板信息

- WLAN业务配置会自动下发给AP，配置完成后，通过执行命令display vap ssid wlan-net查看如下信息，当“Status”项显示为“ON”时，表示AP对应的射频上的VAP已创建成功。

```
[AC-wlan-view] display vap ssid wlan-net
WID : WLAN ID
------------------------------------------------------------------------------------------------------------
AP ID      AP name    RfID  WID  BSSID              Status  Auth type        STA  SSID
------------------------------------------------------------------------------------------------------------
0          area_1     0     1    60DE-4476-E360     ON      WPA/WPA2-PSK     0    wlan-net
0          area_1     1     1    60DE-4476-E370     ON      WPA/WPA2-PSK     0    wlan-net
------------------------------------------------------------------------------------------------------------
Total: 2
```

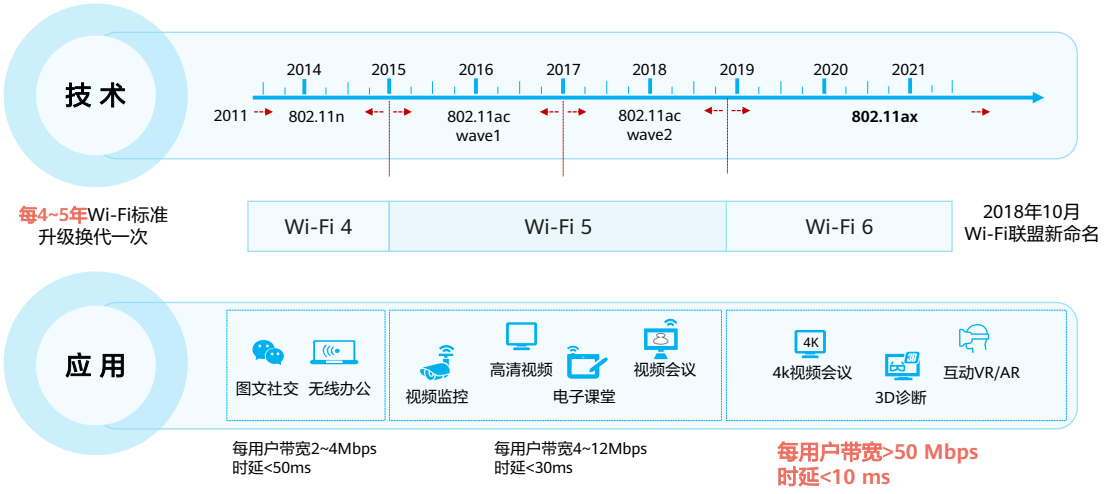

- **display vap**命令输出信息描述：
  - AP ID：AP ID。
  - AP name：AP名称。
  - RfID：射频ID。
  - WID：VAP的ID。
  - SSID：SSID的名称。
  - BSSID：VAP的MAC地址。
  - Status：VAP当前状态：
    - ON：VAP服务开启
    - OFF：VAP服务关闭
  - Auth type：VAP认证方式。
  - STA：当前VAP接入的终端数。

目录

华为WLAN方案满足未来无线建网需求

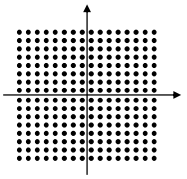
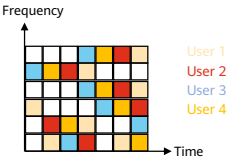
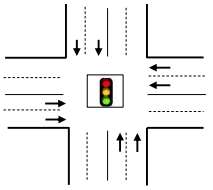
全场景	• 面对复杂多样的应用场景，采用场景定制化解决方案 • 园区网络、分支网络均有完整的WLAN部署、管理方案
大带宽	• 支持802.11ac wave2协议，双5G射频覆盖，无线接入带宽最高可达3.46 Gbps • 华为主导制定下一代802.11ax标准（Wi-Fi 6），单5G速率高达9.6 Gbps • 支持无线漫游及WMM等多种无线QoS协议，保证业务质量
高安全	• 支持WPA/WPA2/WPA3/WAPI等主流认证/加密方式 • 支持无线入侵检测 • 可通过Portal、802.1x技术对用户进行进行身份认证，保护内网安全
易部署	• AP支持即插即用，自动升级，信道自主选择，设备速率和功率动态调整，支持负载均衡 • 物联网融合AP、内置高密天线AP等特色产品，简化安装，快速部署 • 支持云管理模式，AP采用双栈设计，本地管理与云管理平滑切换

双轮驱动：技术与应用发展助推Wi-Fi 6时代到来



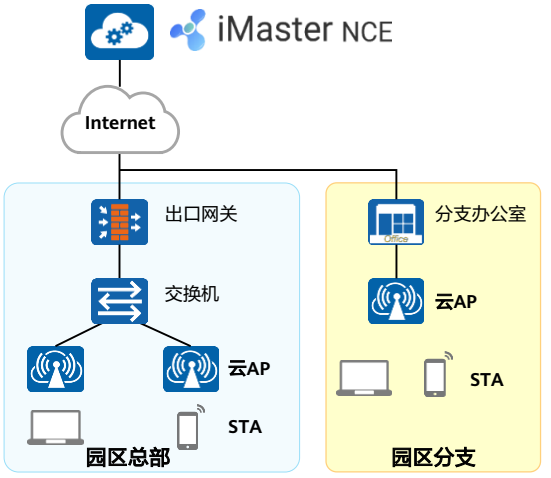
- Wi-Fi 5无法满足4K/8K视频会议场景低业务时延，高带宽需求。
- Wi-Fi 6配合华为SmartRadio智能应用加速，可以将时延降低至10 ms。

Wi-Fi 6 Vs Wi-Fi 5

大带宽	高并发	低时延	低耗电
			
1024-QAM 8x8 MU-MIMO	UL/DL OFDMA UL/DL MU-MIMO	OFDMA Spatial Reuse	TWT 20MHz-Only
● 速率高达 9.6 Gbps ● 带宽提 升 4 倍	● 每AP接入 1024 终端 ● 并发用户数提升 4 倍	● 业务时延低至 20ms ● 平均时延降低 30%	● 目标时间唤醒机制 ● 终端功耗降低 30%

- 带宽提升4倍，是按理论速率进行对比，目前所有Wi-Fi 5产品实现理论（wave2）速率为2.5G；Wi-Fi 6标准理论速率为9.6G。
- 并发用户数提升4倍，实际测试人均2M情况下，Wi-Fi 5并发100用户，Wi-Fi 6并发用户400用户。
- Wi-Fi 6中平均时延在20 ms左右（Wi-Fi 5中平均时延在30 ms左右），后面利用华为SmartRadio智能应用加速技术后，业务时延可再降低至10ms。
- TWT：Wi-Fi 5不支持。

下一代园区网络：智简园区（中小型园区网络）



基本概念

- 通过云管理平台，可以实现任意地点对设备进行集中的管理和维护，大大降低网络部署运维成本。
- 适用范围：中小型企业。

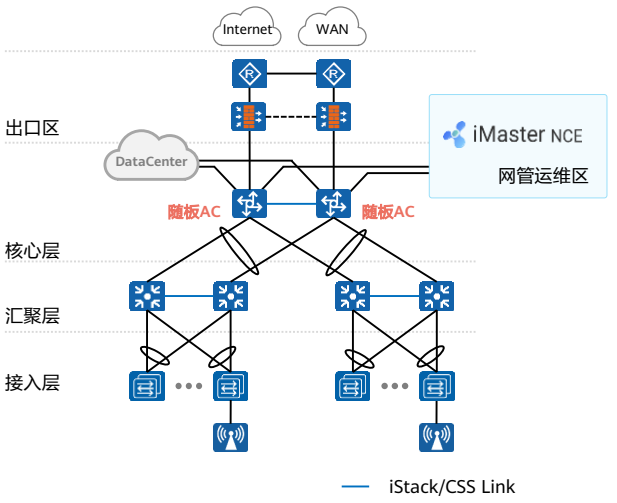
优势（对比AC+FIT AP架构）

- 即插即用，自动开局，减少网络部署成本。
- 统一运维：所有云管理网元统一在云管理平台上进行监控和管理。
- 工具化：通常情况下，云解决方案会在云端提供各类工具，有效降低各类开支。

HUAWEI

- 传统方案的弊端：
  - 在部署网络时，传统网络解决方案会存在部署成本高、后期运维困难等问题，尤其是对于分支站点数量多、站点地域分散的企业，这些问题尤为明显。
- 云管理架构：
  - 云管理架构可以很好的解决以上问题，通过云管理平台，实现任意地点对设备进行集中的管理和维护，大大降低网络部署运维成本。
  - 当云AP布放完成后，无须网络管理员到安装现场对云AP进行软件调试，云AP上电后即可自动连接到指定的云管理平台加载指定的配置文件、软件包和补丁文件等系统文件，实现云AP零配置上线。网络管理员可以随时随地通过云管理平台统一给AP下发配置，使业务批量配置更快捷。

下一代园区网络：智简园区（大中型园区网络）



架构特点

- AP需要配合iMaster NCE使用，由iMaster NCE统一管理和配置，功能丰富，进一步和有线网络融合，结合大数据和AI技术实现园区网络的极简、智慧和安全。
- 适用范围：大中型企业。

HUAWEI

思考题

1. 直连式组网和旁挂式组网各有什么优势?
2. （多选）FIT AP发现AC的方式有哪些?（　　）
   A. 静态发现
   B. DHCP动态发现
   C. FTP动态发现
   D. DNS动态发现

1. 答案：
   - 直连式组网优势：在直连式组网中，多采用直接转发模式，适用于大规模集中部署的WLAN网络，并可以简化网络架构。
   - 旁挂式组网优势：这种方式是常用的组网模式，此时无线用户业务数据无需经过AC集中处理，基本无带宽瓶颈，而且便于继承现有网络的安全策略，故此模式也多是推荐的网络部署方案。
2. ABD

本章总结

- 通过WLAN技术，用户可以方便地接入到无线网络，并在无线网络覆盖区域内自由移动，彻底摆脱有线网络的束缚。
- 本章主要介绍了企业网络WLAN技术，包括：WLAN的基本概念、WLAN的工作原理、WLAN的组网架构、WLAN的配置实现和WLAN技术发展趋势。


谢 谢
www.huawei.com

广域网技术

前言

- 随着经济全球化与数字化变革加速，企业规模不断扩大，越来越多的分支机构出现在不同的地域。每个分支的网络被认为一个LAN（Local Area Network，局域网），总部和各分支机构之间通信需要跨越地理位置。因此，企业需要通过WAN（Wide Area Network，广域网）将这些分散在不同地理位置的分支机构连接起来，以便更好地开展业务。
- 广域网技术的发展，伴随着带宽不断的升级：早期出现的X.25只能提供64 kbit/s的带宽，其后DDN（Digital Data Network，数字数据网）和FR（Frame Relay，帧中继）提供的带宽提高到2 Mbit/s，SDH（Synchronous Digital Hierachy，同步数字结构）和ATM（Asynchronous Transfer Mode，异步传输模式）进一步把带宽提升到10 Gbit/s，最后发展到当前以IP为基础的10 Gbit/s甚至更高带宽的广域网络。
- 本课程主要讲解广域网技术基础概述以及PPP（Point-to-Point Protocol，点对点协议）原理与相关应用。

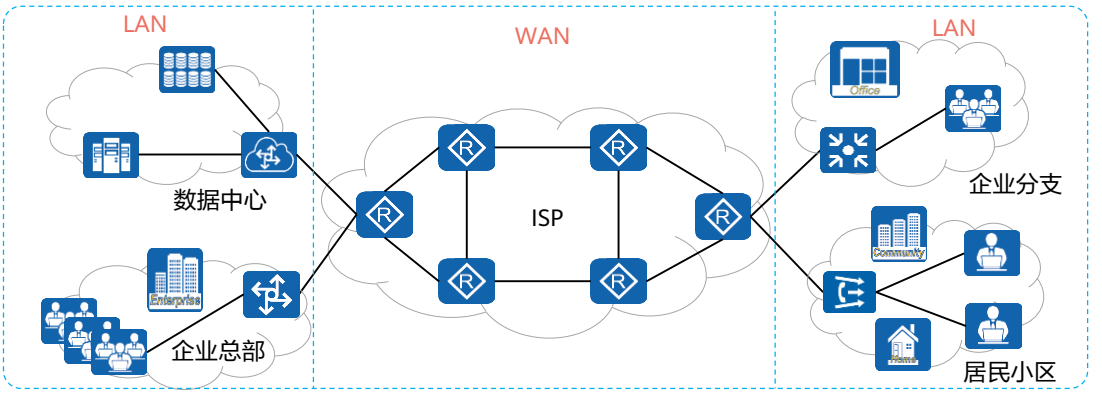
目标

- 学完本课程后，您将能够：
  - 了解广域网基本概念和发展历史
  - 掌握PPP和PPPoE的工作原理
  - 掌握PPP和PPPoE的基本配置
  - 了解MPLS/SR相关技术的基本概念

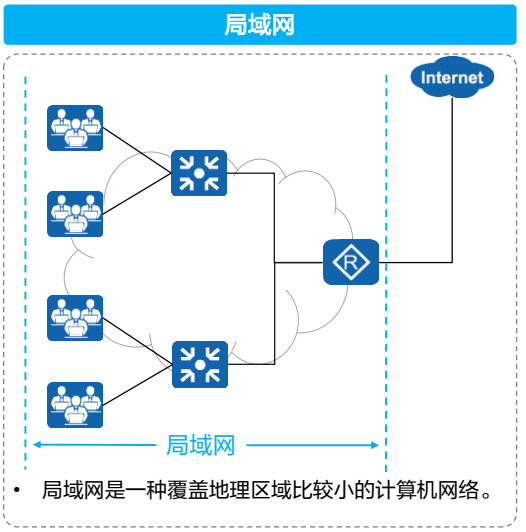
目录

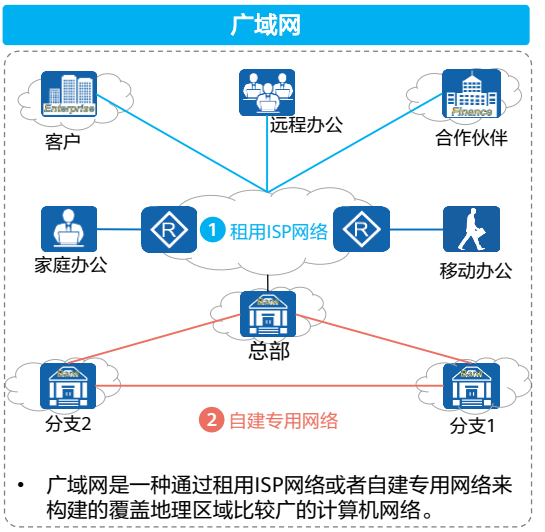
什么是广域网

- **广域网**是连接不同地区局域网的网络，通常所覆盖的范围从几十公里到几千公里。它能连接多个地区、城市和国家，或横跨几个洲提供远距离通信，形成国际性的远程网络。



广域网与局域网区别

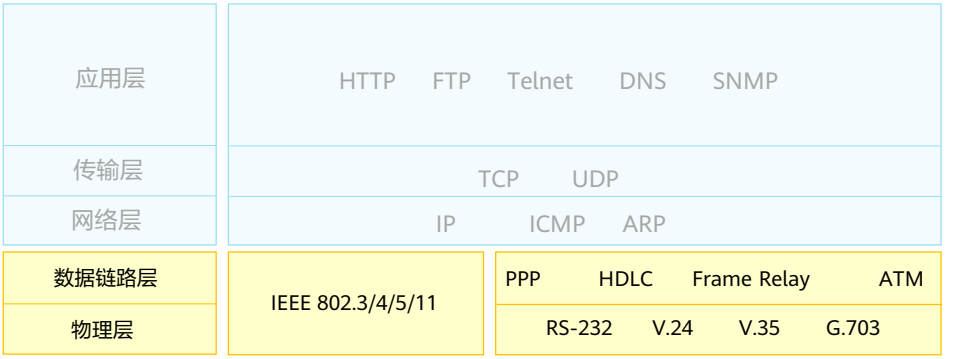




- 广域网与局域网的区别主要体现在以下几个方面：
  - 局域网带宽高但是传输距离短，无法满足广域网长距离传输；
  - 局域网设备通常都是交换机，广域网设备大多都是路由器；
  - 局域网属于某一个单位或者组织，广域网服务大多由ISP提供；
  - 广域网与局域网一般仅在物理层和数据链路层采用不同的协议或技术，其他层次基本没有差异；
  - 银行、政府、军队、大型公司的专用网络也属于广域网，且与Internet实现物理隔离；
  - Internet只是广域网的一种，小企业借用Internet作为广域网连接。

早期广域网技术介绍

- 早期广域网与局域网的区别在于数据链路层和物理层的差异性，在TCP/IP参考模型中，其他各层无差异。

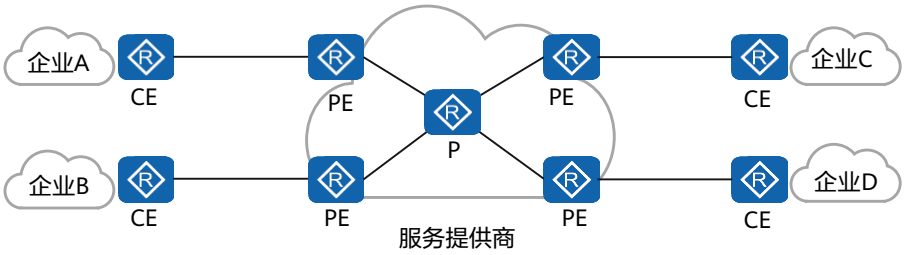


HUAWEI

- 初期广域网常用的物理层标准有EIA（Electronic Industries Alliance，电子工业协会）和TIA（Telecommunications Industry Association，电信工业协会）制定的公共物理层接口标准EIA/TIA-232（即RS-232）、 ITU（International Telecommunication Union，国际电信联盟）制定的串行线路接口标准V.24和V.35，以及有关各种数字接口的物理和电气特性的G.703标准等。
- 广域网常见的数据链路层标准有：HDLC（High-level Data Link Control，高级数据链路控制）、PPP（Point-to-Point Protocol，点到点协议）、FR（Frame Relay，帧中继）、ATM异步传输模式等，其中：
  - HDLC协议是一种通用的协议，工作在数据链路层。数据报文加上头开销和尾开销后封装成HDLC帧，只支持在点到点的同步链路上的数据传输，不支持IP地址协商与认证，过于追求高可靠性，导致数据帧开销较大，传输效率较低。
  - PPP协议工作在数据链路层，主要用在支持全双工的同、异步链路上，进行点到点之间的数据传输。由于它能够提供用户认证，易于扩充，并且支持同、异步通信，因而获得广泛应用。
  - 帧中继是一种工业标准的、交换式的数据链路协议，通过使用无差错校验机制，加快了数据转发速度。
  - ATM是建立在电路交换和分组交换基础上的一种面向连接的交换技术，ATM传送信息的基本载体是53 Byte固定长度ATM信元。

广域网络设备角色介绍

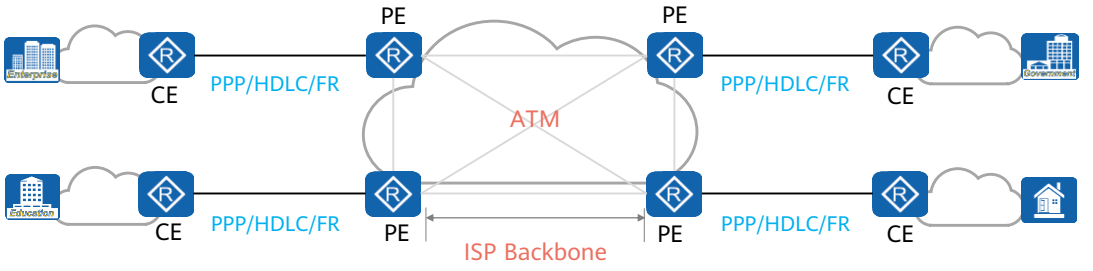
- 广域网络设备基本角色有三种，CE（Customer Edge，用户边缘设备）、PE（Provider Edge，服务提供商边缘设备）和P（Provider，服务提供商设备）。具体定义是：
  - CE：用户端连接服务提供商的边缘设备。CE连接一个或多个PE，实现用户接入。
  - PE：服务提供商连接CE的边缘设备。PE同时连接CE和P设备，是重要的网络节点。
  - P：服务提供商不连接任何CE的设备。



HUAWEI

早期广域网技术的应用

- 早期的广域网技术主要是针对不同的物理链路类型，在数据链路层进行不同的二层封装。在CE与PE之间常用的广域网封装协议有PPP/HDLC/FR等，用于解决用户接入广域网的长距离传输问题。在ISP内部常用的广域网协议主要是ATM，它用于解决骨干网高速转发的问题。



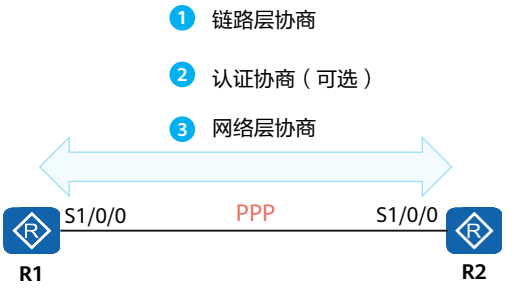
目录

PPP协议概述

- PPP（Point-to-Point Protocol，点到点协议）是一种常见的广域网数据链路层协议，主要用于在全双工的链路上进行点到点的数据传输封装。
- PPP提供了安全认证协议族PAP（Password Authentication Protocol，密码验证协议）和CHAP（Challenge Handshake Authentication Protocol，挑战握手认证协议）。
- PPP协议具有良好的扩展性，例如，当需要在以太网链路上承载PPP协议时，PPP可以扩展为PPPoE。
- PPP协议提供LCP（Link Control Protocol，链路控制协议），用于各种链路层参数的协商，例如最大接收单元，认证模式等。
- PPP协议提供各种NCP（Network Control Protocol，网络控制协议），如IPCP（IP Control Protocol，IP控制协议），用于各网络层参数的协商，更好地支持了网络层协议。

PPP链路建立流程

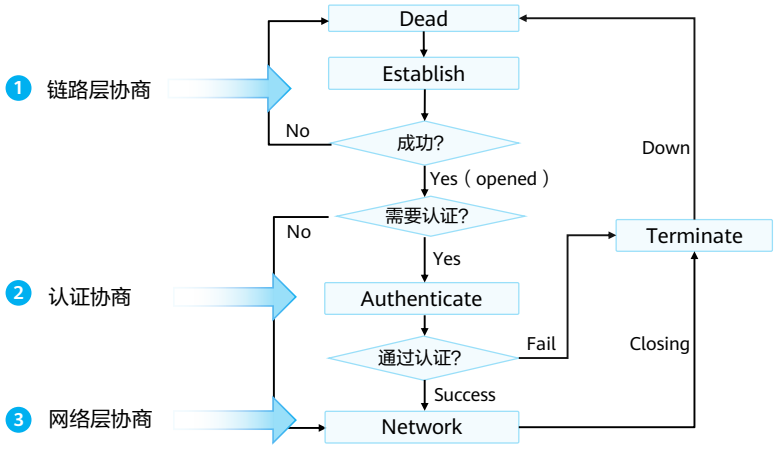
- PPP链路的建立有三个阶段的协商过程，链路层协商、认证协商（可选）和网络层协商。
  - 链路层协商：通过LCP报文进行链路参数协商，建立链路层连接。
  - 认证协商（可选）：通过链路建立阶段协商的认证方式进行链路认证。
  - 网络层协商：通过NCP协商来选择和配置一个网络层协议并进行网络层参数协商。



特性介绍 链路建立 LCP协商 认证协商 NCP协商

PPP链路接口状态机

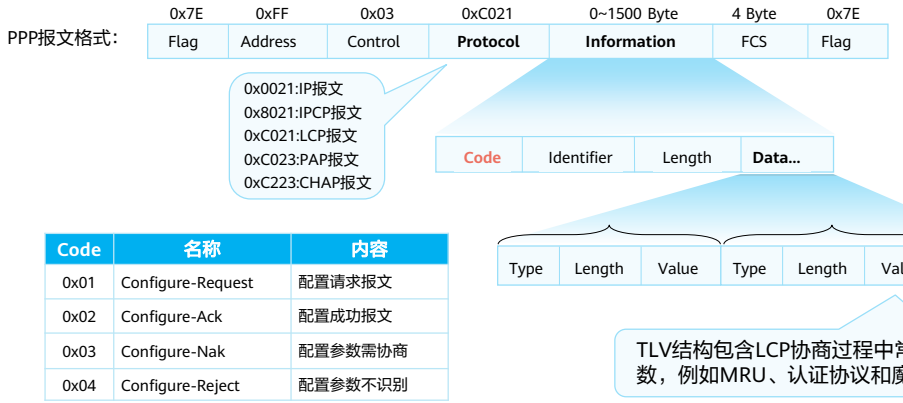
- PPP协商由链路两端的接口完成。接口的状态表示了协议的协商阶段。



- 正常PPP链路建立需要经历链路建立阶段、认证阶段和网络层协商阶段，详细过程如下：
  1. 通信双方开始建立PPP链路时，先进入到Establish阶段。
  2. 在Establish阶段，进行LCP协商：协商通信双方的MRU（Maximum Receive Unit，最大接收单元）、认证方式和魔术字（Magic Number）等选项。协商成功后进入Opened状态，表示底层链路已建立。
  3. 如果配置了认证，将进入Authenticate阶段。否则直接进入Network阶段。
  4. 在Authenticate阶段，会根据连接建立阶段协商的认证方式进行链路认证。认证方式有两种：PAP和CHAP。如果认证成功，进入Network阶段，否则进入Terminate阶段，拆除链路，LCP状态转为Down。
  5. 在Network阶段，PPP链路进行NCP协商。通过NCP协商来选择和配置一个网络层协议并进行网络层参数协商。最常见的NCP协议是IPCP，用来协商IP参数。
  6. 在Terminate阶段，如果所有的资源都被释放，通信双方将回到Dead阶段。
- PPP运行过程中，可以随时中断连接，物理链路断开、认证失败、超时定时器时间到、管理员通过配置关闭连接等动作都可能导致链路进入Terminate阶段。

LCP报文格式

- PPP报文可由Protocol字段标识不同类型的PPP报文。例如，当Protocol字段为0xC021时，代表是LCP报文。此时又由Code字段标识不同类型LCP报文，如下表所示。



Code	名称	内容
0x01	Configure-Request	配置请求报文
0x02	Configure-Ack	配置成功报文
0x03	Configure-Nak	配置参数需协商
0x04	Configure-Reject	配置参数不识别

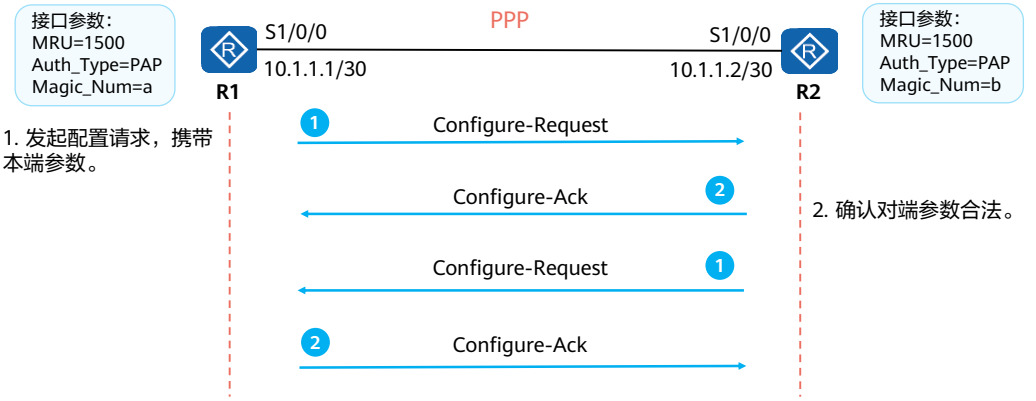
HUAWEI

- PPP帧格式：
  - Flag字段标识一个物理帧的起始和结束，该字节为二进制序列01111110（0X7E）。
  - PPP帧的Address字段字节固定为11111111（0XFF），是一个广播地址。
  - PPP数据帧的Control字段默认为00000011（0X03），表明为无序号帧。
  - 帧校验序列（FCS）字段是个16 bit的校验和，用于检查PPP帧的完整性。
  - Protocol字段用来说明PPP所封装的协议报文类型，0XC021代表LCP报文，0XC023代表PAP报文，0XC223代表CHAP报文。
  - Information字段包含Protocol字段中指定协议的内容，该字段的最大长度被称为最大接收单元MRU，缺省值为1500。
  - 当Protocol字段为0XC021时，Information结构如下：
    - Identifier字段为1个字节，用来匹配请求和响应。
    - Length域的值就是该LCP报文的总字节数据。
    - Data字段则承载各种TLV（Type/Length/Value）参数用于协商配置选项，包括最大接收单元，认证协议等等。
- LCP报文携带的一些常见的配置参数有MRU、认证协议和魔术字。
  - 在VRP（Versatile Routing Platform，通用路由平台）平台上，MRU参数使用接口上配置的MTU（Maximum Transmission Unit，最大传输单元）值来表示。
  - 常用的PPP认证协议有PAP和CHAP，一条PPP链路的两端可以使用不同的认证协议认证对端，但是被认证方必须支持认证方要求使用的认证协议并正确配置用户名和密码等认证信息。
  - LCP使用魔术字来检测链路环路和其他异常情况。魔术字是随机产生的一个数字，随机机制需要保证两端产生相同魔术字的可能性几乎为0。

特性介绍 链路建立 LCP协商 认证协商 NCP协商

LCP协商过程 - 正常协商

- LCP协商由不同的LCP报文交互完成。协商由任意一方发送Configure-Request报文发起。如果对端接收此报文且参数匹配，则通过回复Configure-Ack响应协商成功。

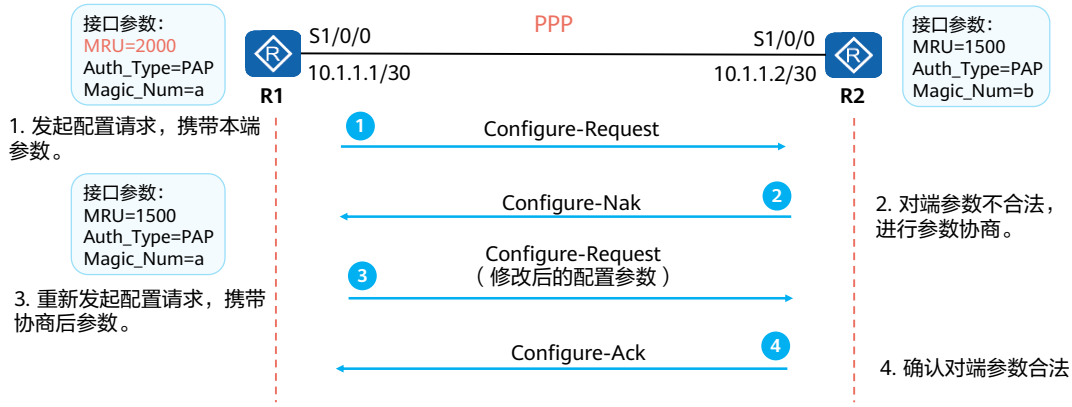


  HUAWEI

- R1和R2使用串行链路相连，运行PPP协议。当物理层链路变为可用状态之后，R1和R2使用LCP协商链路参数。
- 本例中，R1首先发送一个Configure-Request报文，此报文中包含R1上配置的链路层参数。当R2收到此Configure-Request报文之后，如果R2能识别并接受此报文中的所有参数，则向R1回应一个Configure-Ack报文。同样的，R2也需要向R1发送Configure-Request报文，使R1检测R2上的参数是不是可接受的。
- R1在没有收到Configure-Ack报文的情况下，会每隔3秒重传一次Configure-Request报文，如果连续10次发送Configure-Request报文仍然没有收到Configure-Ack报文，则认为对端不可用，停止发送Configure-Request报文。

LCP协商过程 - 参数不匹配

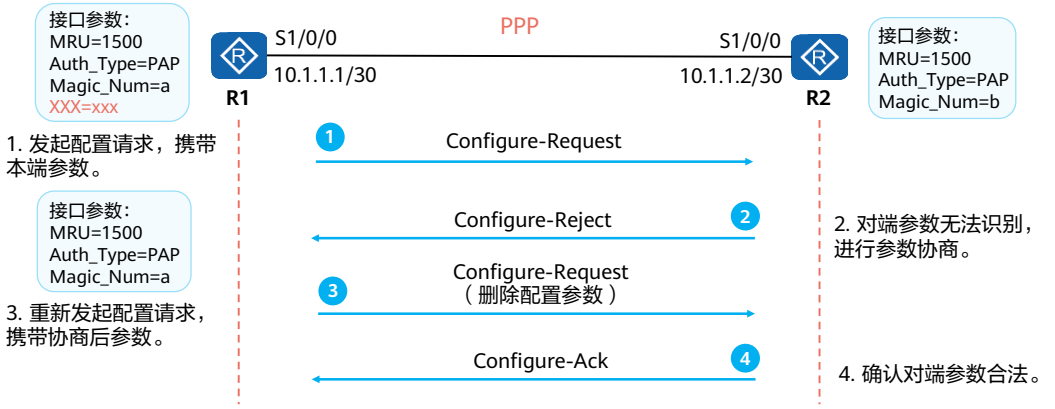
- 在LCP报文交互中出现LCP参数不匹配时，接收方回复Configure-Nak响应告知对端修改参数然后重新协商。



- 当R2收到R1发送的Configure-Request报文之后，如果R2能识别此报文中携带的所有链路层参数，但是认为部分或全部参数的取值不能接受，即参数的取值协商不成功，则R2需要向R1回应一个Configure-Nak报文。
- 在这个Configure-Nak报文中，只包含不能接受的链路层参数，并且此报文所包含的链路层参数将被修改为R2上可以接受的取值（或取值范围）。
- 在收到Configure-Nak报文之后，R1需要根据此报文中的链路层参数重新选择本地配置的其他参数，并重新发送一个Configure-Request。

LCP协商过程 - 参数不识别

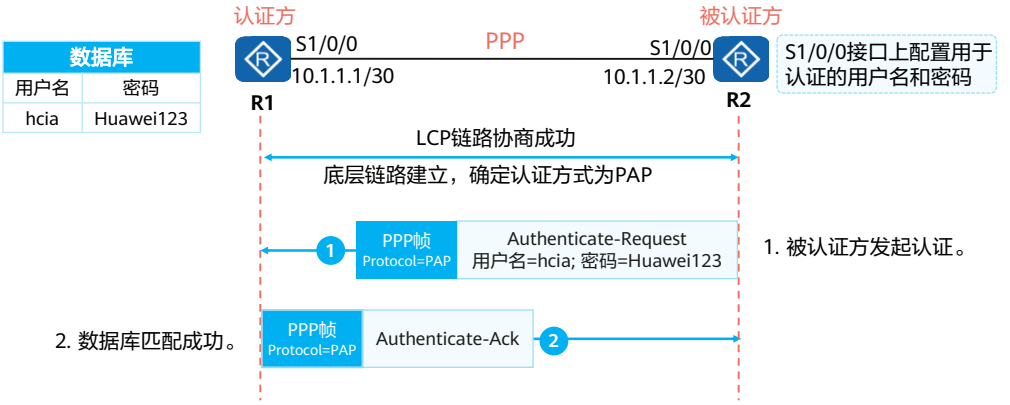
- 在LCP报文交互中出现LCP参数不识别时，接收方回复Configure-Reject响应告知对端删除不识别的参数然后重新协商。



- 当R2收到R1发送的Configure-Request报文之后，如果R2不能识别此报文中携带的部分或全部链路层参数，则R2需要向R1回应一个Configure-Reject报文。在此Configure-Reject报文中，只包含不能被识别的链路层参数。
- 在收到Configure-Reject报文之后，R1需要向R2重新发送一个Configure-Request报文，在新的Configure-Request报文中，不再包含不被对端（R2）识别的参数。

PPP认证模式 - PAP

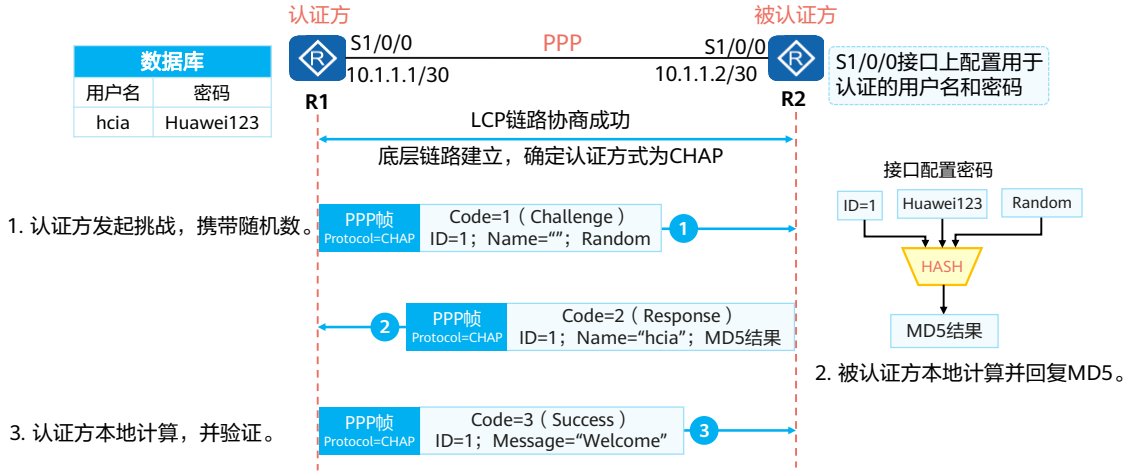
- 链路协商成功后，进行认证协商（此过程可选）。认证协商有两种模式，PAP和CHAP。
- PAP认证双方有两次握手。协商报文以明文的形式在链路上传输。



- LCP协商完成后，认证方要求被认证方使用PAP进行认证。
- PAP认证协议为两次握手认证协议，密码以明文方式在链路上发送，过程如下：
  - 被认证方将配置的用户名和密码信息使用Authenticate-Request报文以明文方式发送给认证方。
  - 认证方收到被认证方发送的用户名和密码信息之后，根据本地配置的用户名和密码数据库检查用户名和密码信息是否匹配；如果匹配，则返回Authenticate-Ack报文，表示认证成功。否则，返回Authenticate-Nak报文，表示认证失败。

PPP认证模式 - CHAP

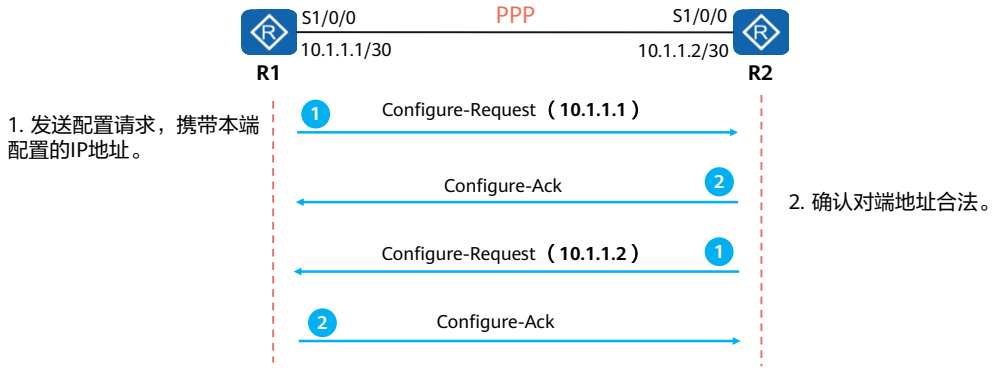
- CHAP认证双方有三次握手。协商报文被加密后再在链路上传输。



- LCP协商完成后，认证方要求被认证方使用CHAP进行认证。
- CHAP认证过程需要三次报文的交互。过程如下：
  - 认证方主动发起认证请求，认证方向被认证方发送Challenge报文，报文内包含随机数（Random）和ID。
  - 被认证方收到此Challenge报文之后，进行一次加密运算，运算公式为MD5{ ID＋随机数＋密码}，意思是将Identifier、随机数和密码三部分连成一个字符串，然后对此字符串做MD5运算，得到一个16 Byte长的摘要信息，然后将此摘要信息和端口上配置的CHAP用户名一起封装在Response报文中发回认证方。
  - 认证方接收到被认证方发送的Response报文之后，按照其中的用户名在本地查找相应的密码信息，得到密码信息之后，进行一次加密运算，运算方式和被认证方的加密运算方式相同；然后将加密运算得到的摘要信息和Response报文中封装的摘要信息做比较，相同则认证成功，不相同则认证失败。
- 使用CHAP认证方式时，被认证方的密码是被加密后才进行传输的，这样就极大的提高了安全性。
- 加密算法声明
  - 使用加密算法时，MD5（数字签名场景和口令加密）加密算法安全性低，存在安全风险，在协议支持的加密算法选择范围内，建议使用更安全的加密算法，例如AES/RSA（2048位以上）/SHA2/HMAC-SHA2。

NCP协商 - 静态IP地址协商

- PPP认证协商后，双方进入NCP协商阶段，协商在数据链路上所传输的数据包的格式与类型。以常见的IPCP协议为例，它分为静态IP地址协商和动态IP地址协商。
- 静态IP地址协商需要手动在链路两端配置IP地址。

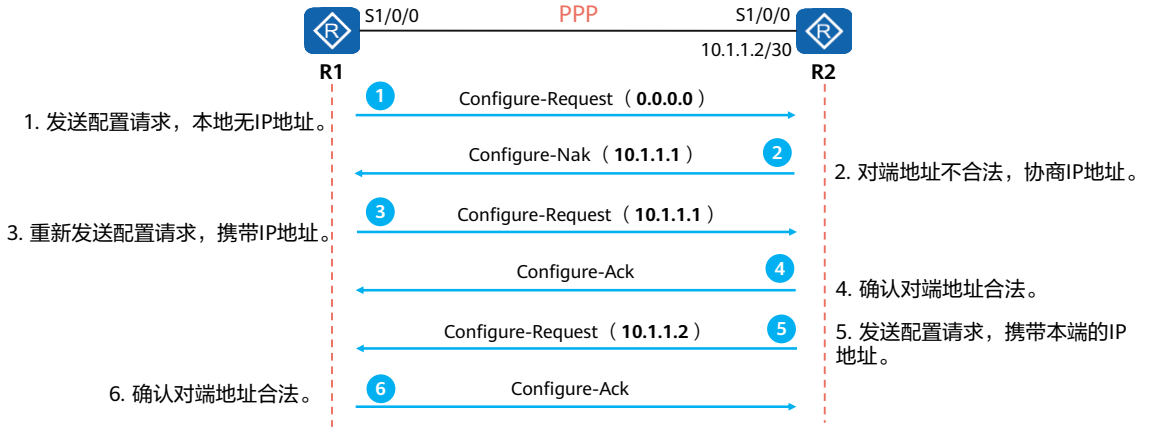


HUAWEI

- NCP主要用来建立和配置不同的网络层协议，协商在该数据链路上所传输的数据包的格式与类型。常见的有IPCP等。
- 静态IP地址协商过程如下：
  - 每一端都要发送Configure-Request报文，在此报文中包含本地配置的IP地址；
  - 每一端接收到此Configure-Request报文之后，检查其中的IP地址，如果IP地址是一个合法的单播IP地址，而且和本地配置的IP地址不同（没有IP冲突），则认为对端可以使用该地址，回应一个Configure-Ack报文。

NCP协商 - 动态IP地址协商

- 动态IP地址协商支持PPP链路一端为对端配置IP地址。



- 动态协商IP地址的过程如下：
  - R1向R2发送一个Configure-Request报文，此报文中会包含一个IP地址0.0.0.0，表示向对端请求IP地址；
  - R2收到上述Configure-Request报文后，认为其中包含的地址（0.0.0.0）不合法，使用Configure-Nak回应一个新的IP地址10.1.1.1；
  - R1收到此Configure-Nak报文之后，更新本地IP地址，并重新发送一个Configure-Request报文，包含新的IP地址10.1.1.1；
  - R2收到Configure-Request报文后，认为其中包含的IP地址为合法地址，回应一个Configure-Ack报文；
  - 同时，R2也要向R1发送Configure-Request报文请求使用地址10.1.1.2，R1认为此地址合法，回应Configure-Ack报文。

目录

PPP基础配置命令

1. 配置接口封装PPP协议

```
[Huawei-Serial0/0/0] link-protocol ppp
```

在接口视图下，将接口封装协议改为ppp，华为串行接口默认封装协议为ppp。

2. 配置协商超时时间间隔

```
[Huawei-Serial0/0/0] ppp timer negotiate seconds
```

在PPP LCP协商过程中，本端设备会向对端设备发送LCP协商报文，如果在指定协商时间间隔内没有收到对端的应答报文，则重新发送。

PAP认证配置命令

1. 配置验证方以PAP方式认证对端

```
[Huawei-aaa] local-user user-name  password { cipher | irreversible-cipher } password
[Huawei-aaa] local-user user-name service-type ppp
```

```
[Huawei-Serial0/0/0] ppp authentication-mode pap
```

配置验证方以PAP方式认证对端，首先需要通过AAA将被验证方的用户名和密码加入本地用户列表，然后选择认证模式。

2. 配置被验证方以PAP方式被对端认证

```
[Huawei-Serial0/0/0] ppp pap local-user user-name password { cipher | simple } password
```

配置本地被对端以PAP方式验证时，本地发送PAP用户名和口令。

CHAP认证配置命令

1. 配置验证方以CHAP方式认证对端

```
[Huawei-aaa] local-user user-name  password { cipher | irreversible-cipher } password
[Huawei-aaa] local-user user-name service-type ppp
```

```
[Huawei-Serial0/0/0] ppp authentication-mode chap
```

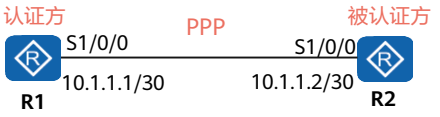
2. 配置被验证方以CHAP方式被对端认证

```
[Huawei-Serial0/0/0] ppp chap user user-name
```

```
[Huawei-Serial0/0/0] ppp chap password { cipher | simple } password
```

配置本地用户名，配置本地被对端以CHAP方式验证时的口令。

配置举例 - PAP认证



- 实验要求：
  1. 在R1与R2之间的PPP链路上启用PAP认证功能；
  2. 将R1配置为认证方；
  3. 将R2配置为被认证方。

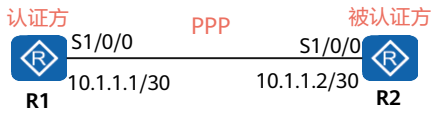
R1的配置如下：

```
[R1]aaa  #添加待认证用户信息
[R1-aaa]local-user huawei password cipher huawei123
[R1-aaa]local-user huawei service-type ppp
#指定认证用户业务类型
[R1]interface Serial 1/0/0
[R1-Serial1/0/0]link-protocol ppp
[R1-Serial1/0/0]ppp authentication-mode pap
#指定认证模式为PAP
[R1-Serial1/0/0]ip address 10.1.1.1 30
```

R2的配置如下：

```
[R2]interface Serial 1/0/0
[R2-Serial1/0/0]link-protocol ppp
[R2-Serial1/0/0]ppp pap local-user huawei password cipher
huawei123          #添加PPP认证的用户信息
[R2-Serial1/0/0]ip address 10.1.1.2 30
```

配置举例 - CHAP认证



- 实验要求：
  1. 在R1与R2之间的PPP链路上启用CHAP认证功能；
  2. 将R1配置为认证方；
  3. 将R2配置为被认证方。

R1的配置如下：

```
[R1]aaa  #添加待认证用户信息
[R1-aaa]local-user huawei password cipher huawei123
[R1-aaa]local-user huawei service-type ppp
#指定认证用户业务类型
[R1]interface Serial 1/0/0
[R1-Serial1/0/0]link-protocol ppp
[R1-Serial1/0/0]ppp authentication-mode chap
#指定认证模式为CHAP
```

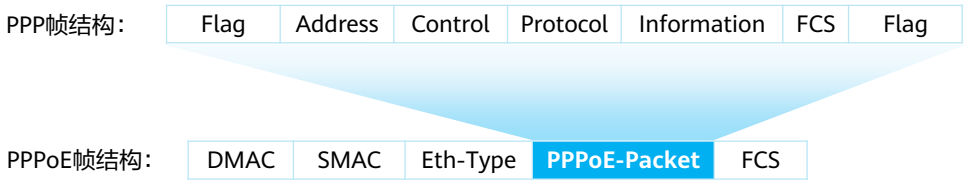
R2的配置如下：

```
[R2]interface Serial 1/0/0
[R2-Serial1/0/0]link-protocol ppp
[R2-Serial1/0/0]ppp chap user huawei
[R2-Serial1/0/0]ppp chap password cipher huawei123
#添加PPP认证的用户信息
```

目录

什么是PPPoE

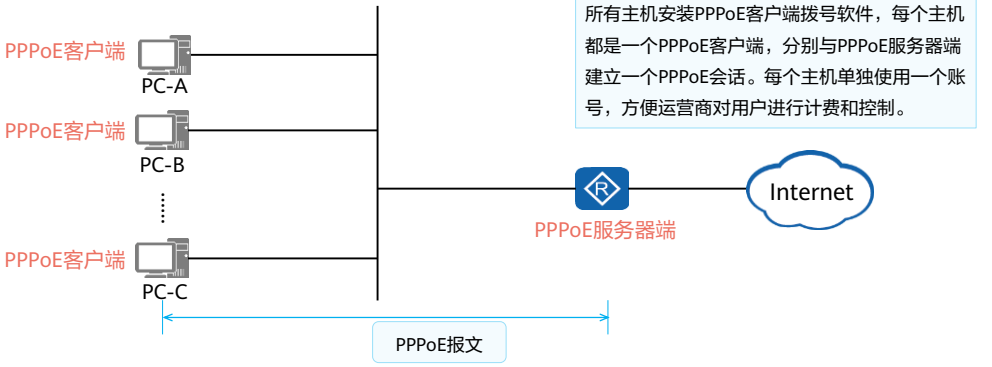
- PPPoE（PPP over Ethernet，以太网承载PPP协议）是一种把PPP帧封装到以太网帧中的链路层协议。PPPoE可以使以太网网络中的多台主机连接到远端的宽带接入服务器。
- PPPoE集中了PPP和Ethernet两个技术的优点。既有以太网的组网灵活优势，又可以利用PPP协议实现认证、计费等功能。



- 运营商希望把一个站点上的多台主机连接到同一台远程接入设备，同时接入设备能够提供与拨号上网类似的访问控制和计费功能。在众多的接入技术中，把多个主机连接到接入设备的比较经济的方法就是以太网，而PPP协议可以提供良好的访问控制和计费功能，于是产生了在以太网上传输PPP报文的技术，即PPPoE。
- PPPoE利用以太网将大量主机组成网络，通过一个远端接入设备接入因特网，并运用PPP协议对接入的每个主机进行控制，具有适用范围广、安全性高、计费方便的特点。

PPPoE应用场景

- PPPoE实现了在以太网上提供点到点的连接。PPPoE客户端与PPPoE服务器端之间建立PPP会话，封装PPP数据报文，为以太网上的主机提供接入服务，实现用户控制和计费，在企业网络与运营商网络中应用广泛。
- PPPoE的常见应用场景有家庭用户拨号上网、企业用户拨号上网等。



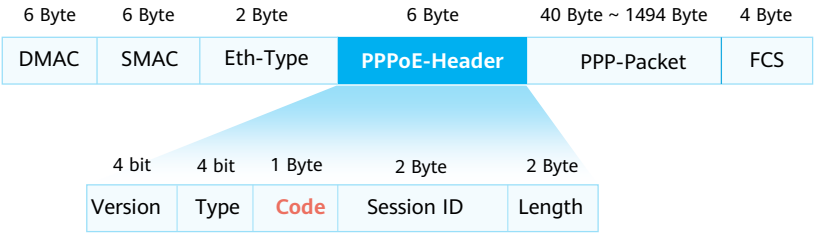
HUAWEI

PPPoE会话建立

- PPPoE的会话建立有三个阶段，PPPoE发现阶段、PPPoE会话阶段和PPPoE终结阶段。

PPPoE报文

- PPPoE会话的建立通过不同的PPPoE报文交互实现。PPPoE报文结构及常见的报文类型如下所示：

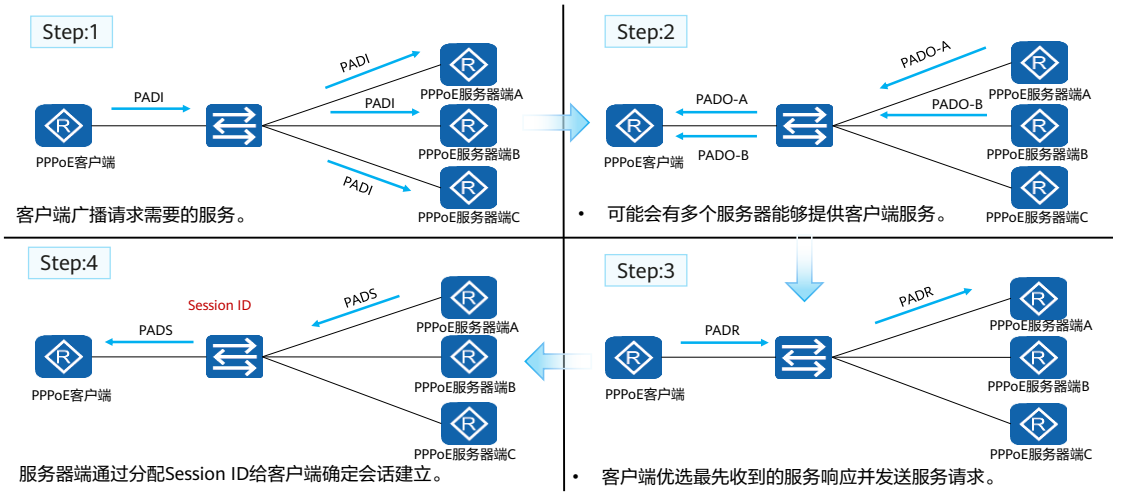


Code	名称	内容
0x09	PADI	PPPoE Active Discovery Initiation，PPPoE激活发现起始报文
0x07	PADO	PPPoE Active Discovery Offer，PPPoE激活发现服务报文
0x19	PADR	PPPoE Active Discovery Request，PPPoE激活发现请求报文
0x65	PADS	PPPoE Active Discovery Session-confirmation，PPPoE激活发现会话确认报文
0xa7	PADT	PPPoE Active Discovery Terminate，PPPoE激活发现终止报文

- PPPoE报文封装在Ethernet帧中，Ethernet中各字段解释如下：
- DMAC：表示目的设备的MAC地址，通常为以太网单播目的地址或者以太网广播地址（0xFFFFFFFF）。
- SMAC：表示源设备的以太网MAC地址。
- Eth-Type：表示协议类型字段，当值为0x8863时表示承载的是PPPoE发现阶段的报文。当值为0x8864时表示承载的是PPPoE会话阶段的报文。
- PPPoE字段中的各个字段解释如下：
  - VER：表示PPPoE版本号，值为0x01。
  - Type：表示类型，值为0x01。
  - Code：表示PPPoE报文类型，不同取值标识不同的PPPoE报文类型。
  - PPPoE会话ID，与以太网SMAC和DMAC一起定义了一个PPPoE会话。
  - Length：表示PPPoE报文的长度。

PPPoE发现阶段

- PPPoE协议发现有四个步骤：客户端发送请求、服务端响应请求、客户端确认响应和建立会话。

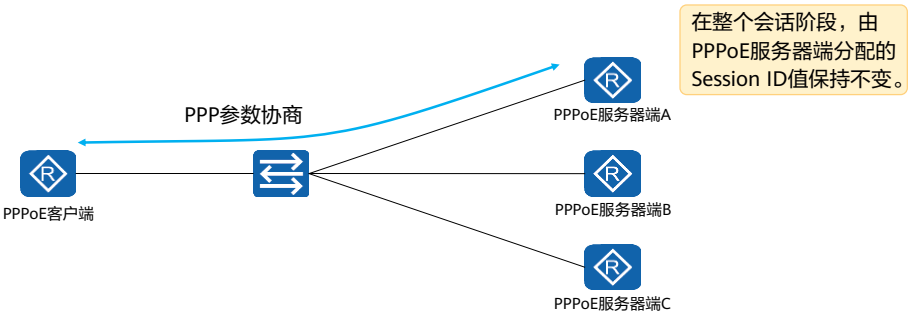


1. PPPoE客户端在本地以太网中广播一个PADI报文，此PADI报文中包含了客户端需要的服务信息。
   - PADI报文的目的MAC地址是一个广播地址，Code字段为0x09，Session ID字段为0x0000。
   - 所有PPPoE服务器端收到PADI报文之后，会将报文中所请求的服务与自己能够提供的服务进行比较。
2. 如果服务器端可以提供客户端请求的服务，就会回复一个PADO报文。
   - PADO报文的目的地址是发送PADI报文的客户端MAC地址，Code字段为0x07，Session ID字段为0x0000。
3. 客户端可能会收到多个PADO报文，此时将选择最先收到的PADO报文对应的PPPoE服务器端，并发送一个PADR报文给这个服务器端。
   - PADR报文的目的地址是选中的服务器端的MAC地址，Code字段为0x19，Session ID字段为0x0000。
4. PPPoE服务器端收到PADR报文后，会生成一个唯一的Session ID来标识和PPPoE客户端的会话，并发送PADS报文。
   - PADS报文的目的地址是PPPoE客户端的MAC地址，Code字段为0x65，Session ID字段是PPPoE服务器端为本PPPoE会话产生的Session ID。

- 会话建立成功后，PPPoE客户端和服务器端进入PPPoE会话阶段。

PPPoE会话阶段

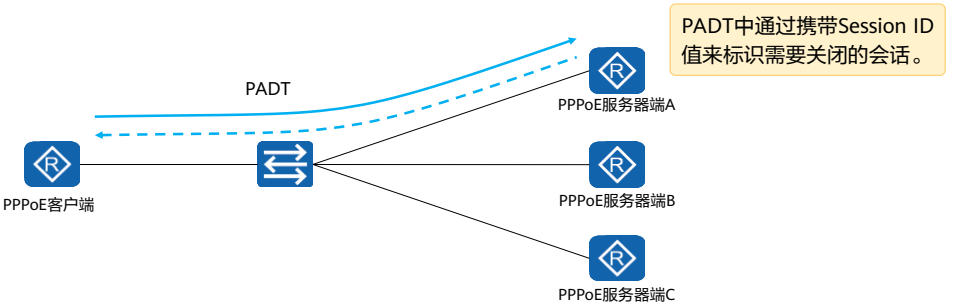
- PPPoE会话阶段会进行PPP协商，分为LCP协商、认证协商、NCP协商三个阶段。



- PPPoE会话阶段可分为两部分：PPP协商阶段和PPP报文传输阶段。
- PPPoE Session上的PPP协商和普通的PPP协商方式一致，分为LCP、认证、NCP三个阶段。
  - LCP阶段主要完成建立、配置和检测数据链路连接。
  - LCP协商成功后，开始进行认证，认证协议类型由LCP协商结果决定。
  - 认证成功后，PPP进入NCP阶段，NCP是一个协议族，用于配置不同的网络层协议，常用的是IP控制协议（IPCP），它负责配置用户的IP地址和DNS服务器地址等。
- PPPoE Session的PPP协商成功后，就可以承载PPP数据报文。在这一阶段传输的数据包中必须包含在发现阶段确定的Session ID并保持不变。

PPPoE会话终结阶段

- 当PPPoE客户端希望关闭连接时，会向PPPoE服务器端发送一个PADT报文，用于关闭连接。
- 同样，如果PPPoE服务器端希望关闭连接时，也会向PPPoE客户端发送一个PADT报文。



- 在PADT报文中，目的MAC地址为单播地址，Session ID为希望关闭的连接的Session ID。一旦收到一个PADT报文之后，连接随即关闭。

目录

PPPoE基础配置

1. 通过拨号规则来配置发起PPPoE会话的条件

   [Huawei] **dialer-rule**

2. 配置拨号接口用户名，此用户名必须与对端服务器用户名相同

   [Huawei-Dialer1]**dialer user** *username*

3. 将接口置于一个拨号访问组

   [Huawei-Dialer1]**dialer-group** *group-number*

4. 指定当前拨号接口使用的拨号绑定

   [Huawei-Dialer1]**dialer-bundle** *number*

5. 将物理端口与dialer-bundle进行绑定

   [Huawei-Ethernet0/0/0]**pppoe-client dial-bundle-number** *number*

配置实例 - PPPoE客户端

- 实验要求：
  1. 将R1设置为PPPoE客户端，R2为PPPoE服务器端；
  2. 在R1上配置PPPoE客户端拨号接口；
  3. 在R1上配置PPPoE客户端拨号接口的认证功能；
  4. R1上的拨号接口获取PPPoE服务器端分配的IP地址；
  5. R1通过拨号接口可以访问服务器端。

1.创建拨号接口并配置被认证方用户名和密码：

```
[R1]dialer-rule
[R1-dialer-rule]dialer-rule 1 ip permit
[R1-dialer-rule]quit
[R1]interface dialer 1
[R1-Dialer1] dialer user enterprise
[R1-Dialer1] dialer-group 1
[R1-Dialer1] dialer bundle 1
[R1-Dialer1] ppp chap user huawei1
[R1-Dialer1] ppp chap password cipher huawei123
[R1-Dialer1] ip address ppp-negotiate
```

2.将拨号接口绑定出接口：

```
[R1]interface GigabitEthernet 0/0/1
[R1-GigabitEthernet0/0/1]pppoe-client dial-bundle-number 1
[R1-GigabitEthernet0/0/1]quit
```

3.配置本端到达服务器端的缺省路由：

```
[R1]ip route-static 0.0.0.0  0.0.0.0  dialer 1
```


- PPPoE客户端配置包括三个步骤。
- 第一步配置一个拨号接口。
  - **dialer-rule**命令用于进入Dialer-rule视图，在该视图下，可以通过拨号规则来配置发起PPPoE会话的条件。
  - **interface dialer** *number*命令用来创建并进入Dialer接口。
  - **dialer user** *user-name*命令用于配置对端用户名。
  - **dialer-group** *group-number*命令用来将接口置于一个拨号访问组。
  - **dialer bundle** *number*命令用来指定Dialer接口使用的Dialer bundle。设备通过Dialer bundle将物理接口与拨号接口关联起来。
  - 注：必须确保命令**dialer-group**中的参数*group-number*和命令**dialer-rule**中的*dialer-rule-number*保持一致。
- 第二步是在接口上将Dialer Bundle和接口绑定：
  - **pppoe-client dial-bundle-number** *number*命令来实现Dialer Bundle和物理接口的绑定，用来指定PPPoE会话对应的Dialer Bundle，其中number是与PPPoE会话相对应的Dialer Bundle编号。
- 第三步配置一条缺省静态路由，该路由允许在路由表中没有相应匹配表项的流量都能通过拨号接口发起PPPoE会话。

配置实例 - PPPoE服务器端

- 实验要求：
  1. 在PPPoE服务器端上创建为客户端分配IP的地址池；
  2. PPPoE服务器端完成PPPoE客户端认证并分配合法的IP地址。

1.创建地址池与虚拟模板：

```
[R2]ip pool pool1      #创建地址池，指定分配的IP地址和网关
[R2-ip-pool-pool1]network 192.168.1.0 mask 255.255.255.0
[R2-ip-pool-pool1]gateway-list 192.168.1.254
[R2]interface Virtual-Template 1          #创建虚拟模板接口
[R2-Virtual-Template1]ppp authentication-mode chap
[R2-Virtual-Template1]ip address 192.168.1.254 255.255.255.0
[R2-Virtual-Template1]remote address pool pool1
```

2.将物理接口与虚拟模板绑定：

```
[R2]interface GigabitEthernet 0/0/0
[R2-GigabitEthernet0/0/0]pppoe-server bind virtual-template 1
[R2-GigabitEthernet0/0/0]quit
```

3.创建访问用户：

```
[R2]aaa     #添加认证用户信息
[R2-aaa]local-user huawei1 password cipher huawei123
[R2-aaa]local-user huawei1 service-type ppp
```


- PPPoE服务器端配置
  - **interface virtual-template**命令用来创建虚拟模板接口，或者进入一个已经创建的虚拟模板接口视图。
  - **pppoe-server bind**命令用来配置PPPoE接入用户上线绑定的虚拟模板接口。

配置验证

1、查看拨号接口详细信息

```
<R1>display interface Dialer 1
Dialer1 current state: UP
Line protocol current state: UP (spoofing)
Description: HUAWEI, AR Series, Dialer1 Interface
Route Port, The Maximum Transmit Unit is 1500, Hold timer is 10(sec)
Internet Address is negotiated, 192.168.10.254/32
Link layer protocol is PPP
LCP initial
Physical is Dialer
Bound to Dialer1:0:
Dialer1:0 current state : UP
Line protocol current state : UP
Link layer protocol is PPP
LCP opened, IPCP opened
```

2、查看PPPoE-client会话初始状态信息

```
[R1]display pppoe-client session summary
PPPoE Client Session:
ID   Bundle  Dialer  Intf      Client-MAC      Server-MAC      State
0    1       1       GE0/0/1   54899876830c    000000000000    IDLE
```

3、查看PPPoE-client会话建立状态信息

```
[R1]display pppoe-client session summary
PPPoE Client Session:
ID   Bundle  Dialer  Intf      Client-MAC      Server-MAC      State
1    1       1       GE0/0/1   00e0fc0308f6    00e0fc036781    UP
```


- **display interface dialer**[**number**]命令用于查看拨号接口的配置，便于定位拨号接口的故障。
- LCP opened, IPCP opened表示链路的状态完全正常。
- **display pppoe-client session summary**命令用于查看PPPoE客户端的PPPoE会话状态和统计信息。
    - ID表示PPPoE会话ID，Bundle ID和Dialer ID的值与拨号参数配置有关。
    - Intf表示客户端侧协商时的物理接口。
    - State表示PPPoE会话的状态，包括以下四种：
        1. IDLE表示当前会话状态为空闲。
        2. PADI表示PPPoE会话处于发现阶段，并已经发送PADI报文。
        3. PADR表示PPPoE会话处于发现阶段，并已经发送PADR报文。
        4. UP表示PPPoE会话建立成功。

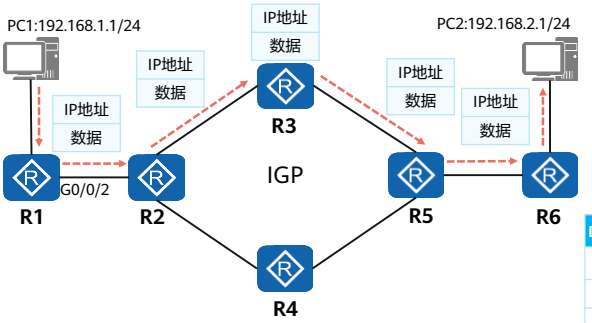
目录

广域网技术的历史演进

- 早期广域网常用的数据链路层协议包括PPP、HDLC和ATM等。后期随着全网IP化的演进，基于IP技术的Internet快速普及，但基于最长匹配算法的IP技术必须使用软件查找路由，转发性能低下，因此IP技术的转发性能成为当时限制网络发展的瓶颈。
- MPLS（Multiprotocol Label Switching，多协议标记交换）最初是为了提高路由器的转发速度而提出的。与传统IP路由方式相比，它在数据转发时，只在网络边缘解析IP报文头，后续节点只基于标签转发，而不用在每一跳都解析IP报文头，减少软件处理流程节约了处理时间。
- 随着路由器性能的提升，路由查找速度已经不是阻碍网络发展的瓶颈。这使得MPLS在提高转发速度方面不再具备明显的优势。但是MPLS支持多层标签和转发平面面向连接的特性，使其在VPN（Virtual Private Network，虚拟专用网）、TE（Traffic Engineering，流量工程）、QoS（Quality of Service，服务质量）等方面得到广泛应用。

传统IP路由转发

- 传统的IP转发采用的是逐跳转发。数据报文经过每一台路由器，都要被解封装查看报文网络层信息，然后根据路由最长匹配原则查找路由表指导报文转发。各路由器重复进行解封装查找路由表和再封装的过程，所以转发性能低。



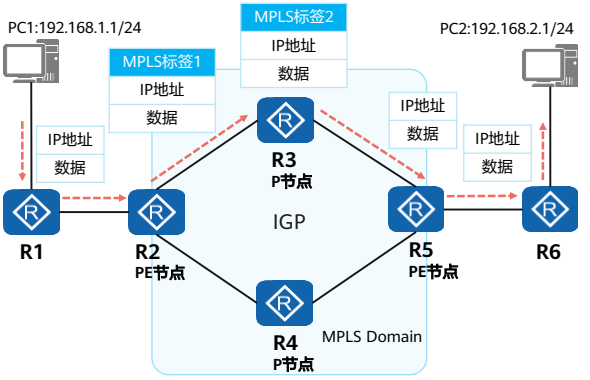
- 传统IP路由转发的特点：
  - 所有路由器需要知道全网的路由。
  - 传统IP转发是面向无连接的，无法提供较好的端到端QoS保证。

R1路由表

Destination/Mask	Protocol	Preference	Cost	NextHop	Interface
192.168.1.0/24	Direct	0	0	192.168.1.254	GE0/0/0
192.168.12.0/24	Direct	0	0	192.168.12.1	GE0/0/2
192.168.2.0/24	OSPF	10	3	192.168.12.2	GE0/0/2

HUAWEI

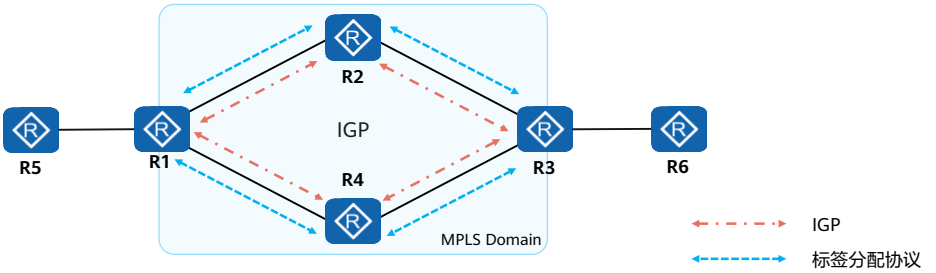
MPLS标签转发



- MPLS是一种IP骨干网技术。
- MPLS是一种隧道技术，在IP路由和控制协议的基础上，向网络层提供面向连接的交换。能够提供较好的QoS保证。
- MPLS标签指导报文转发的过程中，使用本地标签查找替代传统IP转发的路由查找，大大提高转发效率。
- MPLS转发过程中使用的标签，既可以通过手工静态配置，又可以通过动态标签分发协议分配。

MPLS转发存在的问题

- MPLS的标签分发有静态和动态两种方式，均面临着不同的问题：
  - 静态标签分发为手工配置。随着网络规模不断的扩大，网络拓扑易变化，静态手工配置标签不适应大型网络需求。
  - 动态标签分发的问题，一方面在于部分动态标签协议本身并无算路能力，需依赖IGP进行路径计算，同时控制面协议复杂，设备之间需要发送大量的消息来维持邻居及路径状态，浪费了链路带宽及设备资源。另一方面部分标签分发协议虽然支持流量工程，但是配置复杂，不支持负载分担，需要大量协议报文维护路径正常工作；同时每台设备都是独立存在，只知道自己的状态，设备之间需要交互信令报文，也会浪费链路带宽及设备资源。

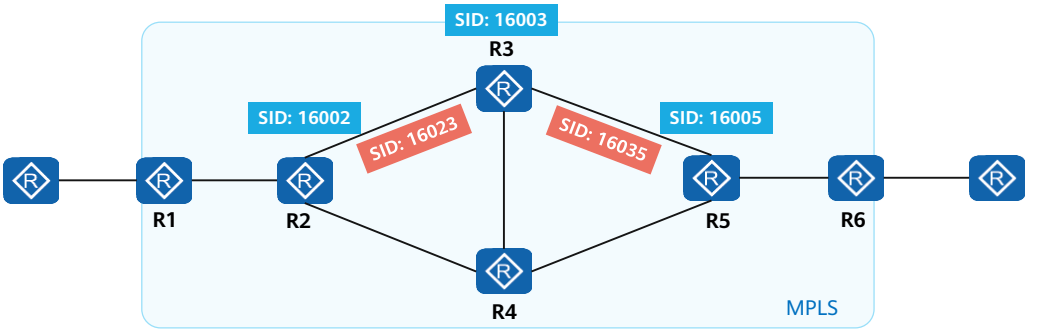


Segment Routing简介

- 为解决传统IP转发和MPLS转发的问题，业界提出了SR （ Segment Routing，分段路由 ）。SR的转发机制有很大改进，主要体现在以下几个方面：
  1. 基于现有协议进行扩展：
     - 扩展后的IGP/BGP具有标签分发能力，因此网络中无需其他任何标签分发协议，实现协议简化。
  2. 引入源路由机制：
     - 基于源路由机制，支持通过控制器进行集中算路。
  3. 由业务来定义网络：
     - 业务驱动网络，由应用提出需求（时延、带宽、丢包率等），控制器收集网络拓扑、带宽利用率、时延等信息，根据业务需求计算显式路径。

Segment Routing转发原理（1）

- SR将网络路径分成一个个的段（Segment），并且为这些段分配SID（Segment ID）。
- SID的分配对象有两种，转发节点或者邻接链路。本例中转发节点SID 1600X，X为路由器编号；邻接链路SID 160XX，XX表示链路两端的节点编号。

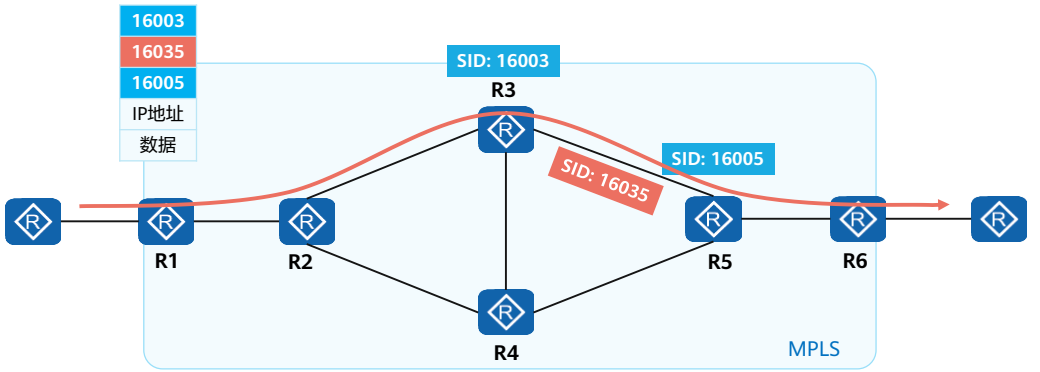


HUAWEI

- SID用于标识Segment，它的格式取决于具体的技术实现，例如可以使用MPLS标签、MPLS标签空间中的索引、IPv6报文头部。例如使用MPLS标签被称为SR-MPLS，使用IPv6被称为SRv6。

Segment Routing转发原理 (2)

- 邻接链路和网络节点的SID有序排列形成段序列（Segment List），它代表一条转发路径。SR由源节点将段序列编码在数据包头部，随数据包传输。SR的本质是指令，指引报文去哪里和怎么去。

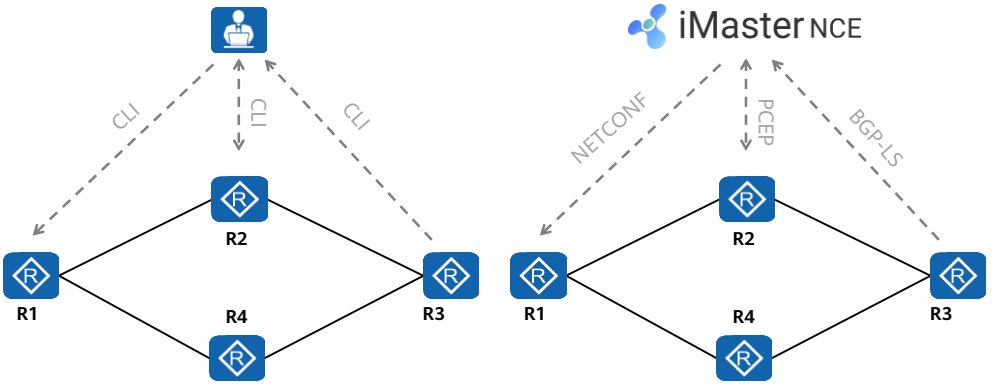


HUAWEI

- 接收端收到数据包后，对段序列进行解析，如果段序列的顶部段标识是本节点时，则弹出该标识，然后进行下一步处理；如果不是本节点，则使用ECMP（Equal Cost Multiple Path）方式将数据包转发到下一节点。

SR的部署方式

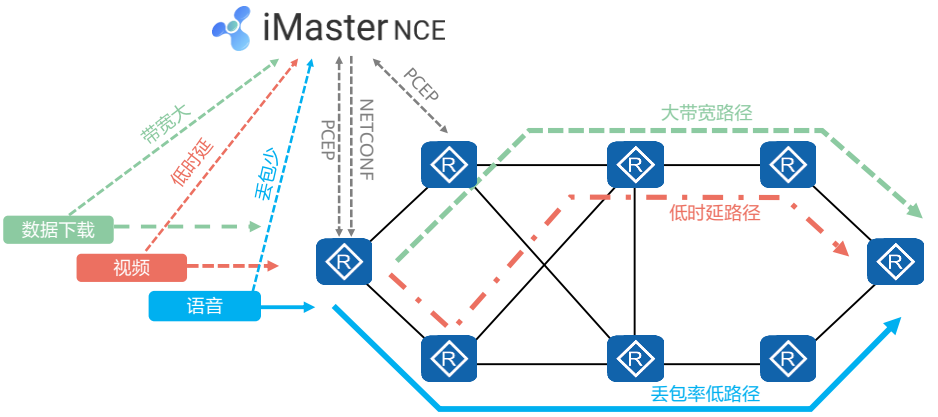
- SR部署分为有控制器部署和无控制器部署。控制器配合方式由控制器收集信息，预留路径资源和计算路径，最后将结果下发到头结点，是更为推荐的部署方式。



- PCEP：Path Computation Element Communication Protocol，路径计算单元通信协议
- NETCONF：Network Configuration Protocol，网络配置协议

Segment Routing的应用

- SR可以简易的指定的报文转发路径，在现网中可以为不同业务定义不同的路径。例如本例定义了数据下载、视频和语音三条显式路径，实现了业务驱动网络。设备由控制器纳管，支持路径实时快速发放。



HUAWEI

思考题

1. （多选）下面关于PPP描述正确的是（　）。

   A. PPP支持将多条物理链路捆绑为逻辑链路以增大带宽。

   B. PPP支持明文和密文认证。

   C. PPP扩展性不好，不可以部署在以太网链路上。

   D. 对物理层而言， PPP支持异步链路和同步链路。

   E. PPP支持多种网络层协议，如IPCP等。

2. （单选）PPPoE客户端向服务器端发送PADI报文，服务器端回复PADO报文。其中，PADO报文是一个什么帧？（　）

   A. 组播　　B. 广播　　C. 单播　　D. 任播

3. （单选）以太网数据帧的Length/Type字段取以下哪个值时，表示承载的是PPPoE发现阶段的报文？（　）

   A. 0x0800　　B. 0x8864　　C. 0x8863　　D. 0x0806

1. ABDE
2. C
3. C

本章总结

- 通过回顾早期广域网技术的类型和应用，介绍了广域网发展演进的历程，从开始的电路交换网络到后期IP化网络，再到MPLS标签交换网，最后引出SR网络，随着网络技术的不断的发展，网络也变得越来高效智能。
- 介绍PPP协议的工作原理，包括PPP链路建立的参数协商，认证协商以及网络层协商的过程。重点分析了PPP的两个认证协议PAP和CHAP，描述了它们的工作过程以及不同之处。
- PPP协议在当前最主要的应用是PPPoE，通过分析PPPoE会话的发现、协商、建立及拆除的过程，全面了解PPPoE的工作机制及配置。

更多信息

- SRv6技术与产业白皮书
  - https://e.huawei.com/cn/material/networking/ne-router/c1e6ffbba36147a1aab69f16a7cf0499
- （多媒体）Segment Routing IPv6进阶系列-01 产生背景
  - https://support.huawei.com/enterprise/zh/doc/EDOC1100086272?idPath=24030814%7C9856750%7C22715517%7C9858933%7C21134118
- （多媒体）Segment Routing IPv6进阶系列-02 基本原理
  - https://support.huawei.com/enterprise/zh/doc/EDOC1100086273?idPath=24030814%7C9856750%7C22715517%7C9858933%7C21134118


谢 谢
www.huawei.com

网络管理与运维

前言

- 随着网络的规模越来越庞大，网络中的设备种类繁多，如何对越来越复杂的网络进行有效的管理，从而提供高质量的网络服务，已成为网络管理所面临的巨大挑战。
- 网络的管理和运维手段多样，本章将对几种常见的网管与运维手段展开介绍。

- 通常意义上，网络管理与运维统称为网管。

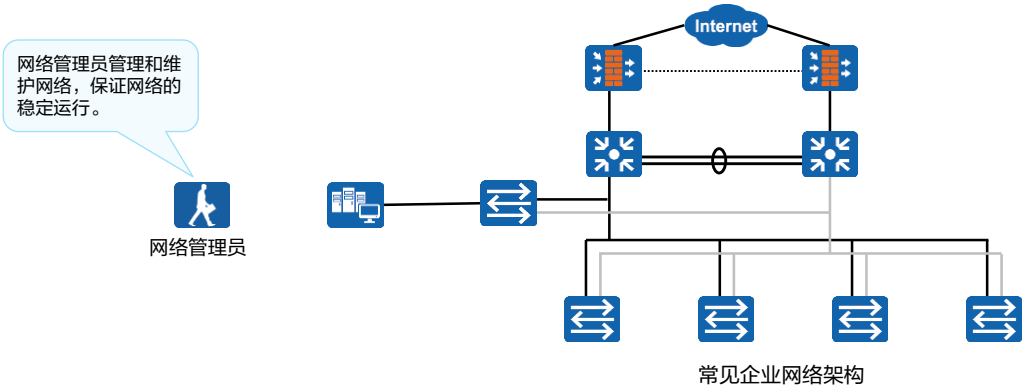
目标

- 学完本课程后，您将能够：
  - 了解网管的基本概念
  - 掌握常见的网管方式
  - 描述网管的基本功能
  - 掌握SNMP协议的工作原理
  - 了解华为iMaster NCE及相关技术

目录

什么是网络管理?

- 网络管理是通过对网络中设备的管理，保证设备工作正常，使通信网络正常地运行，以提供高效、可靠和安全的通信服务，是通信网络生命周期中的重要一环。



常见企业网络架构

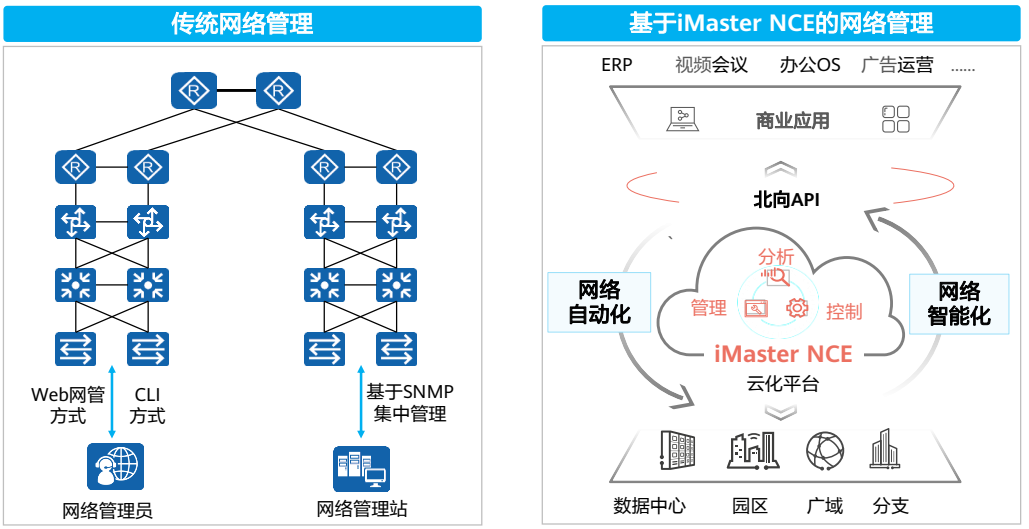
- 网络管理（Network Management）分为两类：
  - 第一类是对网络应用程序、用户账号（例如文件的使用）和存取权限（许可）的管理。它们都是与软件有关的网络管理问题，这里不作深入解释。
  - 第二类是对构成网络的硬件即网元的管理，包括防火墙、交换机、路由器等等。本课程主要针对此类网络管理。
- 一般企业网络中会有专门的部门或者人员负责网络的管理与运维。
- 注：
  - NE（Network Element，网元）：即网络单元，包含硬件设备及运行其上的软件。通常一个网络单元至少具有一块主控板，负责整个网络单元的管理和监控。主机软件运行在主控板上。
  - 通常网络运维的操作都属于网络管理的范畴，本章后续所指的网络管理指对网络进行管理和维护。

网络管理基本功能

OSI定义了网络管理的五大功能模型：

- 配置管理（Configuration Management）：配置管理负责监控网络的配置信息，使网络管理人员可以生成、查询和修改硬件、软件的运行参数和条件，并可以进行相关业务的配置。
- 性能管理（Performance Management）：性能管理以网络性能为准则，保证在使用较少网络资源和具有较小时延的前提下，网络能够提供可靠、连续的通信能力。
- 故障管理（Fault Management）：故障管理的主要目标是确保网络始终可用，并在发生故障时尽快将其修复。
- 安全管理（Security Management）：安全管理可以保护网络和系统免受未经授权的访问和安全攻击。
- 计费管理（Accounting Management）：记录用户使用网络资源的情况并核收费用，同时也统计网络的利用率。

网络管理方式

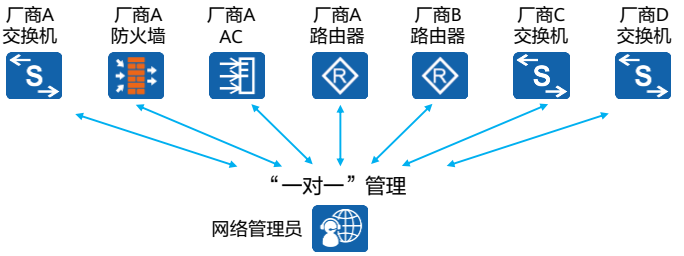


- 传统网络管理：
  - Web网管方式：利用设备内置的Web服务器，为用户提供图形化的操作界面。用户需要从终端通过HTTPS（Hypertext Transfer Protocol Secure，HTTPS 加密协定）登录到设备进行管理。
  - CLI方式：用户利用设备提供的命令行，通过Console口、Telnet或SSH等方式登录到设备，对设备进行管理与维护。此方式可以实现对设备的精细化管理，但是要求用户熟悉命令行。
  - 基于SNMP集中管理：SNMP（Simple Network Management Protocol，简单网络管理协议）提供了一种通过运行网络管理软件的中心计算机（即网络管理站）来管理网元（如路由器、交换机）的方法。此方式可以实现对全网设备集中式、统一化管理，大大提升了管理效率。
- 基于iMaster NCE的网络管理：
  - iMaster NCE是集管理、控制、分析和AI智能功能于一体的网络自动化与智能化平台，包括四大关键能力：全生命周期自动化、基于大数据和AI的智能闭环、开放可编程使能场景化APP生态、超大容量全云化平台。
  - iMaster NCE采用NETCONF（Network Configuration Protocol，网络配置协议）、RESTCONF等协议对设备下发配置，使用Telemetry监控网络流量。

目录

通过CLI或Web进行管理

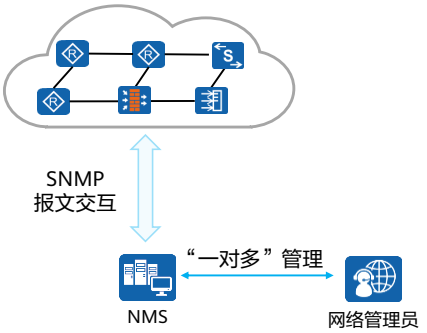
- 当网络规模较小时，CLI和Web方式是常见的网络管理方式。
  - 网络管理员可以通过HTTPS、Telnet、Console等方式登录设备后，对设备逐一进行管理。
  - 这种管理方式不需要在网络中安装任何程序或部署服务器，成本较低。
  - 网络管理员自身需要熟练掌握网络理论知识、各设备厂商网络配置命令。
  - 当网络规模较大，网络拓扑较为复杂时，这种方式的局限性较大。



- 随着网络技术的飞速发展，在网络不断普及地同时也给网络管理带来了一些问题：
  - 网络设备数量成几何级增长，使得网络管理员对设备的管理变得越来越困难；同时，网络作为一个复杂的分布式系统，其覆盖地域不断扩大，也使得对这些设备进行实时监控和故障排查变得极为困难。
  - 网络设备种类多种多样，不同设备厂商提供的管理接口（如命令行接口）各不相同，这使得网络管理变得愈发复杂。

基于SNMP的集中式管理

- SNMP（Simple Network Management Protocol，简单网络管理协议）是广泛用于TCP/IP网络的网络管理标准协议，提供了一种通过运行网络管理软件的中心计算机，即NMS（Network Management Station，网络管理工作站）来管理网元的方法。

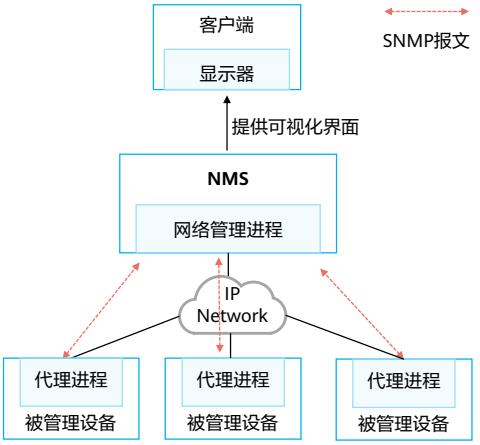


- 网络管理员可以利用NMS在网络上的任意节点完成信息查询、信息修改和故障排查等工作，提升工作效率。
- 屏蔽了不同产品之间的差异，实现了不同种类和厂商的网络设备之间的统一管理。

HUAWEI

- SNMP共有三个版本：SNMPv1、SNMPv2c和SNMPv3。
  - 1990年5月，RFC 1157定义了SNMP的第一个版本SNMPv1。RFC 1157提供了一种监控和管理计算机网络的系统方法。SNMPv1基于团体名认证，安全性较差，且返回报文的错误码也较少。
  - 1996年，IETF颁布了RFC 1901，定义了SNMP的第二个版本SNMPv2c。SNMPv2c中引入了GetBulk和Inform操作，支持更多的标准错误码信息，支持更多的数据类型（Counter64、Counter32）。
  - 鉴于SNMPv2c在安全性方面没有得到改善，IETF又颁布了SNMPv3的版本，提供了基于USM（User-Based Security Model，用户安全模块）的认证加密和VACM（View-based Access Control Model，基于视图的访问控制模型）功能。

SNMP典型架构

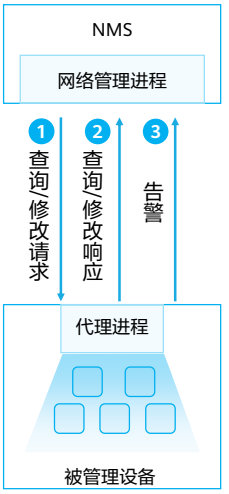


- 在基于SNMP进行管理的网络中，NMS是整个网络的网管中心，在它之上运行管理进程。每个被管理设备需要运行代理（Agent）进程。管理进程和代理进程利用SNMP报文进行通信。
- NMS是一个采用SNMP协议对网络设备进行管理/监控的系统，运行在NMS服务器上。
- 被管理设备是网络中接受NMS管理的设备。
- 代理进程运行于被管理设备上，用于维护被管理设备的信息数据并响应来自NMS的请求，把管理数据汇报给发送请求的NMS。

HUAWEI

- NMS通常是一个独立的设备，运行网络管理应用程序。网络管理应用程序至少能够提供一个人机交互界面，网络管理员通过人机交互界面完成绝大多数网络管理工作。比较常见的人机交互方式为通过Web页面进行交互，即网络管理员通过带显示器的终端，通过HTTP/HTTPS访问NMS提供的Web页面。

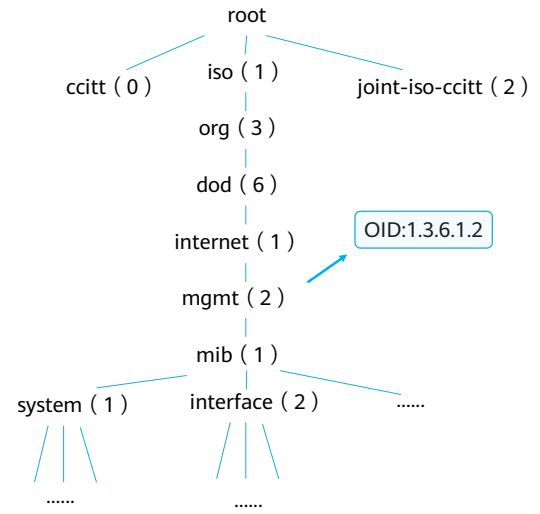
SNMP的信息交互



被管理对象

- NMS和被管理设备的信息交互分为两种：
  - NMS通过SNMP协议给被管理设备发送修改配置信息请求或查询配置信息请求。被管理设备上运行的代理进程根据NMS的请求消息做出响应。
  - 被管理设备可以主动向NMS上报告警信息（Trap）以便网络管理员及时发现故障。
- 被管理对象（Managed object）：每一个设备可能包含多个被管理对象，被管理对象可以是设备中的某个硬件，也可以是在硬件、软件（如路由选择协议）上配置的参数集合。
- SNMP规定通过MIB（Management Information Base，管理信息库）去描述可管理实体的一组对象。

MIB



- MIB是一个数据库，指明了被管理设备所维护的变量（即能够被代理进程查询和设置的信息）。MIB在数据库中定义了被管理设备的一系列属性：
  - 对象标识符（Object IDentifier，OID）
  - 对象的状态
  - 对象的访问权限
  - 对象的数据类型等
- MIB给出了一个数据结构，包含了网络中所有可能的被管理对象的集合。因为数据结构与树相似，MIB又被称为对象命名树。

- MIB的定义与具体的网络管理协议无关。设备制造商可以在产品（如路由器）中包含SNMP代理软件，并保证在定义新的MIB项目后该软件仍遵守标准。用户可以使用同一网络管理客户软件来管理具有不同版本MIB的多个路由器。若一台路由器上不支持此MIB，那么就无法提供相应的功能。
- MIB可以分为公有MIB和私有MIB两种。
  - 公有MIB：一般由RFC定义，主要用来对各种公有协议进行结构化设计和接口标准化处理。大多数的设备制造商都需要按照RFC的定义来提供SNMP接口。
  - 私有MIB：是公有MIB的必要补充，当公司自行开发私有协议或者特有功能时，可以利用私有MIB来完善SNMP接口的管理功能，同时对第三方网管软件管理存在私有协议或特有功能的设备提供支持。如华为公司企业节点为：1.3.6.1.4.1.2011。

常见MIB节点

- 用于查询或修改的节点：

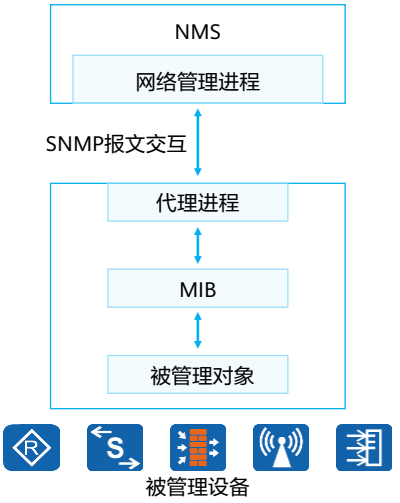
OID	节点名称	数据类型	最大访问权限	含义
1.3.6.1.2.1.2.1	ifNumber	Integer	read-only	系统中网络接口的数量（不关注接口当前状态）。
1.3.6.1.4.1.2011.5.25.41.1.2.1.1.3	hwIpAdEntNetMask	IpAddress	read-create	IP地址的子网掩码。

- 用于告警通知的节点：

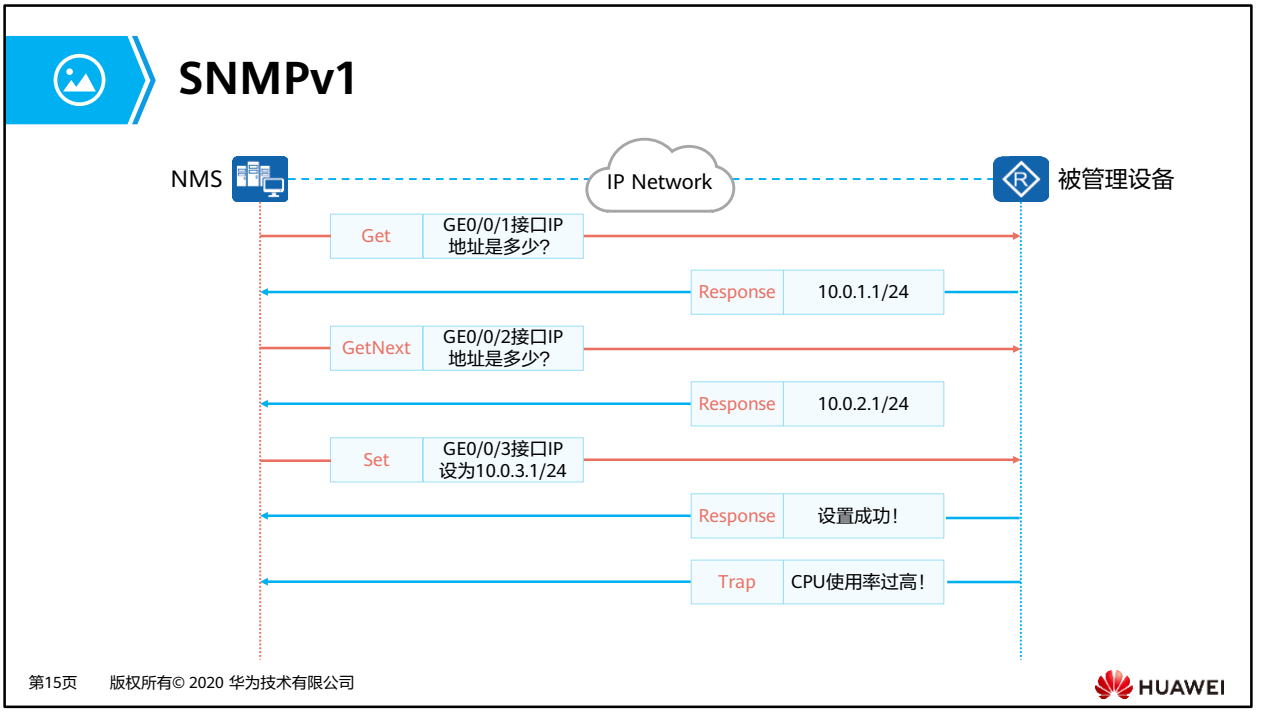
OID	节点名称	绑定变量	含义
3.6.1.6.3.1.1.5.3	linkDown	ifIndex ifAdminStatus ifOperStatus ifDesc	经检测到由于ifOperStatus节点中的其中一条通信链路已经从其他状态（但不是notPresent状态）进入Down状态。这里的其他状态由ifOperStatus的值显示。

- MIB节点的最大访问权限表明网管能够通过该MIB节点对设备进行的操作：
  - not-accessible：无法进行任何操作。
  - read-only：可以读取信息。
  - read-write：可以读取信息和修改配置。
  - read-create：可以读取信息、修改配置、新增配置和删除配置。
- 设备在生成告警时，不仅会上报当前发生的告警类型，同时会绑定一些变量。比如当发送接口linkDown告警时，需要同时绑定接口索引，接口的当前配置状态等变量。
  - ifIndex：接口索引（编号）
  - ifAdminStatus：管理状态，即接口是否被shutdown：1，undo shutdown；2，shutdown
  - ifOperStasuts：接口当前的操作状态，即接口的链路层协议状态：1，Up；2，Down
  - ifDesc：接口描述

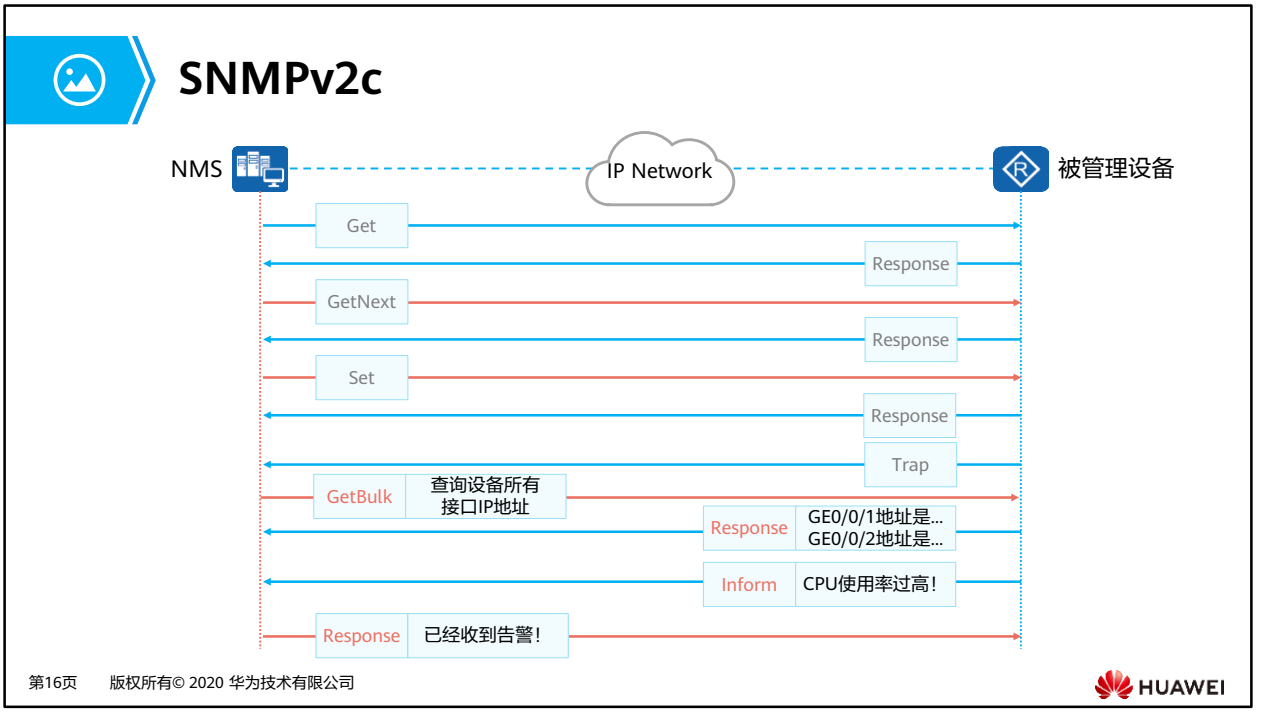
SNMP管理模型



- 查询/修改操作：
  - NMS作为管理者，向代理进程发送SNMP请求报文。
  - 代理进程通过设备端的MIB找到所要查询或修改的信息，向NMS发送SNMP响应报文。
- 告警操作：
  - 设备端的模块由于达到模块定义的告警触发条件，通过代理进程向NMS发送消息，告知设备侧出现的情况，这样便于网络管理人员及时对网络中出现的情况进行处理。



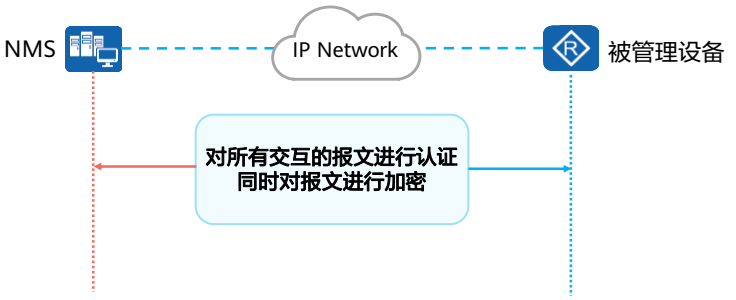
- SNMPv1定义了5种协议操作：
  - Get-Request：NMS从被管理设备的代理进程的MIB中提取一个或多个参数值。
  - Get-Next-Request：NMS从代理进程的MIB中按照字典式排序提取下一个参数值。
  - Set-Request：NMS设置代理进程MIB中的一个或多个参数值。
  - Response：代理进程返回一个或多个参数值。它是前三种操作的响应操作。
  - Trap：代理进程主动向NMS发送报文，告知设备上发生的紧急或重要事件。



- SNMPv2c新增了2种协议操作：
  - GetBulk：相当于连续执行多次GetNext操作。在NMS上可以设置被管理设备在一次GetBulk报文交互时，执行GetNext操作的次数。
  - Inform：被管理设备向NMS主动发送告警。与Trap告警不同的是，被管理设备发送Inform告警后，需要NMS进行接收确认。如果被管理设备没有收到确认信息则会将告警暂时保存在Inform缓存中，并且会重复发送该告警，直到NMS确认收到了该告警或者发送次数已经达到了最大重传次数。

SNMPv3

- SNMPv3与SNMPv1和SNMPv2c的工作机制基本一致但添加了报头数据和安全参数。
- SNMPv3报文具有身份验证和加密处理的功能。
- SNMPv3适用于各种规模的网络，安全性极高。



HUAWEI

- SNMPv3增加了身份验证和加密处理的功能。
  - 身份验证：身份验证是指代理进程（NMS）接收到信息时首先必须确认信息是否来自有权限的NMS（代理进程）并且信息在传输过程中未被改变。
  - 加密处理：SNMPv3报文中添加了报头数据和安全参数字段。比如当管理进程发出SNMPv3版本的Get-Request报文时可以携带用户名、密钥、加密参数等安全参数，代理进程回复Response报文时也采用加密的Response报文。这种安全加密机制特别适用于管理进程和代理进程之间需要经过公网传输数据的场景。

SNMP小结

- SNMP的特点如下：
  - 简单：SNMP采用轮询机制，提供基本的功能集，适合快速、低价格的场景使用，而且SNMP以UDP报文为承载，因而得到绝大多数设备的支持。
  - 强大：SNMP的目标是保证管理信息在任意两点传送，便于管理员在网络上的任何节点检索信息，进行故障排查。
- SNMPv1版本适用于小型网络。组网简单、安全性要求不高或网络环境比较安全且比较稳定的网络，比如校园网，小型企业网。
- SNMPv2c版本适用于大中型网络。安全性要求不高或者网络环境比较安全，但业务比较繁忙，有可能发生流量拥塞的网络。
- SNMPv3版本作为推荐版本，适用于各种规模的网络。尤其是对安全性要求较高，只有合法的管理员才能对网络设备进行管理的网络。

SNMP基本配置（1）

1. 使能SNMP代理功能

   [Huawei] **snmp-agent**

2. 配置SNMP的版本

   [Huawei] **snmp-agent sys-info version [v1 | v2c | v3]**

   用户可以根据自己的需求配置对应的SNMP版本，但设备侧使用的协议版本必须与网管侧一致。

3. 创建或者更新MIB视图的信息

   [Huawei] **snmp-agent mib-view** *view-name* { **exclude** | **include** } *subtree-name* [**mask** *mask*]

4. 增加一个新的SNMP组，将该组用户映射到SNMP视图

   [Huawei] **snmp-agent group v3** *group-name* **{ authentication | noauth | privacy } [read-view** *view-name* **| write-view** *view-name* **| notify-view** *view-name* **]**

   该命令用于SNMPv3版本中创建SNMP组，指定认证加密方式、只读视图、读写视图、通知视图。是安全性需求较高的网管网络中的必需指令。

SNMP基本配置 (2)

5. 为一个SNMP组添加一个新用户

```
[Huawei] snmp-agent usm-user v3 user-name group group-name
```

6. 配置SNMPv3用户认证密码

```
[Huawei] snmp-agent usm-user v3 user-name  authentication-mode { md5 | sha | sha2-256 }
```

7. 配置SNMPv3用户加密密码

```
[Huawei] snmp-agent usm-user v3 user-name privacy-mode { aes128 | des56 }
```

8. 配置设备发送Trap报文的参数信息

```
[Huawei] snmp-agent target-host trap-paramsname paramsname v3 securityname securityname
{ authentication | noauthnopriv | privacy }
```

SNMP基本配置 (3)

9. 配置Trap报文的目的主机

```
[Huawei] snmp-agent target-host trap-hostname hostname address ipv4-address trap-paramsname paramsname
```

10. 打开设备的所有告警开关

```
[Huawei] snmp-agent trap enable
```

注意该命令只是打开设备发送Trap告警的功能，要与**snmp-agent target-host**协同使用，由**snmp-agent target-host**指定Trap告警发送给哪台设备。

11. 配置发送告警的源接口。

```
[Huawei] snmp-agent trap source interface-type  interface-number
```

注意Trap告警无论从哪个接口发出都必须有一个发送的源地址，因此源接口必须是已经配置了IP地址的接口。

SNMP配置举例（网络设备侧）

- 上述路由器R1上使能SNMP功能，配置版本为v3。
- 配置SNMPv3组名为test，加密认证方式为privacy。
- 创建SNMPv3用户，名为R1同时配置认证和加密密码为HCIA-Datacom123。
- 创建名为param的Trap参数信息，securityname为sec
- 设置SNMP告警主机地址为192.168.1.10。
- 打开告警开关，设置发送告警的源接口为GE0/0/1。

R1配置如下：

```
[R1]snmp-agent
[R1]snmp-agent sys-info version v3
[R1]snmp-agent group v3 test privacy
[R1]snmp-agent usm-user v3 R1 test authentication-mode
md5 HCIA@Datacom123 privacy-mode aes128 HCIA-
Datacom123
[R1]snmp-agent target-host trap-paramsname param v3
securityname sec privacy
[R1]snmp-agent target-host trap-hostname nms address
192.168.1.10 trap-paramsname param
[R1]snmp-agent trap source GigabitEthernet 0/0/1
[R1]snmp-agent trap enable
Info: All switches of SNMP trap/notification will be open.
Continue? [Y/N]:y
```

目录

HUAWEI

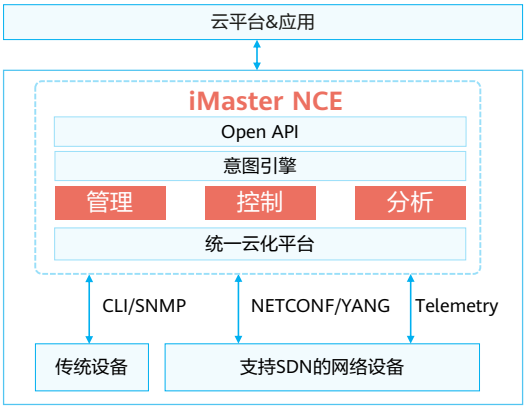
网络产业的变革与挑战

- 伴随5G和云时代的到来，VR/AR、直播、无人驾驶等各类创新性业务大量涌现，整个ICT产业迸发出蓬勃生机。与此同时，整个网络的流量也呈现出爆炸式增长，华为GIV（Global Industry Vision，全球产业展望）预计，2025年新增的数据量将达到180 ZB。业务的动态复杂性也使得整个网络复杂度不断攀升。
- 整体来看，这些问题的源头都指向了现有的网络系统，只有通过构建自动化、智能化的以用户体验为中心的网络系统才能有效应对。

- 泽字节（Zettabyte，ZB），1 ZB=10^{12} GB

华为iMaster NCE

- 华为iMaster NCE是一款集管理、控制、分析和AI智能功能于一体的网络自动化与智能化平台。



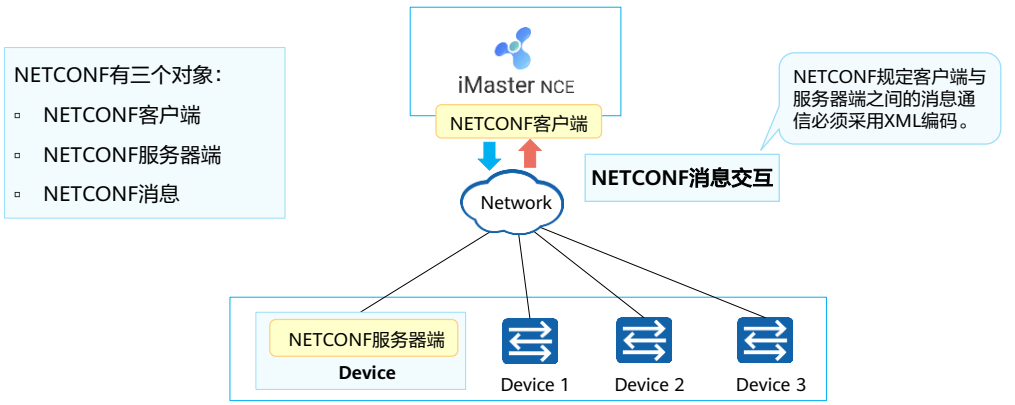
- 在管理与控制方面，iMaster NCE支持：
  - CLI和SNMP等传统技术实现传统设备的管理和控制。
  - NETCONF（基于YANG模型）协议实现对支持SDN的网络的管理和控制。
- iMaster NCE通过SNMP、Telemetry等协议采集网络数据，结合AI算法进行大数据智能分析，通过Dashboard、报表等方式多维度呈现设备及网络状态、帮助运维人员快速发现设备及网络异常情况并处理，保障设备和网络的正常运作。

HUAWEI

- iMaster NCE包含四大关键能力：
  - 全生命周期自动化：以统一的资源建模和数据共享服务为基础，提供跨多网络技术域的全生命周期的自动化能力，实现设备即插即用、网络即换即通、业务自助服务、故障自愈和风险预警。
  - 基于大数据和AI的智能闭环：基于意图、自动化、分析和智能四大子引擎构建完整的智能化闭环系统。基于Telemetry采集并汇聚海量的网络数据，iMaster NCE实现实时网络态势感知，通过统一的数据建模构建基于大数据的网络全局分析和洞察，并注入基于华为30多年电信领域经验积累的AI算法，面向用户意图进行自动化闭环的分析、预测和决策，提升客户满意度，持续提升网络的智能化水平。
  - 开放可编程使能场景化APP生态：iMaster NCE对外提供可编程的集成开发环境Design Studio和开发者社区，实现南向与第三方网络控制器或网络设备对接，北向与云端AI训练平台和IT应用快速集成，并支持客户灵活选购华为原生APP，客户自行开发或寻求第三方系统集成商的支持进行APP的创新与开发。
  - 大容量全云化平台：基于Cloud Native的云化架构，iMaster NCE支持在私有云、公有云中运行，也支持On-premise部署模式，具备大容量和弹性可伸缩能力，支持大规模系统容量和用户接入，让网络从数据分散、多级运维的离线模式转变为数据共享、流程打通的在线模式。

NETCONF简介

- NETCONF（Network Configuration Protocol，网络配置协议），提供一套管理网络设备的机制。用户可以使用这套机制增加、修改、删除网络设备的配置，获取网络设备的配置和状态信息。

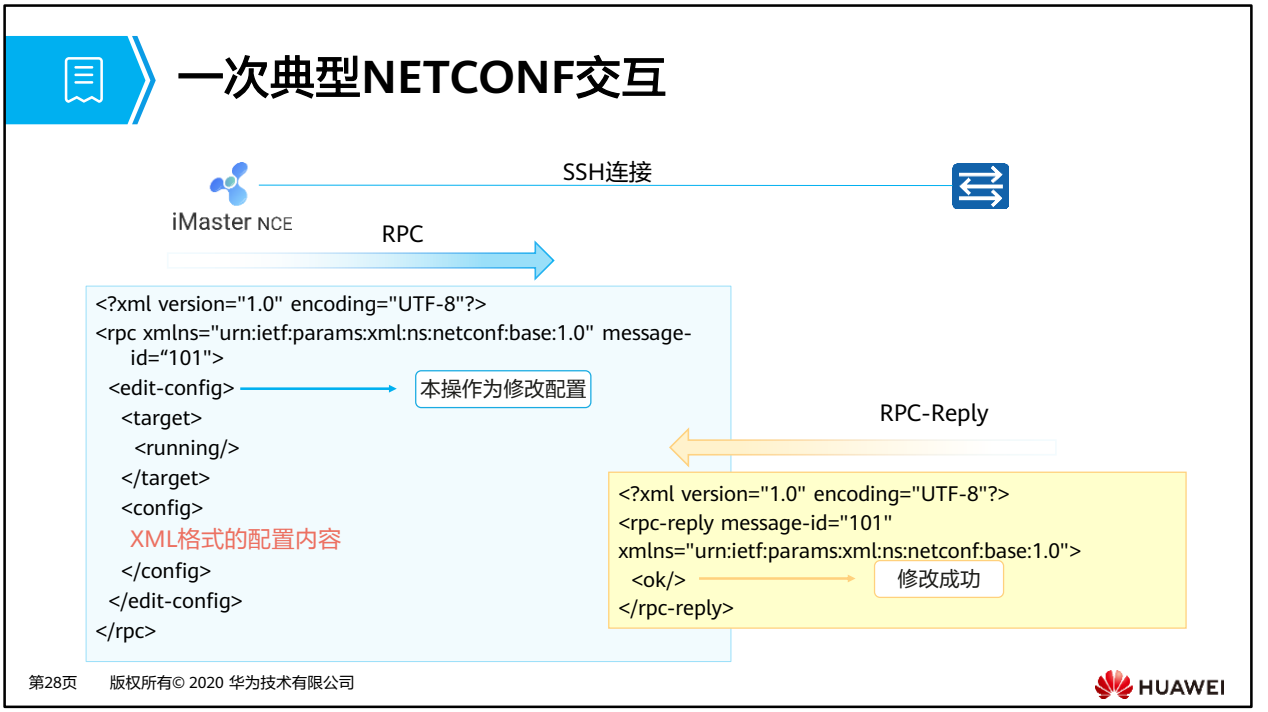


HUAWEI

- NETCONF客户端（Client）：Client 利用NETCONF协议对网络设备进行系统管理。一般由网络管理系统（NMS）作为NETCONF Client。Client向Server发送<rpc>请求，查询或修改一个或多个具体的参数值。Client可以接收Server发送的告警和事件，以获取被管理设备的状态。
- NETCONF服务器端（Server）：Server用于维护被管理设备的信息数据并响应Client的请求，把管理数据汇报给Client。一般由网络设备（例如交换机、路由器等）作为NETCONF Server。Server收到Client 的请求后会进行数据解析，并在CMF（Configuration Manager Frame，配置管理框架）的帮助下处理请求，然后给Client 返回响应。当设备发生故障或其他事件时，Server利用Notification机制将设备的告警和事件通知给Client，向网络管理系统报告设备的当前状态变化。
- Client与Server之间建立基于SSH（Secure Shell，安全外壳）或TLS（Transport Layer Security，传输层安全性协议）等安全传输协议的连接，然后通过Hello报文交换双方支持的能力后建立NETCONF会话，Client即可与Server之间进行交互请求，网络设备必须至少支持一个NETCONF会话。Client从运行的Server上获取的信息包括配置数据和状态数据。

NETCONF的优势

功能	NETCONF	SNMP	CLI
接口类型	机机接口：接口定义完善、规范、标准，便于接口控制和使用。	机机接口	人机接口
操作效率	高：基于对象建模，对象操作一次交互即可，支持过滤、批量等操作。	中	低
扩展能力	可以扩展协议私有能力。	弱	一般
事务处理	支持：试运行、出错回滚、配置回退等事务处理机制。	不支持	部分支持
安全传输	多种安全协议：SSH，TLS，BEEP/TLS，SOAP/HTTP/TLS	仅SNMPv3支持	支持SSH



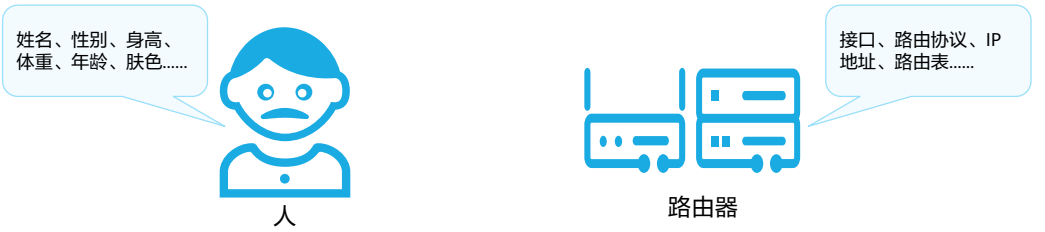
- NETCONF使用SSH实现安全传输，使用RPC（Remote Procedure Call，远程过程调用）实现客户端和服务器端的通信。

YANG语言概述

- YANG（Yet Another Next Generation）是一种数据建模语言，实现了NETCONF数据内容的标准化。
- YANG模型定义了数据的层次化结构，可用于基于NETCONF的操作。建模对象包括配置、状态数据、远程过程调用和通知。它可以对NETCONF客户端和服务器端之间发送的所有数据进行一个完整的描述。

模型（Model）是对“事物”的一种抽象和表达。

数据模型（Data Model）是对数据特征的抽象和表达。

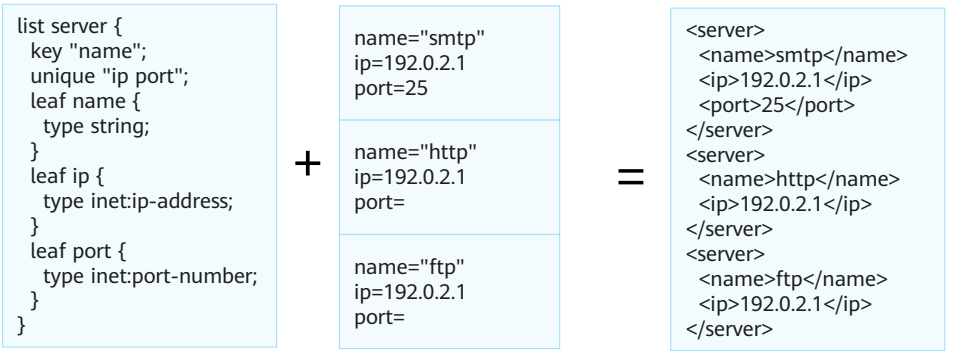


HUAWEI

- YANG起源于NETCONF，但不仅用于NETCONF。虽然统一了YANG建模语言，但是YANG文件没有统一。
- YANG文件可以简单分为三类：
  - 厂家私有YANG文件
  - IETF标准YANG
  - OpenConfig YANG
- YANG模型的最终呈现是.yang为后缀的文件。
- YANG模型的特点：
  - 基于层次化的树状结构建模。
  - 数据模型以模块和子模块呈现。
  - 可以和基于XML的语法的YIN（YANG Independent Notation）模型无损转换。
  - 定义内置的数据类型和允许可扩展类型。

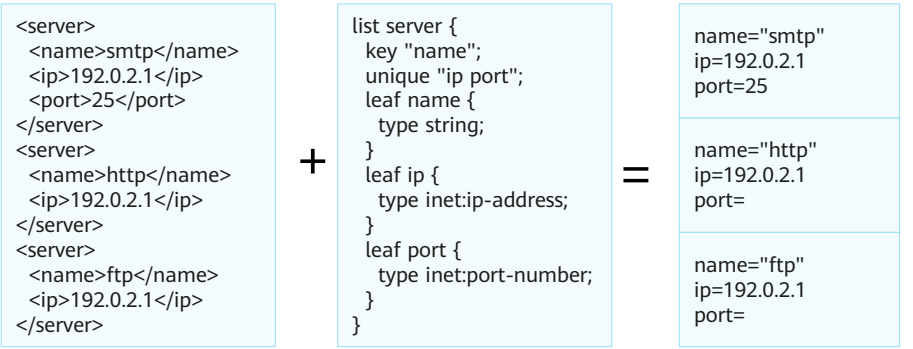
YANG与XML（1）

- 在NETCONF客户端（例如网管平台/SDN控制器）加载YANG文件。
- 通过YANG文件将数据转换为XML格式的NETCONF消息发送到设备。



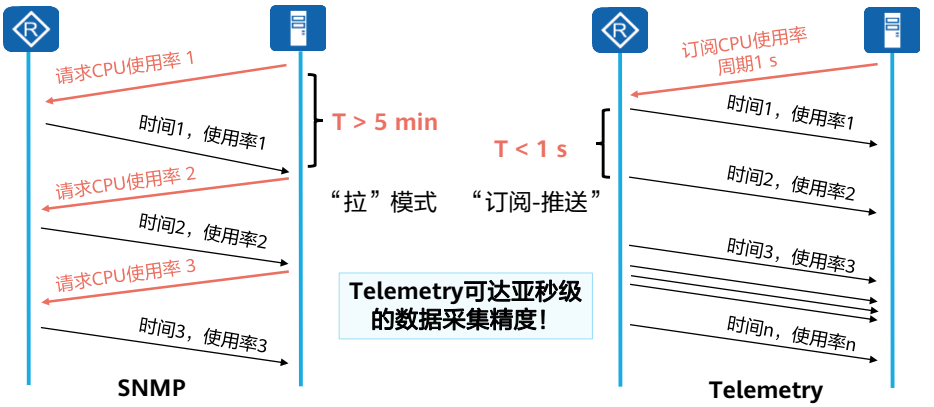
YANG与XML（2）

- 在NETCONF服务器（例如路由器/交换机等）加载YANG文件。
- 通过YANG文件将接收到的XML格式的NETCONF消息转换为数据并做后续处理。



Telemetry基本概述

- Telemetry也作Network Telemetry，即网络遥测技术，是一项远程地从物理设备或虚拟设备上高速采集数据的技术。
- 设备通过推模式（Push Mode）周期性地主动向采集器上送设备的接口流量统计、CPU或内存数据等信息，相对传统拉模式（Pull Mode）的一问一答式交互，提供了更实时更高速的数据采集功能。



- 业界也有一种看法，将SNMP认为是传统的Telemetry技术，把当前Telemetry叫做Streaming Telemetry或Model-Driven Telemetry。
- Telemetry将上送数据打包一起发送，提升传输效率。

思考题

1. （单选）在基于SNMP进行管理的网络中，（　）运行管理进程，对被管理设备进行管理。

   A. NMS

   B. 代理进程

   C. MIB

   D. SNMP

2. （单选）SNMPv1协议中被管理设备上报告警的协议操作是（　）。

   A. Get-Request

   B. Set-Request

   C. Trap

   D. Response

1. A
2. C

思考题

3. YANG是一种数据建模语言。（　　）

   A. True

   B. False

4. Telemetry上限可达到亚秒级的数据采集精度。（　　）

   A. True

   B. False

3. A

4. A

本章总结

- 随着网络技术不断发展，网络的管理与运维手段也越来越多，常见的有：
  - CLI或Web方式
  - SNMP协议方式
  - 通过华为iMaster NCE“管-控-析”智能化运维平台方式


谢 谢
www.huawei.com

IPv6基础

前言

- 20世纪80年代，IETF（Internet Engineering Task Force，因特网工程任务组）发布RFC791，即IPv4协议，标志IPv4正式标准化。在此后的几十年间，IPv4协议成为最主流的协议之一。无数人在IPv4的基础上开发出了各种应用，并且对这个协议做了各种补充和增强，支撑起了今天繁荣的互联网。
- 然而，随着互联网的规模越来越大，以及5G、物联网等新兴技术的发展，IPv4面临的挑战越来越多。IPv6取代IPv4势在必行。
- 本章节描述了为什么需要从IPv4向IPv6进行演进，以及一些关于IPv6的基础知识。

- IPv4（Internet Protocol version 4）：互联网协议（IP）的当前版本。IPv4地址为32 bit编码，通常用4个点分十进制数表示。每个地址由一个网络码、（可选）子网码、主机码组成。网络码和（可选）子网码用于路由，主机码用于在网络或子网内部寻址到一台具体主机。
- IPv6（Internet Protocol version 6）：IETF设计的一套规范，是IPv4的升级版本。它是网络层协议的第二代标准协议，也被称为Ipng（IP Next Generation）。IPv6和IPv4之间最显著的区别就是IP地址的长度从32 bit升为128 bit。

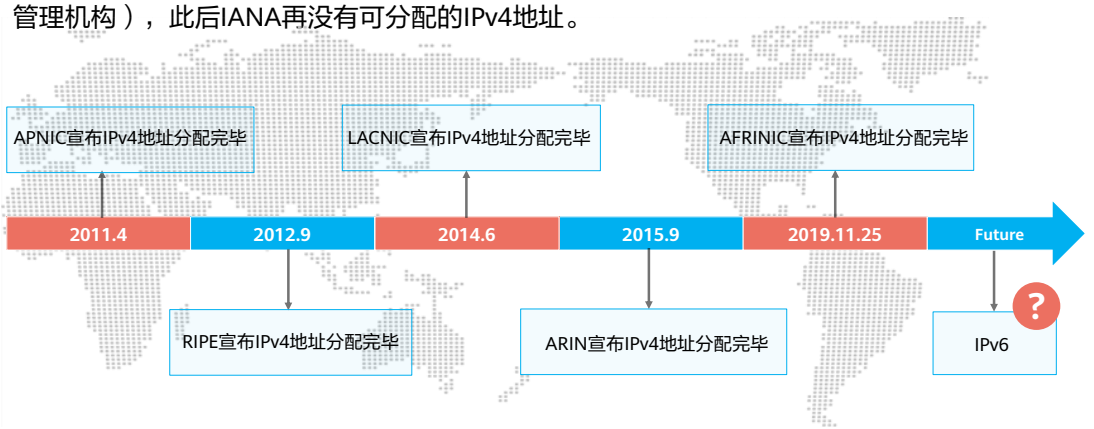
目标

- 学完本课程后，您将能够：
  - 概括IPv6相较于IPv4的优势
  - 描述IPv6的基本概念
  - 描述IPv6报文头部的格式和原理
  - 描述IPv6地址格式和地址类型
  - 描述IPv6地址配置的方法和基本过程
  - 执行IPv6地址以及IPv6静态路由的简单配置

目录

IPv4现状

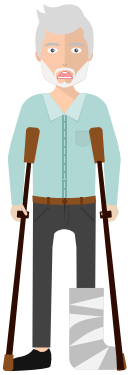
- 2011年2月3日，IANA（Internet Assigned Numbers Authority，因特网地址分配组织）宣布将其最后的468万个IPv4地址平均分配到全球5个RIR（Regional Internet Registry，区域互联网注册管理机构），此后IANA再没有可分配的IPv4地址。



HUAWEI

- IANA，是负责全球互联网IP地址编号分配的机构。IANA将部分IPv4地址分配给大洲级的RIR，再由各RIR进行所辖区域内地址分配，五大RIR包括：
  - RIPE：Reseaux IP Europeans，欧洲IP地址注册中心，服务于欧洲、中东地区和中亚地区；
  - LACNIC：Latin American and Caribbean Internet Address Registry，拉丁美洲和加勒比海Internet地址注册中心，服务于中美、南美以及加勒比海地区；
  - ARIN：American Registry for Internet Numbers，美国Internet编号注册中心，服务于北美地区和部分加勒比海地区；
  - AFRINIC：Africa Network Information Centre，非洲网络信息中心，服务于非洲地区；
  - APNIC：Asia Pacific Network Information Centre，亚太互联网络信息中心，服务于亚洲和太平洋地区。
- 实践证明IPv4是一个非常成功的协议，它本身也经受住了Internet从少量计算机组网发展到目前上亿台计算机互联的考验。但该协议是几十年前基于当时的网络规模而设计的。在今天看来，IPv4的设计者们对于Internet的估计和预想显得很不充分。随着Internet的扩张和新应用的不断推出，IPv4越来越显示出它的局限性。
- Internet规模的快速扩张是当时完全没有预料到的，特别是近十年来，更是爆炸式增长，已经走进了千家万户，人们的日常生活已经离不开它。但正因为发展太快，IP地址空间耗尽的问题迫在眉睫。
- 20世纪90年代，IETF推出NAT（Network Address Translation，网络地址转换）与CIDR（Classless Inter Domain Routing，无类别域间路由）等技术来推迟IPv4地址耗尽发生的时间点。但是这些过渡方案只能减缓地址枯竭的速度，并不能从根本上解决问题。

Why IPv6 ?



IPv4

公网地址枯竭

包头设计不合理

路由表过大，查表效率低

对ARP的依赖，导致广播泛滥

.....

VS

IPv6

“无限”地址

地址层次化分配

即插即用

简化的报文头部

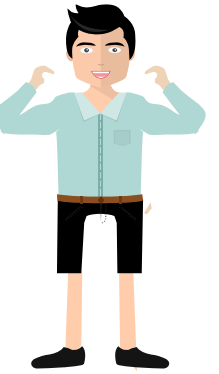
IPv6安全特性

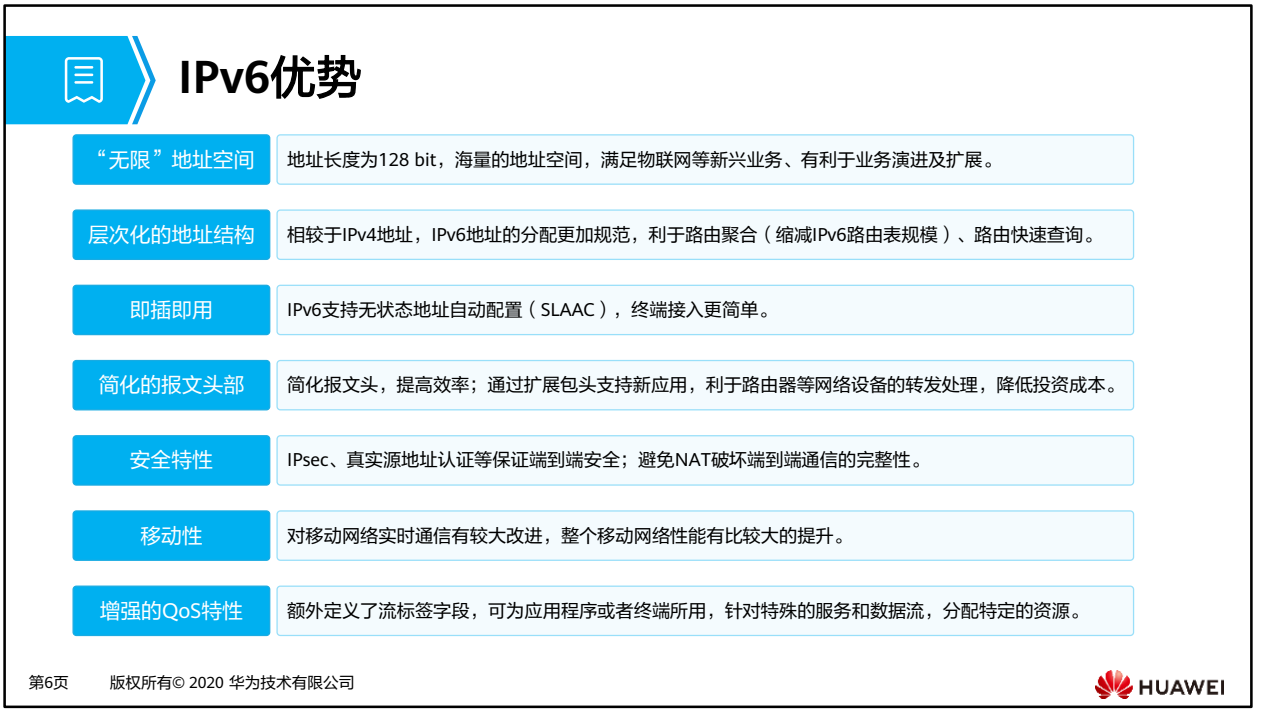
保证端到端通信完整性

对移动性的支持

增强的QoS特性

.....



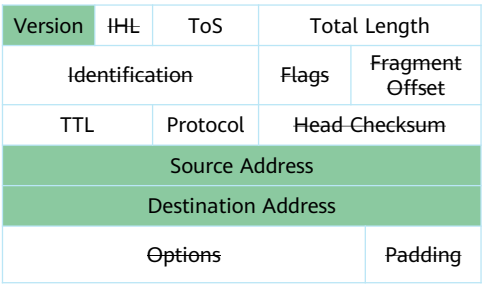


- 近乎无限的地址空间：与IPv4相比，这是最明显的好处。IPv6地址是由128 bit构成，单从数量级来说，IPv6所拥有的地址容量是IPv4的约8×10^{28}倍，号称可以为全世界的每一粒沙分配一个网络地址。这使得海量终端同时在线，统一编址管理，变为可能，为万物互连提供了强有力的支撑。
- 层次化的地址结构：正因为有了近乎无限的地址空间，IPv6在地址规划时就根据使用场景划分了各种地址段。同时严格要求单播IPv6地址段的连续性，禁止出现IPv4的地址“打洞”现象，便于IPv6路由聚合，缩小IPv6地址表规模。
- 即插即用：任何主机或者终端要获取网络资源，传输数据，都必须有明确的IP地址。传统的分配IP地址方式是手工或者DHCP自动获取，除了上述两个方式外，IPv6还支持SLAAC（Stateless Address Autoconfiguration，无状态地址自动配置）。
- 端到端网络的完整性：大面积使用NAT技术的IPv4网络，从根本上破坏了端到端连接的完整性。使用IPv6之后，将不再需要NAT网络设备，上网行为管理、网络监管等将变得简单，与此同时，应用程序也不需要开发复杂的NAT适配代码。
- 安全性得到增强：IPsec（ Internet Protocol Security，因特网协议安全协议）最初是为IPv6设计的，所以基于IPv6的各种协议报文（路由协议、邻居发现等），都可以端到端地加密，当然该功能目前应用并不多。而IPv6的数据面报文安全性，跟IPv4+IPsec的能力，基本相同。
- 可扩展性强：IPv6的扩展属性报文头部，并不是主数据包的一部分，但是在必要的时候，这些扩展头部会插在IPv6基本头部和有效载荷之间，能够协助IPv6完成加密功能、移动功能、最优路径选路、QoS等，并可提高报文转发效率。
- 移动性改善：当一个用户从一个网段移动到另外一个网段，传统的网络会产生经典式“三角式路由”，IPv6网络中，这种移动设备的通信，可不再经过原“三角式路由”，而做直接路由转发，降低了流量转发的成本，提升了网络性能和可靠性。
- QoS可得到进一步增强：IPv6保留了IPv4所有的QoS属性，额外定义了20 Byte的流标签字段，可为应用程序或者终端所用，针对特殊的服务和数据流，分配特定的资源，目前该机制并没有得到充分的开发和应用。

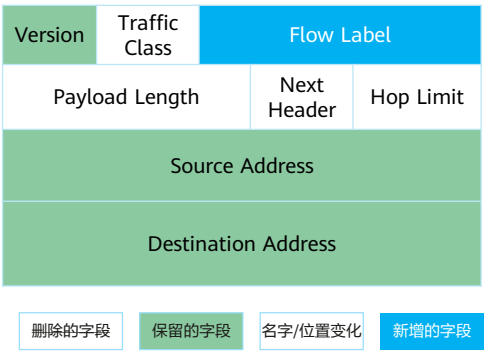
IPv6基本包头

- IPv6包头由一个IPv6基本包头（必须存在）和多个扩展包头（可能不存在）组成。
- 基本包头提供报文转发的基本信息，会被转发路径上的所有设备解析。

IPv4包头（20 Byte ~ 60 Byte ）

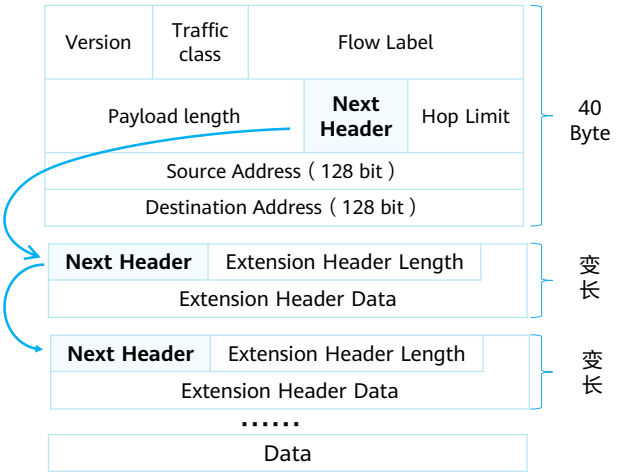


IPv6基本包头（40 Byte ）

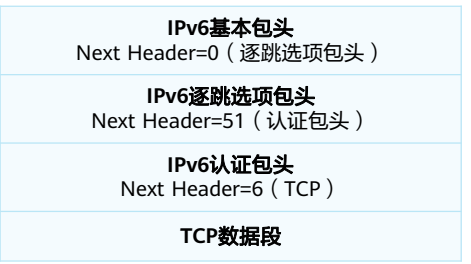


- IPv6基本包头字段解释如下：
  - Version：版本号，长度为4 bit。对于IPv6，该值为6。
  - Traffic Class：流类别，长度为8 bit。等同于IPv4中的ToS字段，表示IPv6数据包的类或优先级，主要应用于QoS。
  - Flow Label：流标签，长度为20 bit。IPv6中的新增字段，用于区分实时流量，不同的流标签+源地址可以唯一确定一条数据流，中间网络设备可以根据这些信息更加高效率的区分数据流。
  - Payload Length：有效载荷长度，长度为16 bit。有效载荷是指紧跟IPv6包头的数据包的其他部分（即扩展包头和上层协议数据单元）。
  - Next Header：下一个包头，长度为8 bit。该字段定义紧跟在IPv6包头后面的第一个扩展包头（如果存在）的类型，或者上层协议数据单元中的协议类型（类似于IPv4的Protocol字段）。
  - Hop Limit：跳数限制，长度为8 bit。该字段类似于IPv4中的Time to Live字段，它定义了IP数据包所能经过的最大跳数。每经过一个路由器，该数值减去1，当该字段的值为0时，数据包将被丢弃。
  - Source Address：源地址，长度为128 bit。表示发送方的地址。
  - Destination Address：目的地址，长度为128 bit。表示接收方的地址。

IPv6拓展包头



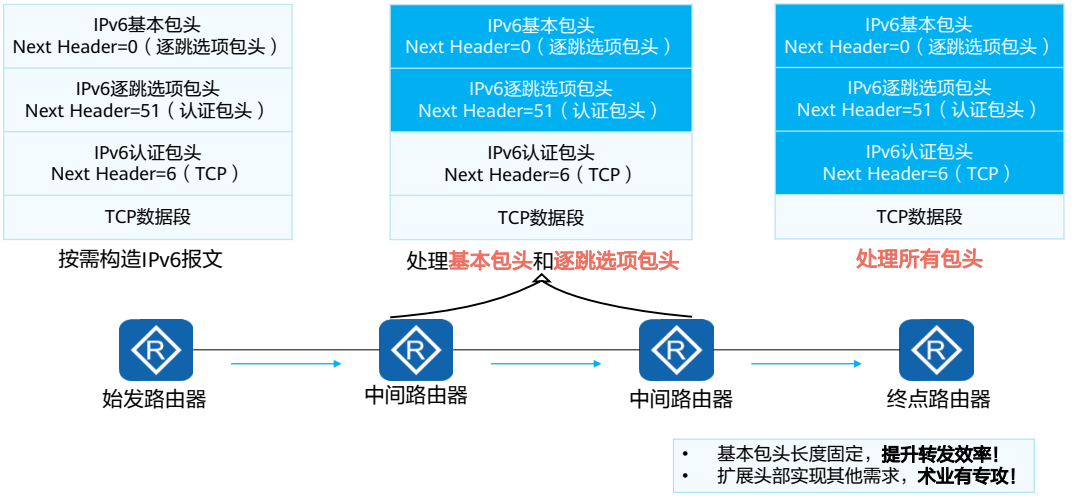
- **Extension Header Length：**扩展包头长度，长度为8 bit。表示扩展包头的长度（不包含Next Header字段）。
- **Extension Header Data：**扩展包头数据，长度可变。扩展包头的内容，为一系列选项字段和填充字段的组合。



IPv6报文示例

- IPv4包头包含可选字段Options，内容涉及Security、Timestamp、Record route等，这些Options可以将IPv4包头长度从20 Byte扩充到60 Byte。携带这些Options的IPv4报文在转发过程中往往需要中间路由转发设备进行软件处理，对于性能是个很大的消耗，因此实际中也很少使用。
- IPv6将这些Options从IPv6基本包头中剥离，放到了扩展包头中，扩展包头被置于IPv6基本包头和上层协议数据单元之间。一个IPv6报文可以包含0个、1个或多个扩展包头，仅当需要路由器或目的节点做某些特殊处理时，才由发送方添加一个或多个扩展头。与IPv4不同，IPv6扩展头长度任意，不受40 Byte限制，这样便于日后扩充新增选项。这一特征加上选项的处理方式使得IPv6选项能得以真正的利用。但是为了提高处理选项头和传输层协议的性能，扩展包头总是8 Byte长度的整数倍。
- 当使用多个扩展包头时，前面包头的Next Header字段指明下一个扩展包头的类型，这样就形成了链状的包头列表。
- 当超过一种扩展包头被用在同一个IPv6报文里时，包头必须按照下列顺序出现：
    1. 逐跳选项包头：主要用于为在传送路径上的每跳转发指定发送参数，传送路径上的每台中间节点都要读取并处理该字段。
    2. 目的选项包头：携带了一些只有目的节点才会处理的信息。
    3. 路由包头：IPv6源节点用来强制数据包经过特定的设备。
    4. 分段包头：当报文长度超过MTU（Maximum Transmission Unit，最大传输单元）时就需要将报文分段发送，而在IPv6中，分段发送使用的是分段包头。
    5. 认证包头（AH）：该包头由IPsec使用，提供认证、数据完整性以及重放保护。
    6. 封装安全净载包头（ESP）：该包头由IPsec使用，提供认证、数据完整性以及重放保护和IPv6数据包的保密。

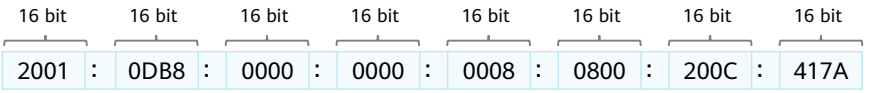
IPv6报文处理机制



- 基本包头长度固定，**提升转发效率！**
- 扩展头部实现其他需求，**术业有专攻！**

IPv6地址

- IPv6地址的长度为128 bit。一般用冒号分割为8段，每一段16 bit，每一段内用十六进制表示。



IPv6地址中的字母大小写不敏感，例如A等同于a。

- 与IPv4地址类似，IPv6也用“IPv6地址/掩码长度”的方式来表示IPv6地址。
  - 例如2001:0DB8:2345:CD30:1230:4567:89AB:CDEF/64

IPv6地址： 2001:0DB8:2345:CD30:1230:4567:89AB:CDEF

子网号： 2001:0DB8:2345:CD30::/64

IPv6地址缩写规范

- 为了书写方便，IPv6可采用以下规则进行缩写。

IPv6地址缩写规范

2001 : 0DB8 : 0000 : 0000 : 0008 : 0800 : 200C : 417A

每组16 bit的单元中的前导0可以省略，但是如果16 bit单元的所有比特都为0，那么至少要保留一个“0”字符；拖尾的0不能被省略。

2001 : DB8 : 0 : 0 : 8 : 800 : 200C : 417A

一个或多个连续的16 bit字符为0时，可用“::”表示，但整个IPv6地址缩写中只允许有一个“::”。

2001 : DB8 : : 8 : 800 : 200C : 417A

若缩写后的IPv6地址出现两个“::”，会导致无法还原为原始IPv6地址。

IPv6地址缩写示例

缩写前 0000:0000:0000:0000:0000:0000:0000:0001
缩写后 ::1

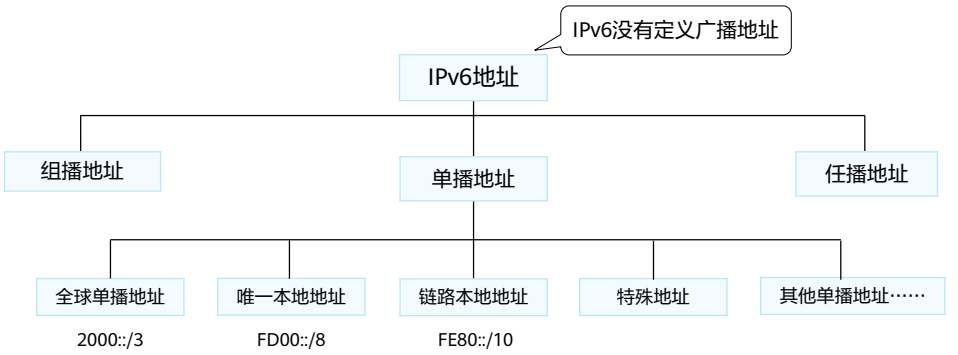
缩写前 2001:0DB8:0000:0000:FB00:1400:5000:45FF
缩写后 2001:DB8::FB00:1400:5000:45FF

缩写前 2001:0DB8:0000:0000:0000:2A2A:0000:0001
缩写后 2001:DB8::2A2A:0:1

缩写前 2001:0DB8:0000:1234:FB00:0000:5000:45FF
缩写后 2001:DB8::1234:FB00:0:5000:45FF
或 2001:DB8:0:1234:FB00::5000:45FF

IPv6地址分类

- 根据IPv6地址前缀，可将IPv6地址分为为单播地址、组播地址和任播地址。

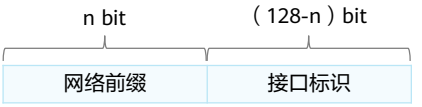


HUAWEI

- 单播地址（Unicast Address）：标识一个接口，目的地址为单播地址的报文会被送到被标识的接口。在IPv6中，一个接口拥有多个IPv6地址是非常常见的现象。除了全球单播地址、唯一本地地址、链路本地地址这三种地址之外，IPv6还有一些特殊单播地址的存在：
  - 未指定地址：0:0:0:0:0:0:0:0/128 或者::/128。该地址作为某些报文的源地址，比如作为重复地址检测时发送的邻居请求报文（NS）的源地址，或者DHCPv6初始化过程中客户端所发送的请求报文的源地址。
  - 环回地址：0:0:0:0:0:0:0:1/128 或者::1/128，与IPv4中的127.0.0.1作用相同，用于本地回环，发往::/1的数据包实际上就是发给本地，可用于本地协议栈环回测试。
- 组播地址（Multicast Address）：标识多个接口，目的地址为组播地址的报文会被送到被标识的所有接口。只有加入相应组播组的设备接口才会侦听发往该组播地址的报文。
- 任播地址（Anycast Address）：任播地址标识一组网络接口（通常属于不同的节点）。目标地址是任播地址的数据包将发送给其中路由意义上最近的一个网络接口。
- IPv6没有定义广播地址（Broadcast Address）。在IPv6网络中，所有广播的应用层场景将会被IPv6组播所取代。

IPv6单播地址结构

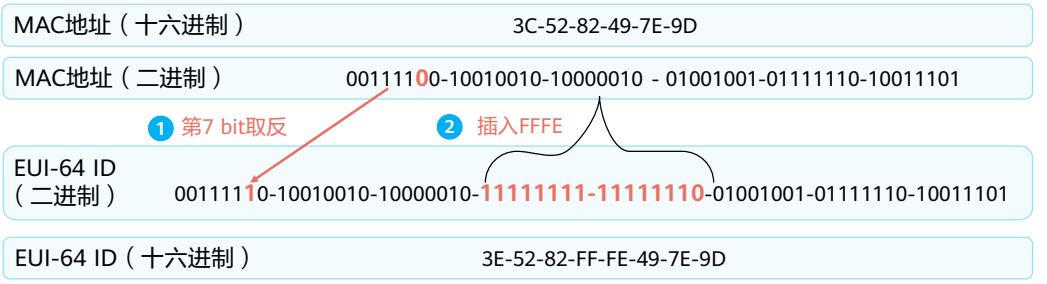
- 一个IPv6单播地址可以分为如下两部分：
  - 网络前缀（Network Prefix）：n bit，相当于IPv4地址中的网络ID。
  - 接口标识（Interface Identify）：（128-n）bit，相当于IPv4地址中的主机ID。
- 常见的IPv6单播地址如全球单播地址、链路本地地址等，要求网络前缀和接口标识必须为64 bit。



- 全球单播地址中，高位前3 bit为000的地址可以采用非64 bit的网络前缀，这部分地址不在本课程涉及的范围中。

IPv6单播地址接口标识

- 接口标识可通过三种方法生成：
  - 手工配置
  - 系统自动生成
  - 通过IEEE EUI-64规范生成
- 其中EUI-64规范最为常用，此规范将接口的MAC地址转换为IPv6接口标识。



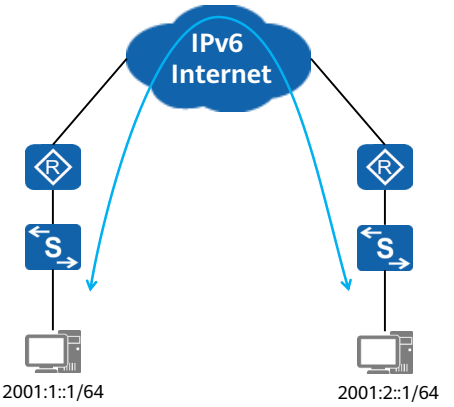
- 接口标识的长度为64 bit，用于标识链路上的接口。在每条链路上，接口标识必须唯一。接口标识有许多用途，最常见的用法就是黏贴在链路本地地址前缀后面，形成接口的链路本地地址。或者在无状态自动配置中，黏贴在获取到的IPv6全球单播地址前缀后面，构成接口的全球单播地址。
- IEEE EUI-64（64-bit Extended Unique Identifier）规范
  - 这种由MAC地址产生IPv6地址接口标识的方法可以减少配置的工作量，尤其是当采用无状态地址自动配置时，只需要获取一个IPv6前缀就可以与接口标识形成IPv6地址。
  - 使用这种方式最大的缺点就是某些恶意者可以通过二层MAC推算出三层IPv6地址。

IPv6常见单播地址 - GUA

- GUA（Global Unicast Address，全球单播地址），也被称为可聚合全球单播地址。该类地址全球唯一，用于需要有互联网访问需求的主机，相当于IPv4的公网地址。

3 bit	45 bit	16 bit	64 bit
001	全局路由前缀	子网ID	接口标识
网络部分			主机部分

- 通常GUA的网络部分长度为64 bit，接口标识也为64 bit。
- 全局路由前缀：由提供商指定给一个组织机构，一般至少为45 bit。
- 子网ID：组织机构根据自身网络需求划分子网。
- 接口标识：用来标识一个设备（的接口）。



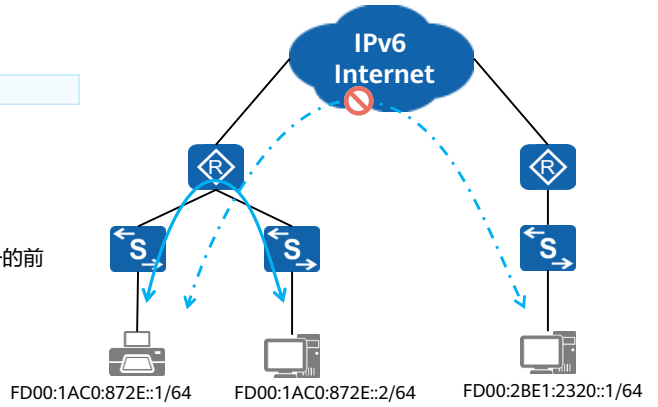
- 可以向运营商申请GUA或者直接向所在地区的IPv6地址管理机构申请。

IPv6常见单播地址 - ULA

- ULA（Unique Local Address，唯一本地地址）是IPv6私网地址，只能够在内网中使用。该地址空间在IPv6公网中不可被路由，因此不能直接访问公网。

8 bit	40 bit	16 bit	64 bit
1111 1101	Global ID	子网ID	接口标识
	伪随机产生		

- 唯一本地地址使用FC00::/7地址块，目前仅使用了FD00::/8地址段。FC00::/8预留为以后拓展用。
- ULA虽然只在有限范围内有效，但也具有全球唯一的前缀（虽然随机方式产生，但是冲突概率很低）。

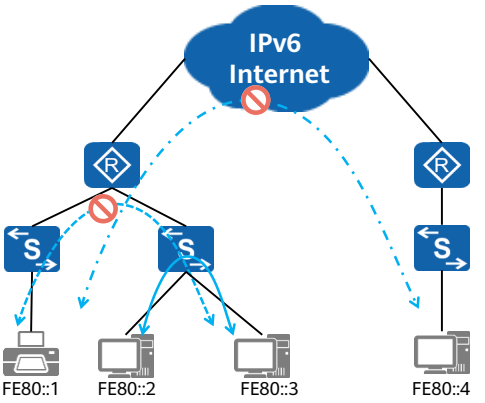


IPv6常见单播地址 - LLA

- LLA（Link-Local Address，链路本地地址）是IPv6中另一种应用范围受限制的地址类型。LLA的有效范围是本地链路，前缀为FE80::/10。

10 bit	54 bit	64 bit
1111 1110 10	0	接口标识
	固定为0	

- LLA用于一条单一链路层面的通信，例如IPv6地址无状态自动配置、IPv6邻居发现等。
- 源或目的IPv6地址为链路本地地址的数据包将不会被转发到始发的链路之外，换句话说，链路本地地址的有效范围为本地链路。
- 每一个IPv6接口都必须具备一个链路本地地址。华为设备支持自动生成和手工指定两种配置方式。



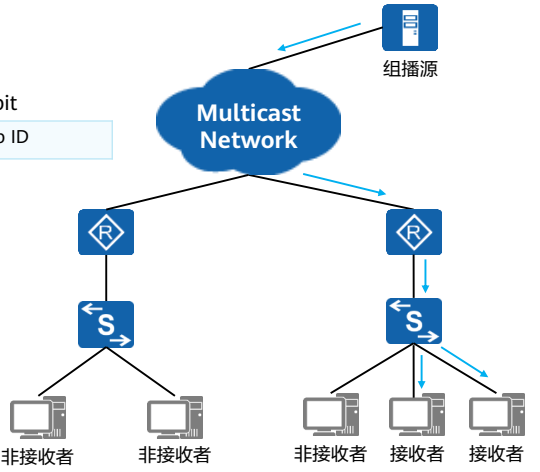
HUAWEI

IPv6组播地址

- IPv6组播地址标识多个接口，一般用于“一对多”的通信场景。
- IPv6组播地址只可以作为IPv6报文的目的地址。

8 bit	4 bit	4 bit	80 bit	32 bit
11111111	Flags	Scope	Reserved（必须为0）	Group ID

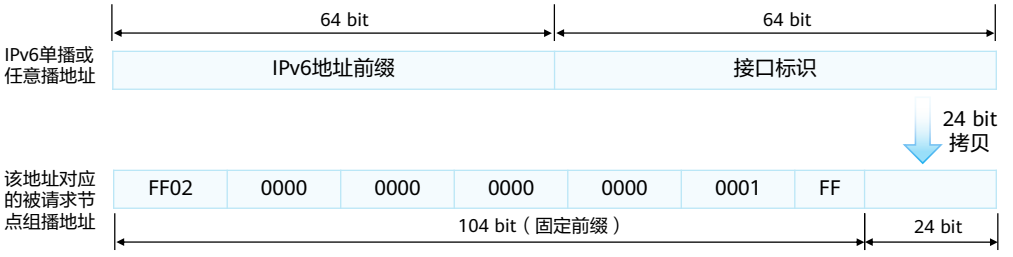
- **Flags：**用来表示永久或临时组播组。
- **Scope：**表示组播组的范围。
- **Group ID：**组播组ID。



- IPv6组播地址各字段值对应的组播组类型和范围：
  - Flags：
    - 0000表示永久分配或众所周知；
    - 0001表示 临时的。
  - Scope：
    - 0：预留；
    - 1：节点本地范围；单个接口有效，仅用于Loopback通讯。
    - 2：链路本地范围；例如FF02::1。
    - 5：站点本地范围；
    - 8：组织本地范围；
    - E：全球范围；
    - F：预留。

被请求节点组播地址

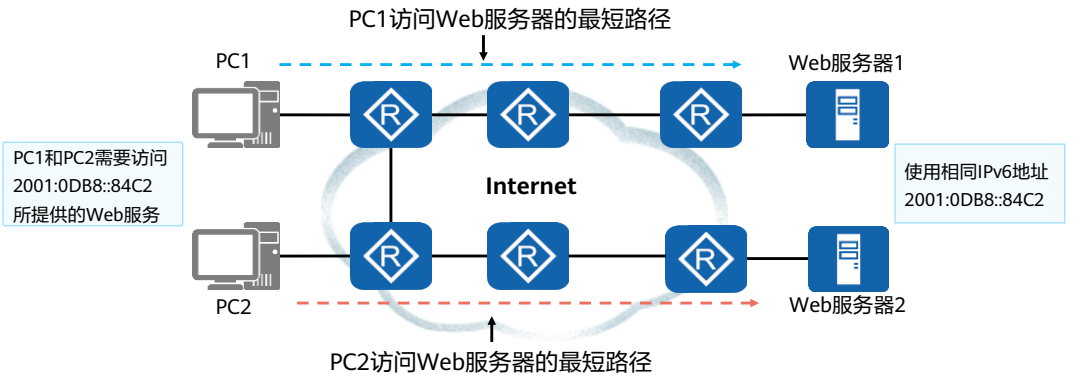
- 当一个节点具有了单播或任播地址，就会对应生成一个被请求节点组播地址，并且加入这个组播组。该地址主要用于**邻居发现机制**和**地址重复检测**功能。被请求节点组播地址的有效范围为**本地链路范围。**



- 被请求节点组播组地址的应用场景举例：在IPv6中，ARP及广播都被取消，当设备需要请求某个IPv6地址对应的MAC地址时，设备依然需要发送请求报文，但是该报文是一个组播报文，其目的IPv6地址是目标IPv6单播地址对应的被请求节点组播地址，由于只有目标节点才会侦听这个被请求节点组播地址，所以该组播报文可以被目标节点所接收，同时不会占用其他非目标节点的网络性能。

IPv6任播地址

- 任播地址标识一组网络接口（通常属于不同的节点）。任播地址可以作为IPv6报文的源地址，也可以作为目的地址。

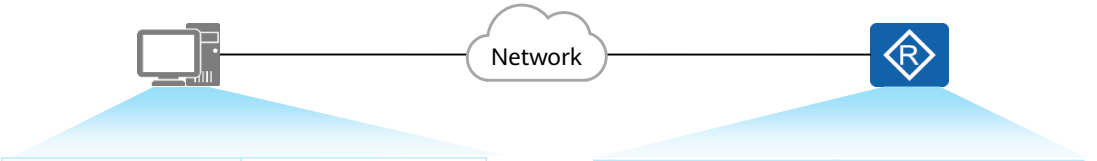


- 任播过程涉及一个任播报文发起方和一个或多个响应方。
  - 任播报文的发起方通常为请求某一服务（例如，Web服务）的主机。
  - 任播地址与单播地址在格式上无任何差异，唯一的区别是一台设备可以给多台具有相同地址的设备发送报文。
- 网络中运用任播地址有很多优势：
  - 业务冗余。比如，用户可以通过多台使用相同地址的服务器获取同一个服务（例如，Web服务）。这些服务器都是任播报文的响应方。如果不是采用任播地址通信，当其中一台服务器发生故障时，用户需要获取另一台服务器的地址才能重新建立通信。如果采用的是任播地址，当一台服务器发生故障时，任播报文的发起方能够自动与使用相同地址的另一台服务器通信，从而实现业务冗余。
  - 提供更优质的服务。比如，某公司在A省和B省各部署了一台提供相同Web服务的服务器。基于路由优选规则，A省的用户在访问该公司提供的Web服务时，会优先访问部署在A省的服务器，提高访问速度，降低访问时延，大大提升了用户体验。

目录

主机和路由器的IPv6地址

- 一般情况下，主机和路由器的单播IPv6地址以及加入的组播地址如下所示：

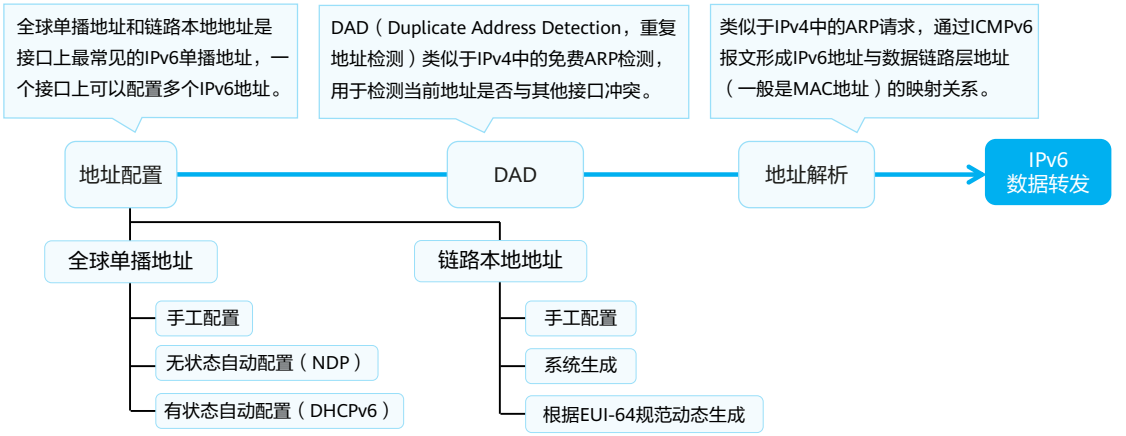


网卡的链路本地地址	FE80::2E0:FCFF:FE35:7287
管理员分配的全球单播地址	2001::1
环回地址	::1
“所有节点”组播地址	FF01::1 及 FF02::1
网卡的每个单播地址对应的被请求节点组播地址	FF02::1:FF35:7287 FF02::1:FF00:1

网卡的链路本地地址	FE80::2E0:FCFF:FE99:1285
管理员分配的全球单播地址	2001::2
环回地址	::1
“所有节点”组播地址	FF01::1 及 FF02::1
“所有路由器”组播地址	FF01::2 及 FF02::2
网卡的每个单播地址对应的被请求节点组播地址	FF02::1:FF99:1285 FF02::1:FF00:2

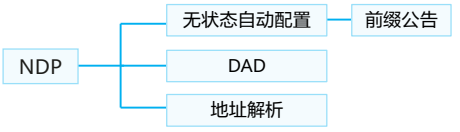
IPv6单播地址业务流程

- 一个接口在发送IPv6报文之前要经历地址配置、DAD、地址解析这三个阶段，NDP（Neighbor Discovery Protocol，邻居发现协议）扮演了重要角色。



NDP

- RFC2461定义了NDP，该RFC后来被RFC4861替代。
- NDP使用ICMPv6报文实现其功能。



NDP使用的ICMPv6报文

ICMPv6 Type	报文名称
133	路由器请求（RS）
134	路由器通告（RA）
135	邻居请求（NS）
136	邻居通告（NA）

机制	RS 133	RA 134	NS 135	NA 136
地址解析			√	√
前缀公告	√	√		
DAD			√	√

HUAWEI

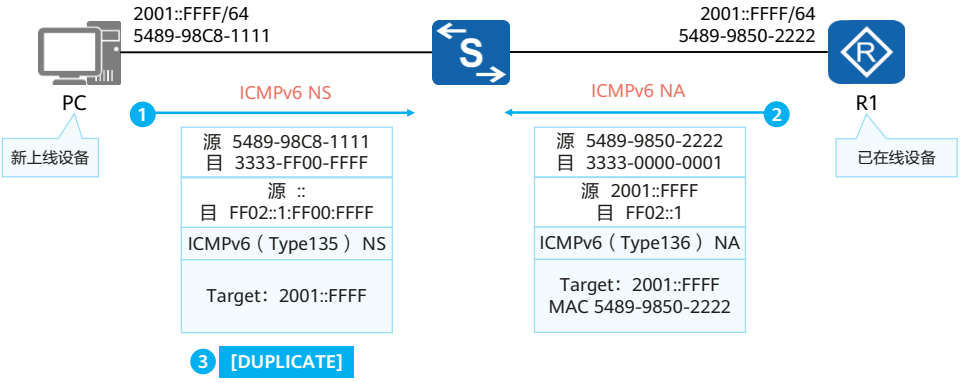
- 无状态自动配置是IPv6的一个亮点功能，它使得IPv6主机能够非常便捷地接入到IPv6网络中，即插即用，无需手工配置繁冗的IPv6地址，无需部署应用服务器（例如DHCP服务器）为主机分发地址。无状态自动配置机制使用到了ICMPv6中的路由器请求报文（Router Solicitation）及路由器通告报文（Router Advertisement）。
- 地址解析过程中使用了两种ICMPv6报文：邻居请求（Neighbor Solicitation）和邻居通告（Neighbor Advertisement）。
- 重复地址检测使用ICMPv6 NS和ICMPv6 NA报文确保网络中无两个相同的单播地址，所有接口在使用单播地址前都需要做DAD。

- IPv6支持地址有状态（stateful）和无状态（stateless）两种自动配置方式，通过ICMPv6 RA报文中的M标记（Managed Address Configuration Flag）和O标记（Other Stateful Configuration Flag）来控制终端自动获取地址的方式。
- 有状态地址配置（DHCPv6 ），M=1，O=1：
  - 采用DHCPv6协议，IPv6客户端将从DHCPv6服务器端获取完整的128 bit IPv6地址，同时包括DNS、SNTP服务器等地址参数。
  - 此外，DHCPv6服务器端将会记录该地址的分配情况（这也是为什么被称为有状态）。
  - 此方法配置较为复杂，且对DHCPv6服务器端的性能要求较高。
  - 有状态地址配置多用于公司内部有线终端的地址分配，便于对地址进行管理。
- 无状态地址配置，M=0，O=0：
  - 采用ICMPv6协议
    - 使能了ICMPv6 RA功能的路由器会周期性的通告该链路上的IPv6地址前缀。
    - 另一种情况，主机发送路由器查询（ICMPv6 RS）报文，路由器回复RA报文告知该链路IPv6地址前缀。
  - 主机根据路由器回应的RA报文，获得IPv6地址前缀信息，使用该地址前缀，加上本地产生的接口标识，形成单播IPv6地址。
  - 若主机还想获得其他配置信息，可以通过DHCPv6来获得除地址外的其他信息。当使用这种方式时，M=0，O=1。

    - 无状态地址配置的关键在于路由器完全不关心主机的状态如何，是否在线等，所以称为无状态。
    - 无状态地址配置多用于物联网等终端较多，且终端不需要除地址外其他参数的场景。
- DNS（Domain Name System，域名系统）：一种将用户容易记忆的域名映射为网络设备能够识别的IPv6地址的机制。
- NIS（Network Information System，网络信息服务）：用来管理电脑网络中所有与电脑系统管理相关配置文件的系统。
- SNTP（Simple Network Time Protocol，简单网络时钟协议 ）：由 NTP 改编而来，主要用来同步因特网中计算机的时钟。

DAD

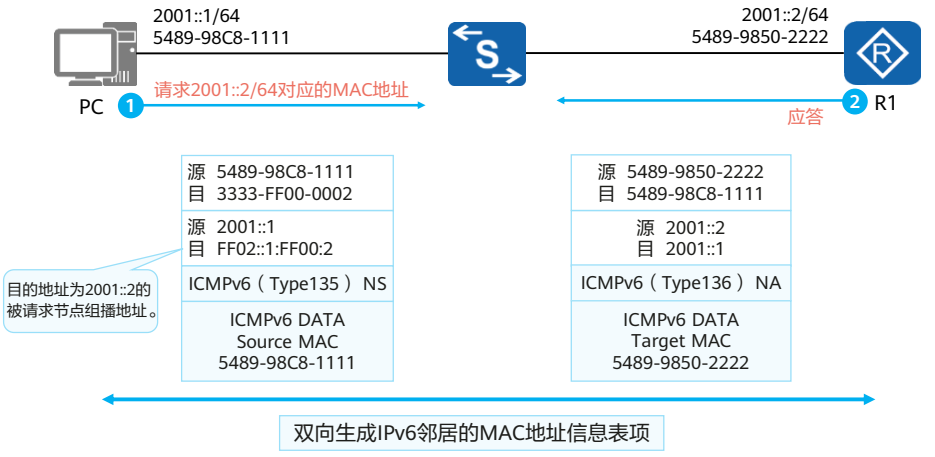
- 无论通过何种方式配置了IPv6单播地址，主机或路由器都会：
  - 通过ICMPv6报文进行DAD
  - 仅当DAD通过之后才会使用该单播地址



- 假设R1为已在线设备，IPv6地址为2001::FFFF/64。PC上线之后，也配置了相同的IPv6地址，在正式使用这个地址之前，PC会对此地址做DAD，过程如下:
  1. PC向链路上以组播的方式发送一个NS报文，该NS的源IPv6地址为“::”，目的IPv6地址为要进行DAD的2001::FFFF对应的被请求节点组播地址，也就是FF02::1:FF00:FFFF。这个NS里包含着要做DAD的目标地址2001::FFFF。
  2. 链路上的节点都会收到这个组播的NS报文，没有配置2001::FFFF的节点接口由于没有加入该地址对应的被请求节点组播组，因此在收到这个NS的时候默默丢弃。而R1在收到这个NS后，由于它的接口配置了2001::FFFF地址，因此接口会加入组播组FF02::1:FF00:FFFF，而此刻所收到的报文又是以该地址为目的地址，因此它会解析该报文，它发现对方进行DAD的目标地址与自己本地接口地址相同，于是立即回送一个NA报文，该报文的目的地址是FF02::1，也就是所有节点组播地址，同时在报文内写入目标地址2001::FFFF，以及自己接口的MAC地址。
  3. 当PC收到这个NA后，它就知道2001::FFFF在链路上已经有人在用了，因此将该地址标记为Duplicate（重复的），该地址将不能用于通信。若未收到NA报文，则PC判断这个IPv6地址可以用，DAD机制有点类似于IPv4中的免费ARP检测重复地址。

地址解析

- IPv6使用ICMPv6的NS和NA报文来取代ARP在IPv4中的地址解析功能。



- IPv6的地址解析不再使用ARP，也不再使用广播方式，而采用和DAD相同的NS和NA报文解析数据链路层地址。
- 假设PC想要解析R1的2001::2这个地址对应的MAC地址，详细过程如下：
  1. PC将发送一个NS报文达到这个目的。这个NS报文的源地址是2001::1，目的地址则是2001::2对应的被请求节点组播地址。
  2. R1接收此NS报文，根据报文内的源IPv6地址和源MAC，记录下PC这个邻居，同时根据自身的IPv6和MAC，回复单播NA报文。
  3. PC收到此NA报文之后，获取其中的源IPv6地址和源MAC。这样双方都可以建立一条关于对方的邻居信息表项。

目录

IPv6基本配置（1）

1. 使能IPv6

```
[Huawei] ipv6
```

使能设备转发IPv6单播报文，包括本地IPv6报文的发送与接收。

```
[Huawei-GigabitEthernet0/0/0] ipv6 enable
```

在接口视图下，在接口上使能该接口的IPv6功能。

2. 配置接口的链路本地地址

```
[Huawei-GigabitEthernet0/0/0] ipv6 address ipv6-address  link-local
```

```
[Huawei-GigabitEthernet0/0/0] ipv6 address auto link-local
```

在接口视图下，通过手工或者自动的方式，配置接口的链路本地地址。

3. 配置接口的全球单播地址

```
[Huawei-GigabitEthernet0/0/0] ipv6 address { ipv6-address prefix-length | ipv6-address/prefix-length }
```

```
[Huawei-GigabitEthernet0/0/0] ipv6 address auto { global | dhcp }
```

在接口视图下，通过手工或者自动（有状态或无状态）的方式，配置接口的全球单播地址。

IPv6基本配置 (2)

4. 配置IPv6静态路由

[Huawei] **ipv6 route-static** *dest-ipv6-address prefix-length* { *interface-type interface-number* [*nexthop-ipv6-address*] | *nexthop-ipv6-address* } [**preference** *preference*]

5. 查看接口的IPv6信息

[Huawei] **display ipv6 interface** [*interface-type interface-number* | **brief**]

6. 查看邻居表项信息

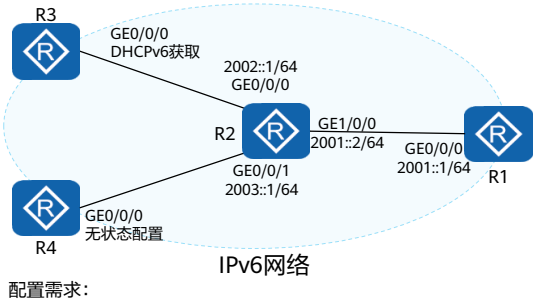
[Huawei] **display ipv6 neighbors**

7. 使能系统发布RA报文功能

[Huawei-GigabitEthernet0/0/0] **undo ipv6 nd ra halt**

默认情况下，华为路由器接口不发送ICMPv6 RA报文，则该接口所连链路上的其他设备无法进行无状态地址自动配置。若想进行IPv6无状态地址配置，需要手工开启发送RA报文。

案例：配置一个小型IPv6网络 (1)



- 配置需求：
  - R1和R2之间使用静态IPv6地址互联。
  - R2作为DHCPv6服务器给R3的GE0/0/0分配全球单播地址。
  - R4的GE0/0/0接口通过R2的RA进行无状态地址自动配置。
  - 配置静态路由，实现各设备之间互访。

1.在R1、R2、R3、R4全局和相关接口使能IPv6功能，同时自动生成链路本地地址（以R1配置为例）

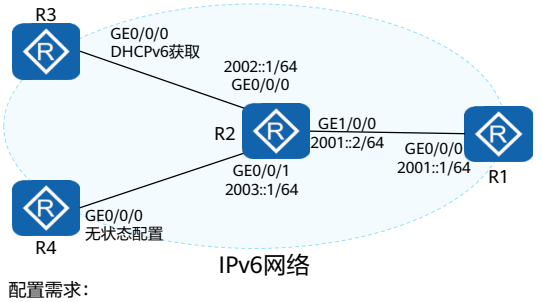
```
[R1]ipv6
[R1]interface GigabitEthernet 0/0/0
[R1-GigabitEthernet0/0/0]ipv6 enable
[R1-GigabitEthernet0/0/0]ipv6 address auto link-local
```

2.在R1、R2相应接口配置静态IPv6全球单播地址

```
[R1]interface GigabitEthernet 0/0/0
[R1-GigabitEthernet0/0/0]ipv6 address 2001::1 64
```

```
[R2]interface GigabitEthernet 1/0/0
[R2-GigabitEthernet1/0/0]ipv6 address 2001::2 64
[R2-GigabitEthernet1/0/0]interface GigabitEthernet 0/0/0
[R2-GigabitEthernet0/0/0]ipv6 address 2002::1 64
[R2-GigabitEthernet0/0/0]interface GigabitEthernet 0/0/1
[R2-GigabitEthernet0/0/1]ipv6 address 2003::1 64
```

案例：配置一个小型IPv6网络 (2)



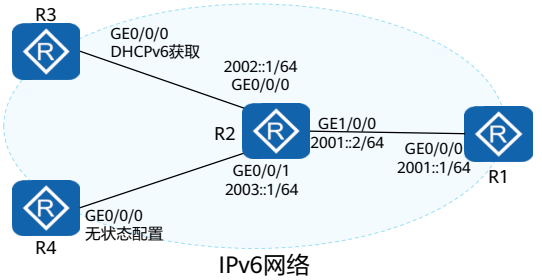
- 配置需求：
  - R1和R2之间使用静态IPv6地址互联。
  - R2作为DHCPv6服务器给R3的GE0/0/0分配全球单播地址。
  - R4的GE0/0/0接口通过R2的RA进行无状态地址自动配置。
  - 配置静态路由，实现各设备之间互访。

3.在R2上配置DHCPv6服务器功能；R3接口通过DHCPv6方式获取全球单播地址，并学习到IPv6网关R2的缺省路由

```
[R2]dhcp enable
[R2]dhcpv6 pool pool1
[R2-dhcpv6-pool-pool1]address prefix 2002::/64
[R2]interface GigabitEthernet 0/0/0
[R2-GigabitEthernet0/0/0]dhcpv6 server pool1
[R2-GigabitEthernet0/0/0] undo ipv6 nd ra halt
[R2-GigabitEthernet0/0/0] ipv6 nd autoconfig managed-address-flag
[R2-GigabitEthernet0/0/0] ipv6 nd autoconfig other-flag
[R2-GigabitEthernet0/0/0] quit
```

```
[R3]dhcp enable
[R3]interface GigabitEthernet 0/0/0
[R3-GigabitEthernet0/0/0]ipv6 address auto dhcp
[R3-GigabitEthernet0/0/0]ipv6 address auto global default
```

案例：配置一个小型IPv6网络 (3)



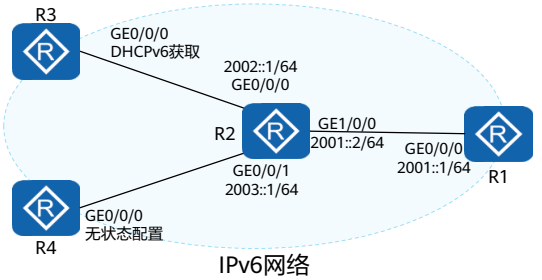
- 配置需求：
  - R1和R2之间使用静态IPv6地址互联。
  - R2作为DHCPv6服务器给R3的GE0/0/0分配全球单播地址。
  - R4的GE0/0/0接口通过R2的RA进行无状态地址自动配置。
  - 配置静态路由，实现各设备之间互访。

4.在R2使能发布RA报文的功能，R4通过无状态地址配置的方式获取地址

```
[R2]interface GigabitEthernet 0/0/1
[R2-GigabitEthernet0/0/1]undo ipv6 nd ra halt
```

```
[R4]interface GigabitEthernet 0/0/0
[R4-GigabitEthernet0/0/0]ipv6 address auto global
```

案例：配置一个小型IPv6网络 (4)



- 配置需求：
  - R1和R2之间使用静态IPv6地址互联。
  - R2作为DHCPv6服务器给R3的GE0/0/0分配全球单播地址。
  - R4的GE0/0/0接口通过R2的RA进行无状态地址自动配置。
  - 配置静态路由，实现各设备之间互访。

5.在R4上配置静态路由

```
[R4]ipv6 route-static 2001:: 64 2003::1
[R4]ipv6 route-static 2002:: 64 2003::1
```

6.在R1上配置聚合后的静态路由

```
[R1]ipv6 route-static  2002:: 15 2001::2
```

思考题

1. 2001:0DB8:0000:0000:032A:0000:0000:2D70，此IPv6地址压缩到最短是多少?
2. IPv6主机无状态地址自动配置的过程是什么?

HUAWEI

1. 2001:DB8::32A:0:0:2D70或2001:DB8:0:0:32A::2D70。
2. IPv6主机首先通过路由器接口发送的RA报文来获取地址前缀信息，之后通过向接口已有的48 bit MAC地址中插入16 bit的FFFE生成接口标识，在生成了IPv6地址后会通过重复地址检测来确认地址是否唯一。

本章总结

对比项	IPv6	IPv4
地址长度	128 bit	32 bit
报文格式	固定40 Byte的基本包头，变长的拓展字段来实现一些IPv6的特性	通过在基本头部上增加option字段的方式支持拓展特性
地址类型	单播、组播、任播	单播、组播、广播
地址配置	静态、DHCP、SLAAC	静态、DHCP
重复地址检测	通过ICMPv6实现	通过免费ARP实现
地址解析	通过ICMPv6实现	通过ARP实现


谢 谢
www.huawei.com

SDN与NFV概述

前言

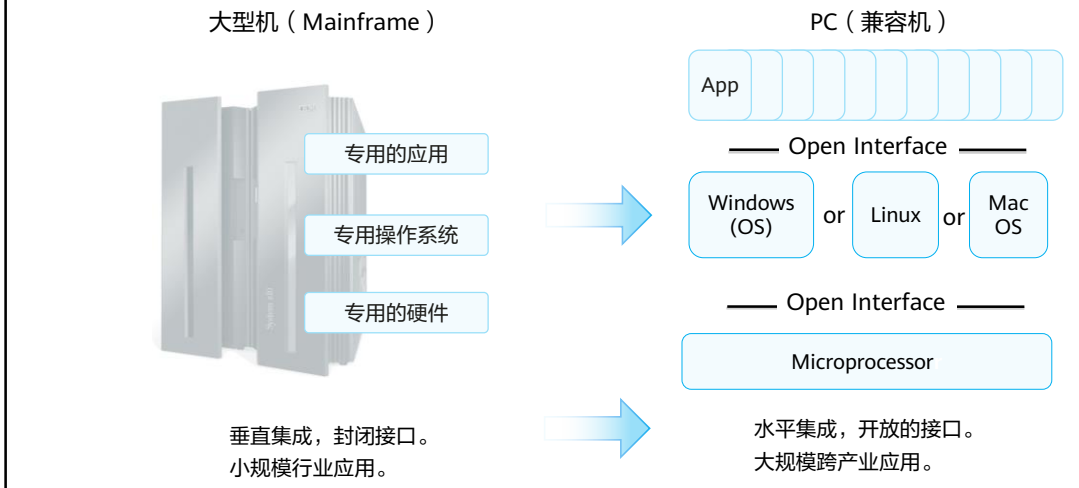
- 计算产业的开放生态带来了通用硬件、操作系统、虚拟化、中间件、云计算、软件应用等多领域的蓬勃发展。网络产业也在不断寻求变革与发展，其中SDN（Software Defined Networking，软件定义网络）与NFV（Network Functional Virtualization，网络功能虚拟化），是备受瞩目的两个概念。
- 本课程定位于帮助工程师了解SDN与NFV发展历史，并初步介绍华为SDN解决方案与NFV解决方案。

目标

- 学完本课程后，您将能够：
  - 描述SDN与NFV的发展历史
  - 了解OpenFlow的基本原理
  - 了解华为SDN解决方案
  - 了解标准NFV架构
  - 了解华为NFV解决方案

目录

计算机时代的演进



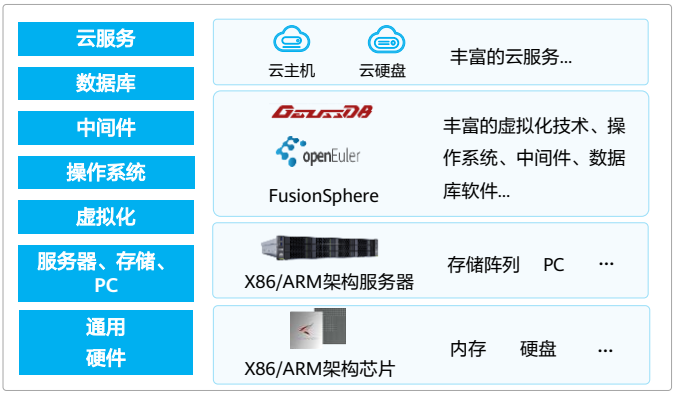
HUAWEI

- 1964年IBM公司花费50亿美金开发出了IBM SYSTEM/360大型机，开始了大型机的历史。大型机通常采用集中式体系架构，这种架构的优势之一是其出色的I/O处理能力，因而最适合处理大规模事务数据。与PC生态系统比较，大型机拥有专用的硬件、操作系统和应用。
- PC生态从硬件、操作系统到应用，经历了多次革新。每一次革新都带来了巨大变化和发展。支撑整个PC生态系统快速革新的三个因素是：
  - Hardware Substrate，硬件底层化。PC工业已经找到了一个简单、通用的硬件底层，x86指令集；
  - Software-definition，软件定义。上层应用程序和下层基础软件（OS，虚拟化）都得到了极大的创新；
  - Open-source，开源。Linux的蓬勃发展已经验证了开源文化和市集模式发展思路的正确性。成千上万的开发者可以快速制定标准，加速创新。

网络产业发展：来自IT行业的启示

- IT产业的变革引发了网络产业的思考。业界开始提出SDN（Software Defined Networking）的概念，并不断在其商用化进程上作出尝试。目的是希望网络变得更开放、灵活和简单。

计算产业开放，促进生态蓬勃发展

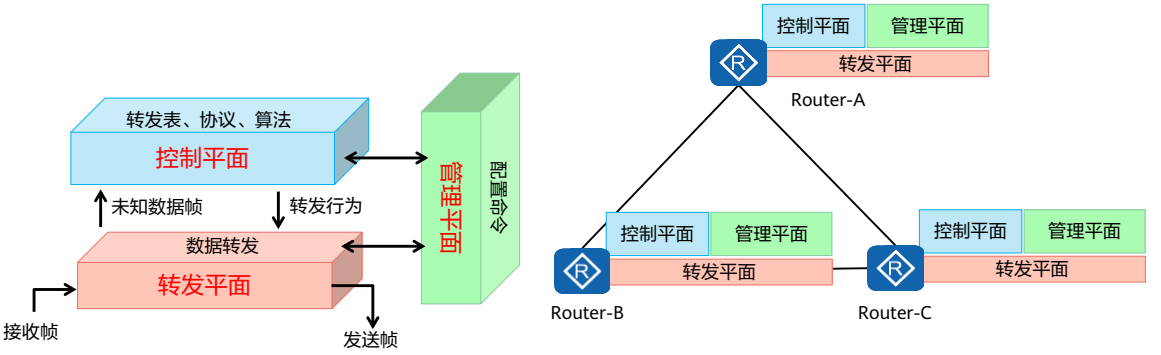


网络产业会如何变化?

- 网络产业是否参考计算产业，打造分层、开放的生态架构?

网络界的现状：经典IP网络 - 分布式网络

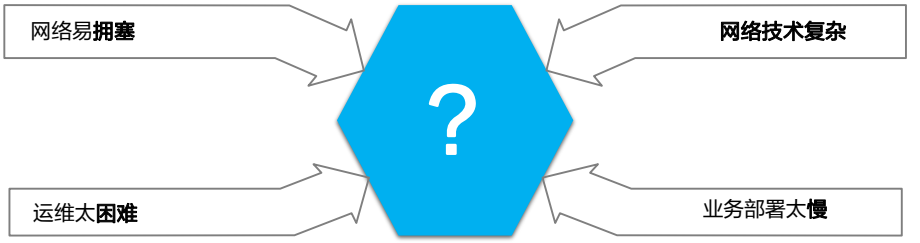
- 经典的IP网络是一个分布式的、对等控制的网络。每台网络设备存在独立的数据平台、控制平面和管理平面。设备的控制平面对等的交互路由协议，然后独立的生成数据平面指导报文转发。
- 经典IP网络的优势在于设备与协议解耦，厂家之间兼容性较好且故障场景下协议保证网络收敛。



HUAWEI

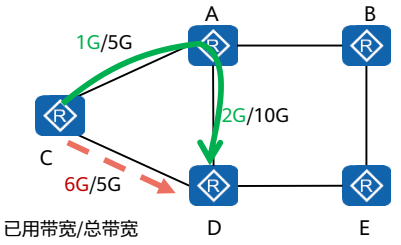
- 以交换机为例介绍转发平面、控制平面和管理平面：
  - 交换机转发平面：转发平面提供高速无阻塞数据通道，实现各个业务模块之间的业务交换功能。交换机的基本任务是处理和转发交换机各不同端口上各种类型的数据。L2/L3/ACL/QoS/组播/安全防护等各种具体的数据处理转发过程，都属于交换机转发平面的任务范畴。
  - 交换机控制平面：控制平面完成系统的协议处理、业务处理、路由运算、转发控制、业务调度、流量统计、系统安全等功能。交换机的控制平面用于控制和管理所有网络协议的运行。控制平面提供了数据平面数据处理转发前所必须的各种网络信息和转发查询表项。
  - 交换机管理平面：管理平面完成系统的运行状态监控、环境监控、日志和告警信息处理、系统加载、系统升级等功能。交换机的管理平面是提供给网络管理人员使用TELNET、WEB、SSH、SNMP、RMON等方式来管理设备，并支持、理解和执行管理人员对于网络设备各种网络协议的设置命令。管理平面必须预先设置好控制平面中各种协议的相关参数，并支持在必要时刻对控制平面的运行进行干预。
- 在华为产品具体的实现上，部分系列产品将功能组合区分为数据平面、管理平面和监控平面。

网络界的思考：经典网络面临的问题



网络易拥塞

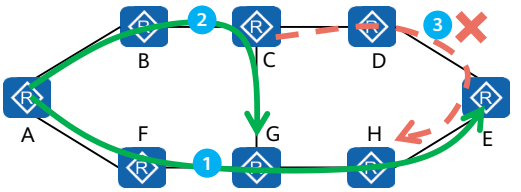
基于带宽固定选路的问题和解决思路



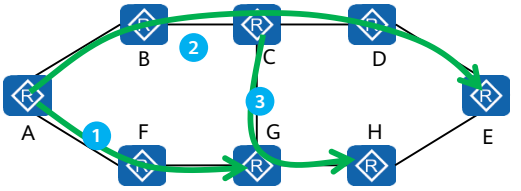
网络基于带宽计算转发路径。路由器C到路由器D的链路为最短转发路径。C-D的业务流量开始超过带宽出现丢包现象。虽然其他链路空闲，但是算法依然选择最短路径转发。如果可以全局考虑，此时最优的流量转发路径为C-A-D。

基于固定顺序建立隧道的问题和解决思路

顺序建立隧道：1. A-E；2. A-G；3. C-H。带宽不足隧道3建立失败。



全局计算，最优调整隧道路径：

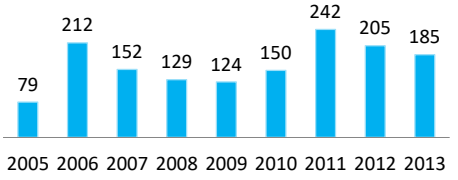


HUAWEI

网络技术太复杂

网络协议多：如果您准备成为一名网络技术专家，您需要阅读网络设备相关RFC 2500篇。如果一天阅读一篇，需要长达6年时间， 而这只是整个RFC的 1/3，其数量还在持续增加。

网络设备相关RFC增长数量

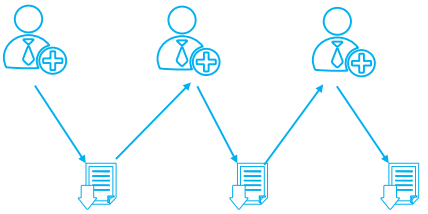


网络配置难：如果您准备成为某个设备商设备的百事通，您需要掌握的命令行超过10000条，而其数量还在增加。

网络故障定位、诊断困难

故障发现难

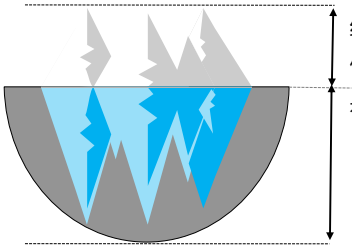
人工故障识别 人工抓包定位 人工问题诊断



- 传统运维网络故障依靠人工故障识别、人工定位和人工诊断。
- 超过85%的网络故障业务投诉后才发现。无法有效主动识别、分析问题。

故障定位难

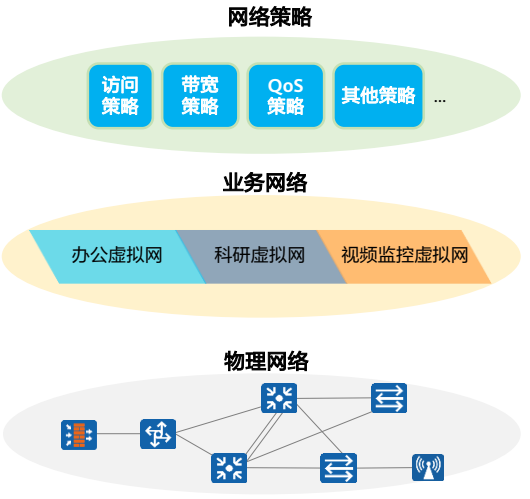
异常流占全网流**3.65%**



- 传统运维仅监控设备指标，存在指标正常，但用户体验差的情况。缺少用户和网络的关联分析。
- 数据中心网络统计，一个故障定位平均耗时76 min。

HUAWEI

网络业务的部署速度太慢



网络策略变更复杂、不灵活：
网络策略无法细化到用户。策略变更复杂，无法灵活调整。

基于IP、位置固定、命令行配置

新业务部署周期长：
新业务部署涉及端到端设备配置修改。

命令行端到端配置

物理网络部署效率低：
物理网络无零配置部署能力。

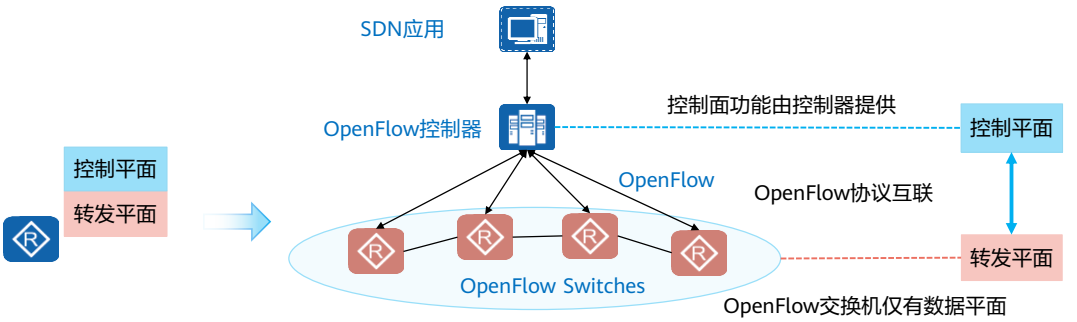
设备逐台命令行配置

HUAWEI

- 网络业务部署的愿景是：网络策略实现业务随行，与物理位置无关；新业务实现快速部署；物理网络支持零配置部署，设备即插即用。

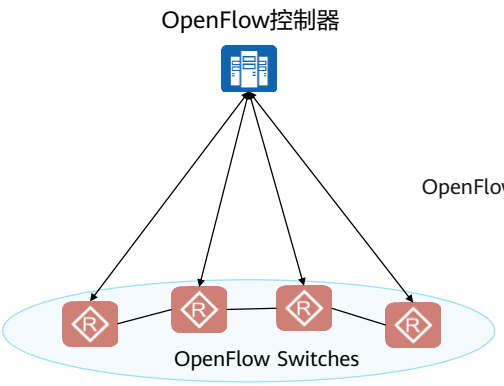
SDN的起源

- SDN（Software Defined Networking）即软件定义网络。是由斯坦福大学Clean Slate研究组提出的一种新型网络创新架构。 其核心理念通过将网络设备控制平面与数据平面分离，从而实现了网络控制平面的集中控制，为网络应用的创新提供了良好的支撑。
- SDN起源提出了三个特征，“转控分离”、“集中控制”和“开放可编程接口”。



OpenFlow基本概念

- OpenFlow是控制器与交换机之间的一种南向接口协议。它定义了三种类型的消息，Controller-to-Switch、Asynchronous 和 Symmetric。每一种消息又包含了更多的子类型。



OpenFlow

Controller-to-Switch

该消息由Controller发送。用于管理Switch和查询Switch的相关信息。

Asynchronous

该消息由Switch发起。当Switch状态发生改变时，发送该消息告诉Controller状态变化。

Symmetric

该消息没有固定发起方，可由Switch或者Controller发起。例如Hello、Echo、Error等。

HUAWEI

- Controller-to-Switch子类型：
  - Features消息：在SSL/TCP会话建立后，Controller给Switch发送Features请求Switch的相关信息。Switch必须应答自己支持的功能，包括接口名、接口MAC地址、接口支持的速率等等基本信息。
  - Configuration消息：Controller可以设置或查询Switch的状态。
  - Modify-State消息：Controller发送该消息给Switch，来管理Switch的状态，即增加/删除、更改流表，并设置Switch的端口属性。
  - Read-State消息：Controller用该消息收集Switch上的统计信息。
  - Send-Packet消息：Controller发送该消息到Switch的特定端口。
- Asynchronous子类型：
  - Packet-in消息：当Flow Table中没有匹配的表项或者匹配“send to Controller”，Switch将给Controller发送packet-in消息。
  - Packet-out消息：从控制器回复的消息。
  - Flow-Removed消息：当给Switch增加一条表项时，会设定超时周期。当时间超时后，该条目就会被删除。这时Switch就会给Controller发送Flow-Removed消息；当流表中有条目要删除时，Switch也会给Controller发送该消息。
  - Port-status消息：当数据路径接口被添加、删除、修改的时候，此消息用于通知控制器。

- Symmetric子类型：
    - Hello消息：当一个OpenFlow连接建立时，Controller和Switch都会立刻向对端发送OFPT_HELLO消息，该消息中的version域填充发送方支持的OpenFlow协议最高的版本号；接收方收到该消息后，接收方会计算协议版本号，即在发送方和接收方的版本号中选择一个较小的；如果接收方支持该版本，则继续处理连接，连接成功；否则，接收者回复一个OFPT_ERROR消息，类型域中填充ofp_error_type.OFPET_HELLO_FAILED
    - Echo消息： Switch和Controller任何一方都可以发起Echo request消息，但收到的一方必须回应Echo reply消息。这个消息可以来测量latency、Controller-Switch之间的连接性，即心跳消息；
    - Error消息：当交换机需要通知控制器发生问题或错误时，Switch给Controller 发送Error消息。
- OpenFlow协议仍在持续更新。更多更全的消息类型请参考ONF最新发布《 OpenFlow Switch Specification 》标准。

Flow Table简介

- OpenFlow交换机基于流表（Flow Table）转发报文。
- 每个流表项由匹配字段、优先级、计数器、指令、超时、Cookie、Flags这七部分组成。其中关于转发的关键的两个内容是**匹配字段**和**指令**。
  - 匹配字段是匹配规则，支持自定义。
  - 指令是用来描述匹配后的处理方式。

匹配域	优先级	计数	指令	Timeout	Cookie	Flags

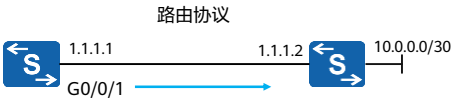
流表字段支持自定义，例如本例匹配项字段：

Ingress Port	Ether Source	Ether Dst	Ether Type	VLAN ID	VLAN Priority	IP Src	IP Dst	TCP Src Port	TCP Dst Port
3	MAC1	MAC2	0x8100	10	7	IP1	IP2	5321	8080

- Match Fields：流表项匹配项（OpenFlow 1.5.1版本支持45个可选匹配项），可以匹配入接口、物理入接口，流表间数据，二层报文头，三层报文头，四层端口号等报文字段等。
- Priority：流表项优先级，定义流表项之间的匹配顺序，优先级高的先匹配。
- Counters：流表项统计计数，统计有多少个报文和字节匹配到该流表项。
- Instructions：流表项动作指令集，定义匹配到该流表项的报文需要进行的处理。当报文匹配流表项时，每个流表项包含的指令集就会执行。这些指令会影响到报文、动作集以及管道流程。
- Timeouts：流表项的超时时间，包括了Idle Time和Hard Time。
  - Idle Time：在Idle Time时间超时后如果没有报文匹配到该流表项，则此流表项被删除。
  - Hard Time：在Hard Time时间超时后，无论是否有报文匹配到该流表项，此流表项都会被删除。
- Cookie：Controller下发的流表项的标识。
- Flags：该字段改变流条目的管理方式。

转发方式对比

经典路由协议基于路由表转发

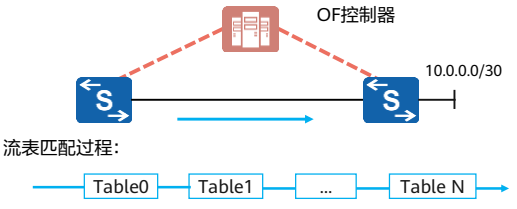


路由表

目的网络	来源	下一跳	出接口
10.0.0.0/30	OSPF	1.1.1.2	G0/0/1

- 经典的网络转发方式是网络设备通过查询路由表指导流量转发。
- 路由表的条目由网络设备之间运行路由协议而计算生成。
- 路由表是定长的。路由表通过最长匹配原则执行报文转发。一台网络设备只有一张路由表。

OpenFlow基于流表转发



流表

匹配域	优先级	计数	指令	Timeout	Cookie

- OpenFlow是一个网络协议。运行OpenFlow的交换机通过查询流表指导流量转发。
- 流表一般是由OF控制器统一计算，然后下发到交换机。
- 流表是变长的，拥有丰富的匹配规则和转发规则。一台网络设备有多张流表。

HUAWEI

- 流表的匹配原则是对于存在的“table0-table255”，优先从table0开始匹配。同一table内部按照优先级匹配，优先级高优先匹配。
- 当前OpenFlow的主流应用是用于数据中心的软件交换机，例如OVS、CE1800V等，而不是实现硬件交换机的转控分离。

SDN的本质诉求

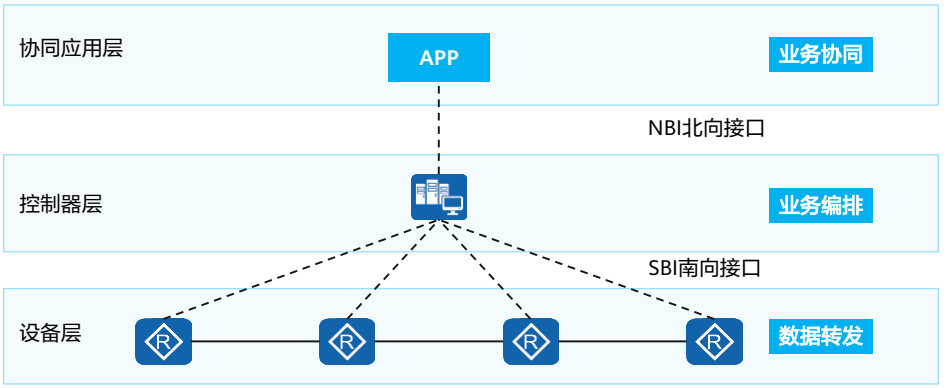
- SDN的本质诉求是让网络更加开放、灵活和简单。它的实现方式是为网络构建一个集中的大脑，通过全局视图集中控制，实现或业务快速部署、或流量调优、或网络业务开放等目标。
- SDN的价值是：
  - 集中管理，简化网络管理与运维；
  - 屏蔽技术细节，降低网络复杂度，降低运维成本；
  - 自动化调优，提高网络利用率；
  - 快速业务部署，缩短业务上线时间；
  - 网络开放，支撑开放可编程的第三方应用。

SDN带来了网络架构的变革。

- SDN是一个更为广泛的概念而不局限于OpenFlow。转控分离是实现SDN的一种方法而不是本质。

SDN网络架构

- SDN网络架构分为协同应用层、控制器层和设备层。不同层次之间通过开放接口连接。以控制器层为主要视角，区分面向设备层的南向接口和面向协同应用层的北向接口。OpenFlow属于南向接口协议的一种。

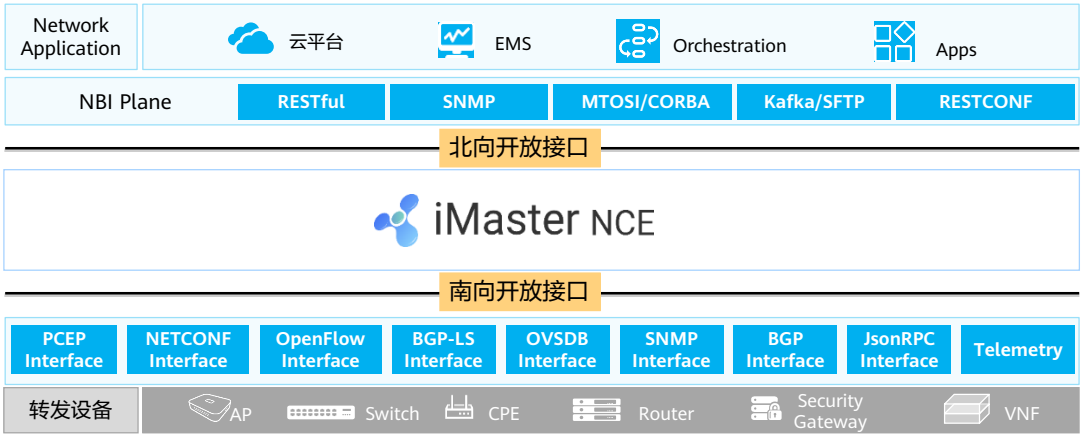


HUAWEI

- 协同应用层：主要完成用户意图的各种上层应用，典型的协同层应用包括OSS、OpenStack等。OSS可以负责整网的业务协同，OpenStack云平台一般用于数据中心负责网络、计算、存储的业务协同。还有其他的协同层应用，比如用户希望部署一个安全APP，这个安全APP不关心设备具体部署位置，只是调用了控制器的北向接口，例如Block（Source IP，DestIP），然后控制器会给各网络设备下发指令。这个指令根据南向协议不同而不同。
- 控制器层：控制器层的实体就是SDN控制器，是SDN网络架构下最核心的部分。控制层是SDN系统的大脑，其核心功能是实现网络业务编排。
- 设备层：网络设备接收控制器指令，执行设备转发。
- NBI北向接口：北向接口为控制器对接协同应用层的接口，主要为RESTful。
- SBI南向接口：南向接口为控制器与设备交互的协议，包括NETCONF、SNMP、OpenFlow、OVSDB等。

华为SDN网络架构

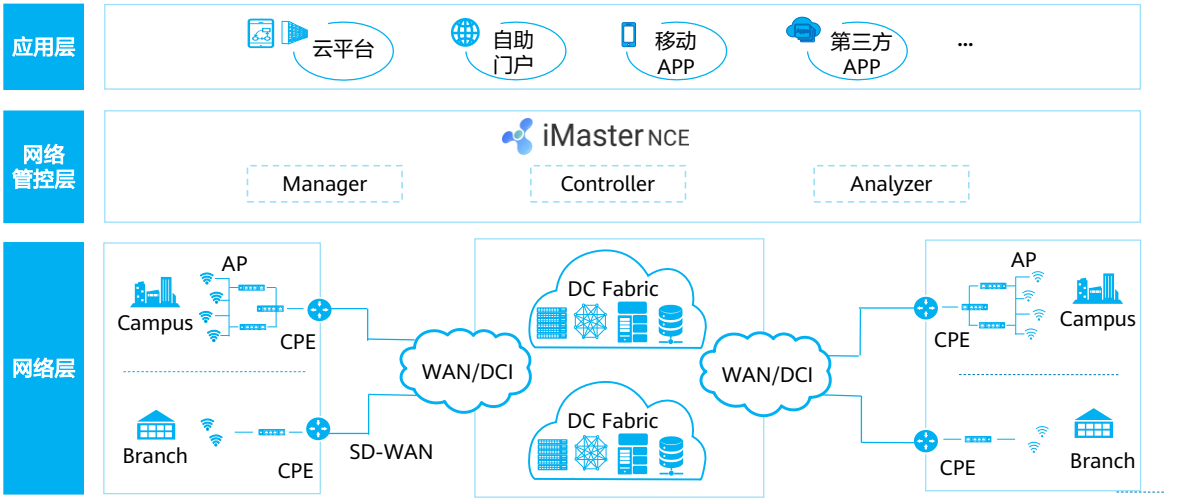
- 华为SDN网络架构支持丰富的南北向接口，包括OpenFlow、OVSDB、NETCONF、PCEP、 RESTful 、SNMP、BGP、JsonRPC、RESTCONF等。



HUAWEI

- 云平台：云数据中心内资源管理平台。云平台包含对网络资源、计算资源和存储资源的管理。OpenStack是最主流的开源云平台。
- EMS（Element Management System，网元管理系统）是管理特定类型的一个或多个电信NE（Network Element，网络单元）的系统。
- Orchestration（容器编排）：容器编排工具也可以包含网络业务编排功能。Kubernetes是主流的工具。
- MTOSI/CORBA用于对接BSS/OSS。Kafka/SFTP可用于对接大数据平台。

华为SDN解决方案 - 管、控、析构建智简网络



HUAWEI

什么是iMaster NCE？

- iMaster NCE，自动驾驶网络管理与控制系统，是华为集管理、控制、分析和AI智能功能于一体的网络自动化与智能化平台。

HUAWEI

- iMaster NCE能做什么？它有效连接了物理网络与商业意图。南向实现全局网络的集中管理、控制和分析。面向商业和业务意图使能资源云化、全生命周期网络自动化，以及数据分析驱动的智能闭环。北向提供开放网络API与IT快速集成。
- iMaster NCE用在哪里？可以在企业领域数据中心网络（DCN）、企业园区（Campus）、企业分支互联（SD-WAN）等场景，让企业网络更加简单、智慧、 开放和安全，加速企业的业务转型和创新。

iMaster NCE全新启航

iMaster NCE

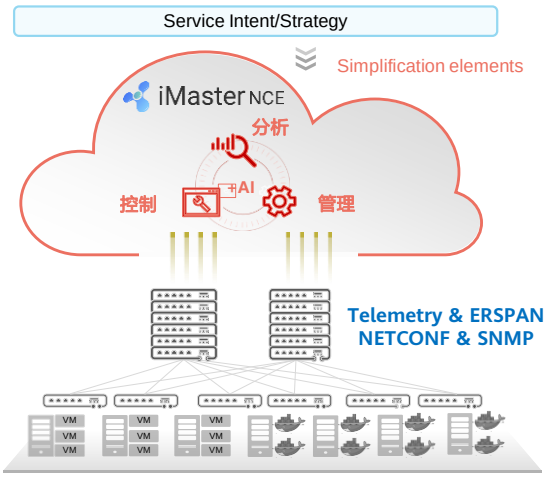
- 数据中心 iMaster NCE-Fabric *
- 企业园区 iMaster NCE-Campus *
- SD-WAN iMaster NCE-WAN
- 广域IP iMaster NCE-IP
- 广域传输 iMaster NCE-T

* 本课程介绍

HUAWEI

华为数据中心CloudFabric自动驾驶解决方案

- 基于iMaster NCE-Fabric，为数据中心网络提供从规划-建设-运维-调优全生命周期服务。



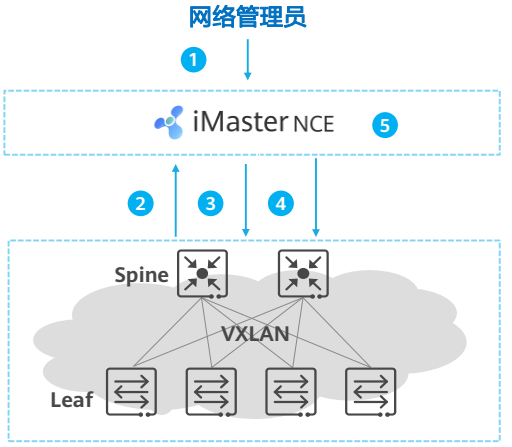
规建一体：
- 规划工具对接NCE，实现规划建设一体化。
- ZTP（Zero Touch Provisioning，零配置开局）

极简部署：
- 业务意图自理解和转换部署。
- 网络变更仿真评估，杜绝人为错误。

智能运维：
- 基于知识图谱和专家规则的快速故障发现定位
- 基于专家规则和仿真分析的快速故障恢复

实时调优：
- 面向AI-Fabric的流量本地推理，在线模型训练调优
- 用户行为预测、资源调优建议

关键特性：极简ZTP部署



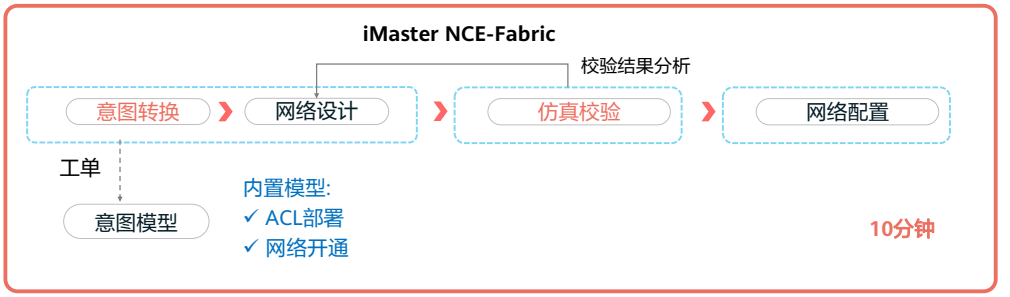
ZTP部署流程：

1. 网络管理员点击启动ZTP任务。
2. 设备自动获取IP地址访问控制器。
3. 控制器判断设备角色（Spine or Leaf），对上线设备下发管理IP、SNMP、NETCONF等配置，并通过管理IP纳管设备。
4. 控制器全局下发互联配置及OSPF、BGP等配置。
5. 设备上线成功，管理员NCE查看全网信息。

注：Spine-Leaf是数据中心网络架构

关键特性：网络意图自理解，业务快速部署

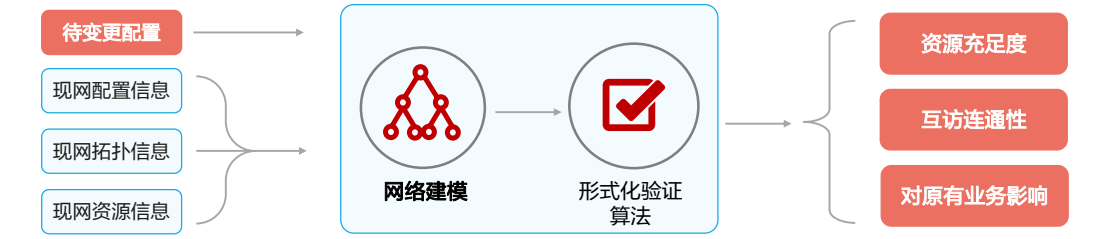
网络设计	配置下发	业务验证
2~3天	10分钟	1~2天



华为iMaster NCE-Fabric支持虚拟化、云计算和容器网络的自动化快速部署。

- iMaster NCE-Fabric支持对接用户IT系统，为用户意图匹配意图模型，通过NETCONF下发配置到设备上实现业务快速部署。
- iMaster NCE-Fabric支持对接主流云平台（OpenStack）、虚拟化平台（vCenter/SystemCenter）和容器编排平台（Kubernetes）。

关键特性：网络变更仿真，预判变更风险



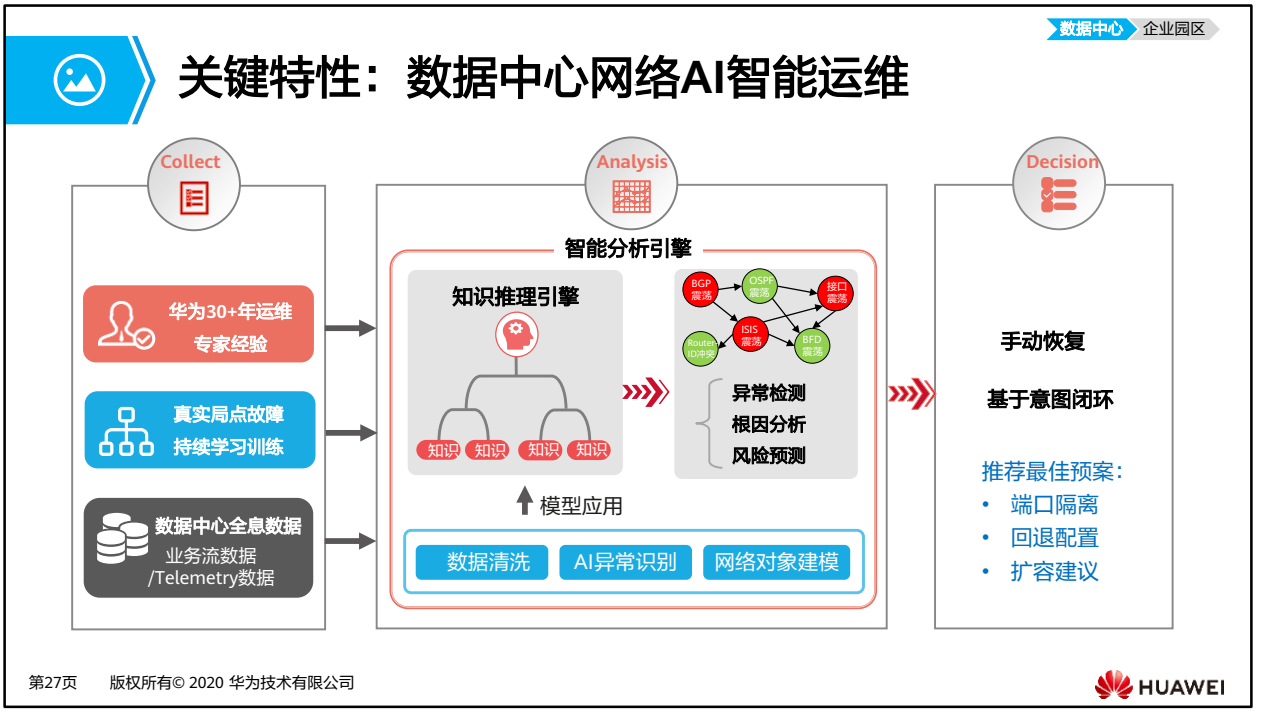
数据收集/上传

建模求解

- 建立物理/逻辑/应用网络模型
- 通过形式化验证算法求解

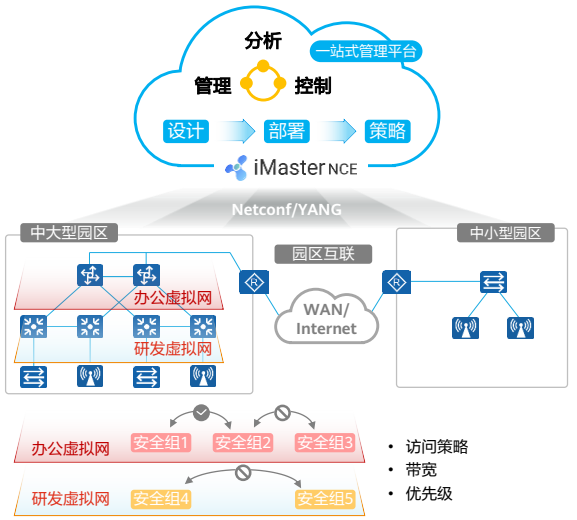
校验结果

- 校验现网资源是否足够、连通性等
- 变更对原有业务影响分析和呈现



- 数据中心AI智能运维能力由iMaster NCE-FabricInsight提供。

华为园区网络CloudCampus自动驾驶解决方案



网络开通”快”，部署效率提升600%

- 设备即插即用：设备极简开局，场景导航，模板配置
- 网络极简部署：网络资源池化，一网多用，业务自动化发放

业务发放”快”，用户体验提升100%

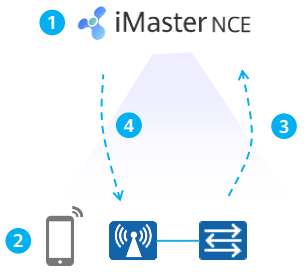
- 业务随行：图形化策略配置，用户随时随地接入，漫游权限不变，体验不变
- 终端智能识别：终端接入防仿冒，终端智能识别准确率95+%
- 智能HQos：基于应用调度和整形，带宽精细化管理，保证关键用户业务体验

智能运维“快”，整网性能提升50%+

- 实时体验可视：基于Telemetry的每时刻、每用户、每区域的网络体验可视
- 精准故障分析：主动识别85%的典型网络问题并给出建议，实时数据对比分析故障预测
- 智能网络调优：基于历史数据的无线网络预测性调优，整网性能提升50%+（来源：Tolly认证）

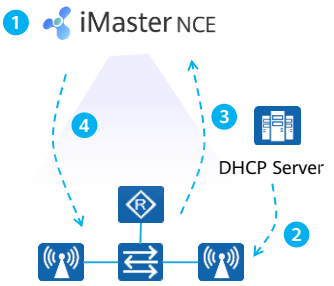
关键特性：设备即插即用

APP扫码开局



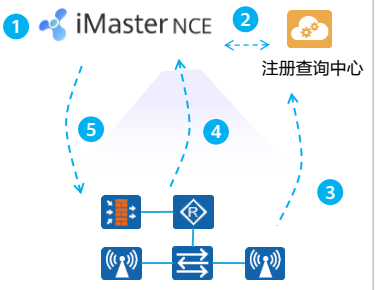
1. 预配置
2. APP扫码部署
3. 设备自动注册上线
4. 配置自动化下发

DHCP开局



1. 预配置
2. 通过DHCP server获取注册信息
3. 设备自动注册上线
4. 配置自动化下发

注册查询中心方式

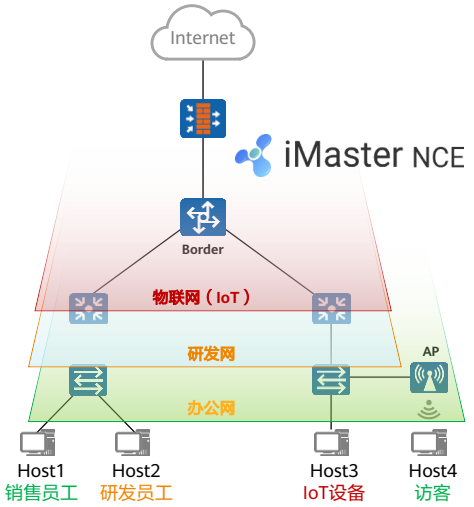


1. 预配置
2. 同步信息
3. 通过注册查询中心获取注册信息
4. 设备自动注册上线
5. 配置自动化下发

HUAWEI

- 设备即插即用包括但不限于APP扫码开局、DHCP开局和注册查询中心。
- 注册中心：华为设备注册查询中心，简称注册中心，是华为云管理网络解决方案的主要部件之一，用于设备的管理模式和注册归属查询。设备根据查询结果确定是否切换到云管理模式，需要注册到哪个云管理平台。
  以AP为例，对于华为支持云管理特性的设备均会预置华为设备注册中心的URL（register.naas.huawei.com）和端口号（10020）。

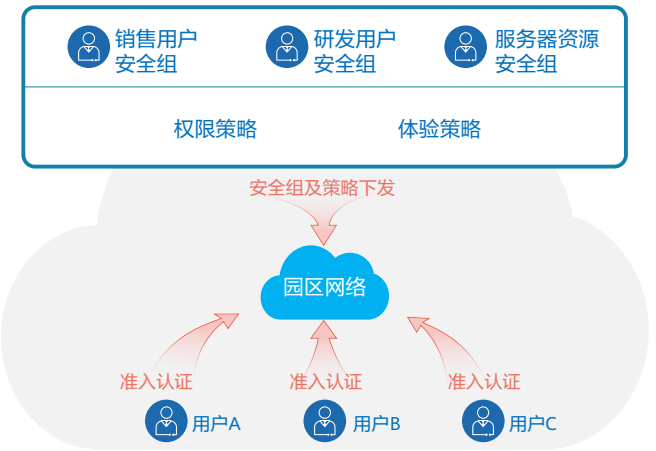
关键特性：构建一网多用的虚拟化园区



- 通过引入虚拟化技术，在园区网络中，基于一张物理网络创建多张虚拟网络（VN，Virtual Network）。不同的虚拟网络应用于不同的业务，例如办公、研发或物联网等。
- 通过iMaster NCE实现全网设备集中管理，管理员通过图形化界面实现网络配置。
- iMaster NCE将管理员的网络业务配置意图“翻译”成设备命令，通过NETCONF协议将配置下发到各台设备，实现网络的自动驾驶。

关键特性：业务随行，基于安全组的策略管理

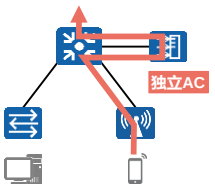
- 业务随行：不管用户身处何地，使用哪个IP地址，都保证该用户拥有相同的网络权限和一致的用户策略。



1. 引入安全组。安全组即拥有相同网络访问策略的一组用户。
2. 定义基于安全组的权限控制策略、用户体验策略，将策略下发到网络设备。
3. 用户执行准入认证后，获得授权的安全组。
4. 用户的流量进入网络后，网络设备根据流量所述的源、目的安全组执行策略。

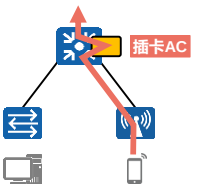
关键特性：有线与无线融合

无线网络构建方式1：独立AC



- 独立AC存在无线流量瓶颈并增加故障节点；
- 有线及无线管理上独立；
- 有线、无线认证点分离。

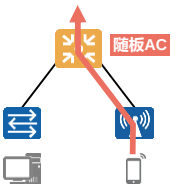
无线网络构建方式2：插卡AC



- AC作为交换机的一块插卡，安装在交换机上；
- 仅仅是硬件层面的融合。

有线及无线认证点分离、策略控制分散、流量转发分离、故障排除困难、管理困难

有线无线融合（随板AC，Native AC）



交换机融合AC功能，无线流量转发无瓶颈，并且减少故障点，有线无线集中管理：

- 有线及无线业务统一管理、融合转发
- 有线及无线用户融合管理、网关融合
- 有线及无线认证点融合
- 有线及无线统一策略执行

关键特性：终端智能识别，安全接入

需求&挑战

某高校

50+类
智能终端

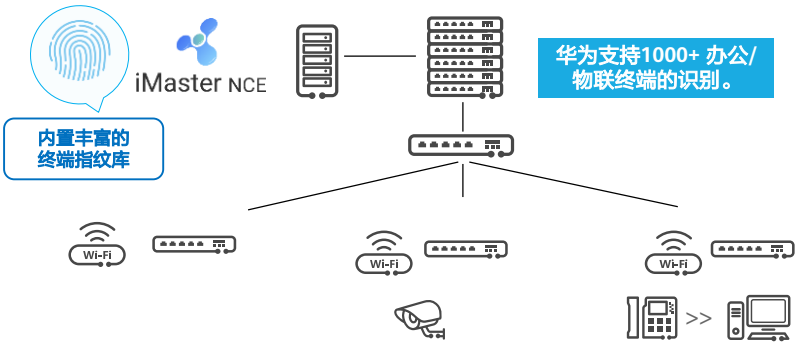
终端由学院采集：
MAC采集难、易错

某企业

100+/天
认证报障

接入仿冒难定位

>>



基于终端类型的

自动认证

识别为打印机

- 自动(MAC)认证，免MAC录入

基于终端类型的

自动授权

识别为摄像头

- 自动加入“视频监控”组
- 设置为“VIP”用户

基于终端类型的

仿冒检测

先识别为IP电话，后识别为PC

- 上报终端仿冒告警

关键特性：园区网络AI智能运维

AS-IS：以设备为中心的网络管理

传统网管

网络管理系统
- 拓扑管理
- 性能管理
- 告警管理
- 配置管理

SNMP
分钟级网络数据采集

- **以设备为中心**，无法感知用户体验
- **“救火式”响应**，无法识别潜在故障
- 依靠专业工程师**现场定位故障**

TO-BE：以用户体验为中心的AI智能运维

iMaster NCE CampusInsight

网络智能分析器
- 体验可视化管理
- 用户旅程回放
- 潜在故障识别
- 故障根因定位
- 网络预测性调优

Telemetry
秒级网络数据采集

体验可视化：基于Telemetry秒级数据采集，每用户每应用每时刻体验可视

分钟级潜在故障识别和根因定位
- 基于动态基线、大数据关联等识别潜在故障
- KPI关联分析和协议回放，精准定位问题根因

网络预测性调优：通过AI 智能分析AP的负载趋势，完成无线网络的预测性调优闭环

利用算法提升效率，通过场景化的持续学习和专家经验，智能运维将运维人员从复杂的告警和噪声解放出来，使运维更加自动化和智能化。

目录

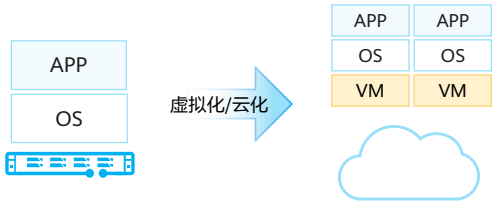
HUAWEI

NFV的背景：同样来自IT行业变革的启示

- 来自IT界的启示，给网络产业带来了网络架构和设备架构两个层面的思考。网络架构层面引入对SDN控制器的思考，设备架构层面引入对设备部署形态的思考。

IT产业的变革

- 近年虚拟化和云计算等IT技术蓬勃发展，传统部署于硬件的应用逐渐云化。应用以软件的方式部署于私有云、公有云或者混合云上。



网络?

- 网络产业思考：能否以软件化的方式部署网络应用呢?
- 这些思考引发了NFV（Network Functions Virtualization，网络功能虚拟化）。

HUAWEI

- 网络功能虚拟化被称为NFV（Network Functions Virtualization），而虚拟化之后的网络功能被称为VNF（Virtualized Network Function）。当我们谈“VNF”时，我们指运营商IMS、CPE这些传统网元在虚拟化之后的实现。在硬件通用化后，传统的网元不再是嵌入式的软硬结合的产品，而是以纯软件的方式安装在通用硬件（即NFVI）上。

NFV的起源

- 2012年10月，13家Top运营商（AT&T、Verizon、VDF、DT、T-Mobile、BT、Telefonica等）在SDN和Open Flow世界大会上发布NFV （Network Functions Virtualization）第一版白皮书，同时成立了ISG（Industry Specification Group）来推动网络虚拟化的需求定义和系统架构制定。
- 2013年，ETSI下NFV ISG（行业规范工作组）进行第一阶段研究，已完成相关标准制定。主要定义网络功能虚拟化的需求和架构，并梳理不同接口的标准化进程。

  HUAWEI

- 2015年，NFV研究进入第二阶段。其主要研究目标是建设一个可互操作的NFV生态，推动更广泛的行业参与，并且确保满足阶段一中定义的需求。同时明确NFV与SDN等相关标准、开源项目的协作关系等。NFV阶段二主要分为5个工作组：IFA（架构与接口）、 EVE（生态圈）、 REL（可靠性）、 SEC（安全）、TST（测试、执行、开源）。各工作组主要讨论交付件文档框架和交付计划。
- ETSI NFV标准组织与Linux基金会合作，启动开源项目OPNFV（NFV开源项目，提供一个集成、开放的参考平台），汇聚业界的优势资源，积极打造NFV产业生态。2015年OPNFV发布了首个版本，进一步促进NFV商用部署。
- NFV相关的标准组织主要有：
  - ETSI NFV ISG：制定NFV的需求和功能框架。
  - 3GPP SA5工作组：重点关注3GPP网元的虚拟化管理（MANO相关）的技术标准和规范。
  - OPNFV：加速NFV市场化节奏的开源平台项目。

NFV的价值

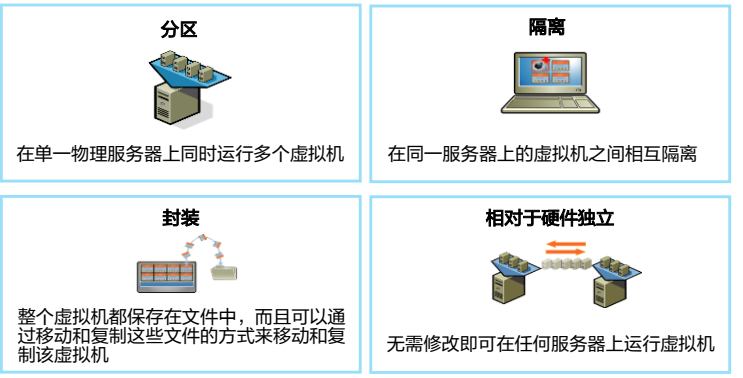
- NFV（Network Functions Virtualization ）是运营商为了解决电信网络硬件繁多、部署运维复杂、业务创新困难等问题而提出的。NFV在重构电信网络的同时，给运营商带来的价值如下：
  - 缩短业务上线时间
  - 降低建网成本
  - 提升网络运维效率
  - 构建开放的生态系统

HUAWEI

- 缩短业务上线时间：在NFV架构的网络中，增加新的业务节点变得异常简单。不再需要复杂的工勘、硬件安装过程。业务部署只需申请虚拟化资源（计算/存储/网络等），加载软件即可，网络部署变得更加简单。同时，如果需要更新业务逻辑，也只需要更新软件或加载新业务模块，完成业务编排即可，业务创新变得更加简单。
- 降低建网成本：首先，虚拟化后的网元能够合并到通用设备（COTS）中，获取规模经济效应。其次，提升网络资源利用率和能效，降低整网成本。NFV采用云计算技术，利用通用化硬件构建统一的资源池，根据业务的实际需要动态按需分配资源，实现资源共享，提高资源使用效率。如通过自动扩缩容解决业务潮汐效应下资源利用问题。
- 提升网络运维效率：自动化集中式管理提升运营效率，降低运维成本。例如数据中心的硬件单元的集中管理的自动化，基于MANO的应用生命周期管理的自动化，基于NFV/SDN协同的网络自动化。
- 构建开放的生态系统 ： 传统电信网络专有软硬件的模式，决定了它是一个封闭系统。NFV架构下的电信网络，基于标准的硬件平台和虚拟化的软件架构，更易于开放平台和开放接口，引入第三方开发者，使得运营商可以共同和第三方合作伙伴共建开放的生态系统。

NFV关键技术：虚拟化

- 在NFV的道路上，虚拟化是基础，云化是关键。
- 传统电信网络中，各个网元都是由专用硬件实现，成本高、运维难。虚拟化具有分区、隔离、封装和相对于硬件独立的特征，能够很好匹配NFV的需求。运营商引入此模式，将网元软件化，运行在通用基础设施上。



HUAWEI

- 传统电信网络中，各个网元都是由专用硬件实现的。这种方式的问题在于，一方面搭建网络需要进行大量不同硬件的互通测试及安装配置，费时费力。另一方面，业务创新需要依赖于硬件厂商的实现，通常耗时较长，难以满足运营商对业务创新的需求。在这种背景下，运营商希望引入虚拟化的的模式，将网元软件化，运行在通用基础设施上（包括通用的服务器、存储、交换机等）。
- 使用通用硬件，首先运营商可以减少采购专用硬件的成本。其次，业务软件可以快速的进行迭代开发，也使得运营商可以快速进行业务创新、提升自身的竞争力。最后，这也赋予了运营商进入云计算市场的能力。

NFV关键技术：云化

- 美国国家标准与技术研究院（NIST）定义：云计算是一种模型，它可以实现随时随地，便捷地，随需应变地从可配置计算资源共享池中获取所需的资源（例如，网络、服务器、存储、应用、及服务），资源能够快速供应并释放，使管理资源的工作量和与服务提供商的交互减小到最低限度。
- 云计算拥有诸多好处。运营商网络中网络功能的云化更多的是利用了资源池化和快速弹性伸缩两个特征。

云计算的特征

HUAWEI

- 美国国家标准和技术研究院的定义，云计算服务应该具备以下几条特征：
  - 按需自助服务（On-demand Self-service）：云计算实现了IT资源的按需自助服务，不需要IT管理员的介入即可申请和释放资源。
  - 广泛网络接入（Broad Network Access）：有网络即可随时、随地的使用。
  - 资源池化（Resource Pooling）：资源池中的资源包括网络、服务器、存储等资源，提供给用户使用。
  - 快速弹性伸缩（Rapid Elasticity）：资源能够快速的供应和释放。申请即可使用，释放立即回收资源。
  - 可计量服务（Measured Service）：计费功能。计费依据就是所使用的资源可计量。例如按使用小时为时间单位，以服务器CPU个数、占用存储的空间、网络的带宽等综合计费。

NFV架构简介

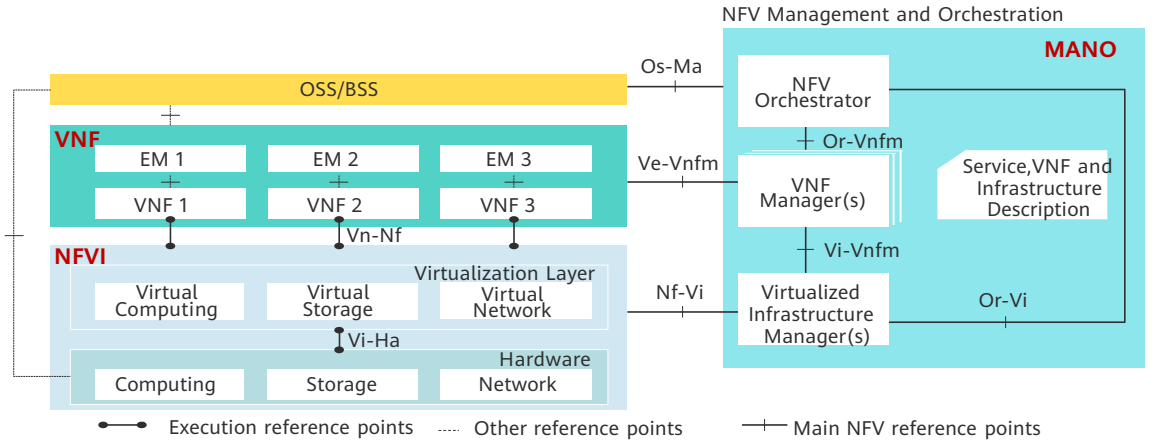
- NFV架构分**NFVI**（Network Functions Virtualization Infrastructure，基础设施层）、**VNF**（Virtualized Network Function，虚拟化网络功能层）和**MANO**（Management and Orchestration，管理编排域），同还要支持现有的BSS/OSS（Business support system/ Operation support system）。

HUAWEI

- NFV架构每一层都可以由不同的厂商提供解决方案，在提高系统开发性的同时增加了系统集成的复杂度。
- NFV价值是通过设备归一和软硬件解耦实现资源的高效利用，可以降低运营商TCO，缩短业务上线时间，打造开放的产业生态。
- NFVI包含硬件层和虚拟化层，业界也有说法称作COTS和CloudOS：
  - COTS（commercial off-the-shelf，商用现货），即通用硬件，强调了易获得性和通用性。例如Huawei FusionServer系列硬件服务器。
  - CloudOS：设备云化的平台软件，可以理解为电信业的操作系统。CloudOS提供了硬件设备的虚拟化能力，将物理的计算/存储/网络资源变成虚拟资源供上层的软件使用。例如华为的云操作系统FusionSphere。
- VNF：VNF可以理解为各种不同网络功能的APP，是运营商传统网元（IMS，EPC，BRAS，CPE...）的软件实现。
- MANO：MANO的引入是要解决NFV多CT/IT厂家环境下的网络业务的发放问题，包括：分配物理/虚拟资源，垂直打通管理各层，快速适配对接新厂家新网元。MANO包括NFVO（Network Functions Virtualization Orchestrator，负责网络服务的生命周期的管理）、VNFM（Virtualized Network Function Manager, 负责VNF的生命周期管理）、VIM（Virtualized Infrastructure Manager，负责NFVI的资源管理）三部分。

NFV的标准架构

- ETSI定义了NFV标准架构，由NFVI、VNF以及MANO主要组件组成。NFVI包括通用的硬件设施及其虚拟化，VNF使用软件实现虚拟化网络功能，MANO实现NFV架构的管理和编排。



HUAWEI

NFV架构功能模块

- NFV标准架构定义的主要功能模块：

OSS/BSS 服务提供商的管理功能，不属于NFV框架内的功能组件，但MANO和网元需要提供对 OSS/BSS 的接口支持。

MANO NFV管理和编排。包括VIM，VNFM及NFVO，提供对VNF及I层统一的管理和编排功能。

- **VIM**：Virtualized Infrastructure Managers, NFVI管理模块，通常运行于对应的基础设施站点中，主要功能包括：资源的发现、虚拟资源的管理分配、故障处理等。
- **VNFM**：VNF Managers ，VNF管理模块，主要对VNF的生命周期（实例化、配置、关闭等）进行控制。
- **NFVO**：NFV Orchestration，实现对整个NFV基础架构、软件资源、网络服务的编排和管理。

VNF 指虚拟机及部署在虚拟机上的业务网元、网络功能软件等。

NFVI NFV基础设施，包括所需的硬件及软件。为VNF提供运行环境。

- **Hardware**：硬件层，包括提供计算、网络、存储资源能力的硬件设备。
- **Virtualization Layer**：虚拟化层，主要完成对硬件资源的抽象，形成虚拟资源，如虚拟计算资源、虚拟存储资源、虚拟网络资源。其虚拟化功能由Hypervisor[1] 实现。

- BSS：Business support system 业务支撑系统。
- OSS：Operation support system 运营支撑系统。
- Hypervisor：是一种运行在物理服务器和虚拟机操作系统之间的中间软件层，可允许多个操作系统和应用共享一套基础物理硬件。因此也可以看作是虚拟环境中的“元”操作系统，它可以协调访问服务器上的所有物理设备和虚拟机，也叫虚拟机监视器（Virtual Machine Monitor VMM）。Hypervisor是所有虚拟化技术的核心。目前主流的Hypervisor有KVM，VMWare ESXi，Xen，HyperV等。

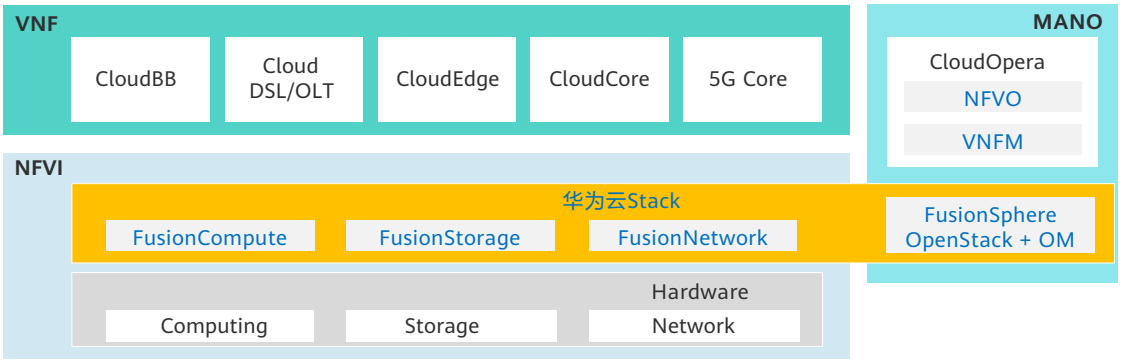
NFV架构接口

- NFV标准架构的主要接口：

接口类型	功能描述
Vi-Ha	虚拟化层与基础硬件之间的接口。虚拟化层满足基础硬件兼容性要求。
Vn-Nf	虚拟机（VM）与NFVI之间的接口。它确保虚拟机可以部署在NFVI上，满足性能、可靠性和可扩展性要求。NFVI满足虚拟机操作系统兼容性要求。
Nf-Vi	虚拟化层管理软件与NFVI之间的接口，提供NFVI虚拟计算、存储和网络系统管理；提供虚拟基础架构配置和连接；提供系统利用率、性能监控和故障管理。
Ve-Vnfm	VNFM与VNF之间的接口，实现VNF生命周期管理、VNF配置、VNF性能和故障管理。
OS-Ma	实现网络服务生命周期管理，VNF生命周期管理。
Vi-Vnfm	提供业务应用管理系统/业务编排系统与虚拟化层管理软件之间交互接口。
Or-Vnfm	给VNFM发送配置信息，对VNFM进行配置，完成Orchestrator与VNFM的对接；分配给一个VNF的NFVI资源的交互；VNF信息的交换。
Or-Vi	Orchestrator需要的资源预定及资源分配的请求；虚拟硬件资源配置及状态信息的交换。

华为NFV解决方案

- 华为NFV架构中，虚拟化层及VIM的功能由华为云Stack NFVI平台实现。华为云Stack可以实现计算资源、存储资源和网络资源的全面虚拟化，并能够对物理硬件虚拟化资源进行统一的管理、监控和优化。
- 华为提供运营商无线网、承载网、传输网、接入网、核心网等全面云化的解决方案。



- DSL（Digital Subscriber Line）数字用户线路
- OLT（Optical Line Terminal）光线路终端

FAQ

- Q1： 业界SDN与NFV是什么关系?
- A：虽然两者都是Network相关的变革，且NFV概念在SDN和OpenFlow世界大会上提出，但是两者彼此独立，没有必然关系。SDN主要影响网络架构，NFV主要影响网元的部署形态。

- Q2：华为的解决方案中SDN与NFV是什么关系?
- A：在华为的解决方案中SDN与NFV是不同的解决方案，但是有一定的关联关系。华为NFVI解决方案由华为云Stack提供。在其中网络方案选择中，可以选择SDN解决方案和非SDN解决方案。

思考题

1. （多选）以下对于华为SDN解决方案说法正确的是？（　　）

   A. 支持丰富的南向接口协议，例如RESTful、NETCONF和OVSDB。

   B. 支持OpenFlow作为南向接口协议。

   C. 管、控、析构建至简网络。

   D. 开放可编程网络接口，支持第三方应用开发和系统对接。

2. 请简述NFV的意义与价值。

1. BCD
2. NFV（Network Functions Virtualization ）是运营商为了解决电信网络硬件繁多、部署运维复杂、业务创新困难等问题而提出的。NFV在重构电信网络的同时，给运营商带来的价值包括但不限于：
   - 缩短业务上线时间
   - 降低建网成本
   - 提升网络运维效率
   - 构建开放的生态系统

本章总结

- 网络产业的变革与发展，提出了SDN（Software Defined Networking）软件定义网络与NFV（Network Functions Virtualization）网络功能虚拟化两个重要概念。
- SDN是网络架构的革新，以控制器为主体，让网络更加开放、灵活和简单。
- NFV是电信网络设备部署形态的革新，以虚拟化为基础，云计算为关键实现电信网络的重构。

更多信息

- 更多OpenFlow相关问题请参考 https://www.opennetworking.org/ 。
- 更多华为SDN解决方案内容请参考HCIP课程。

学习推荐

- https://e.huawei.com/cn/talent/#/home


谢 谢
www.huawei.com

网络编程与自动化

前言

- 网络工程领域不断出现新的协议、技术、交付和运维模式。传统网络面临着云计算、人工智能等新连接需求的挑战。企业也在不断追求业务的敏捷、灵活和弹性。在这些背景下，网络自动化变得越来越重要。
- 网络编程与自动化旨在简化工程师网络配置、管理、监控和操作等相关工作，提高工程师部署和运维效率。本课程定位于指导网络工程师初步了解Python编程实现网络自动化。

目标

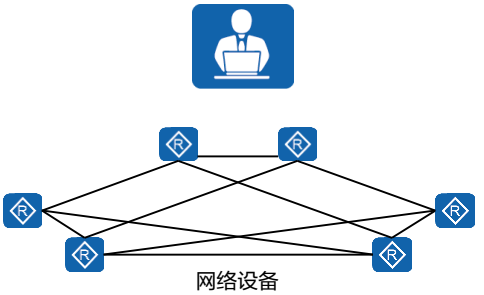
- 学完本课程后，您将能够：
  - 描述传统网络运维的困境
  - 了解网络自动化的实现方式
  - 了解编程语言的分类
  - 掌握Python编码规范
  - 掌握Python telnetlib的基本用法

目录

背景：传统网络运维困境

- 传统的网络运维工作需要网络工程师手动登录网络设备，人工查看和执行配置命令，肉眼筛选配置结果。这种严重依赖“人”的工作方式操作流程长，效率低下，而且操作过程不易审计。

设备多！操作烦琐！效率低！



网络设备

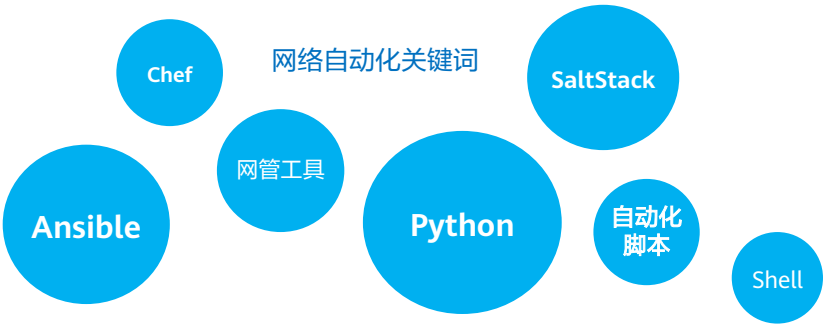
经典运维场景

在工作中你是否遇到过这样的场景：

1. 设备升级：现网有数千台网络设备，你需要周期性、批量性地对设备进行升级。
2. 配置审计：企业年度需要对设备进行配置审计。例如要求所有设备开启sTelnet功能，以太网交换机配置生成树安全功能。你需要快速地找出不符合要求的设备。
3. 配置变更：因为网络安全要求，需要每三个月修改设备账号和密码。你需要在数千台网络设备上删除原有账号并新建账号。

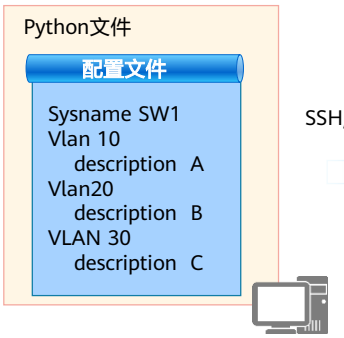
网络自动化

- 网络自动化，通过工具实现网络自动化地部署、运行和运维，逐步减少对“人”的依赖。这能够很好地解决传统网络运维的问题。
- 业界有很多实现网络自动化的开源工具，例如Ansible、SaltStack、Puppet、Chef等。从网络工程能力构建的角度考虑，更推荐工程师具备代码编程能力。



基于编程实现的网络自动化

- 近几年随着网络自动化技术的兴起，以Python为主的编程能力成为了网络工程师的新技能要求。
- Python编写的自动化脚本能够很好的执行重复、耗时、有规则的操作。



举例：Python实现设备自动化配置

- 网络自动化能做什么？最直观的一个网络自动化例子就是自动化配置设备。我们可以把这个过程分为两个步骤：编写配置文件和编写Python代码将配置文件推送到设备上。
- 首先用命令行方式写配置脚本，然后通过Telnet/SSH将它传到设备上运行。这种方式对于初学网络编程与自动化的网络工程师来说，比较容易理解。本章节主要介绍这种方式实现网络自动化。

- 业界也有很多基于开源工具的网络自动化，例如Ansible、SaltStack、Puppet、Chef等。网络工程师能力构建上更推荐具备代码编程能力。

目录

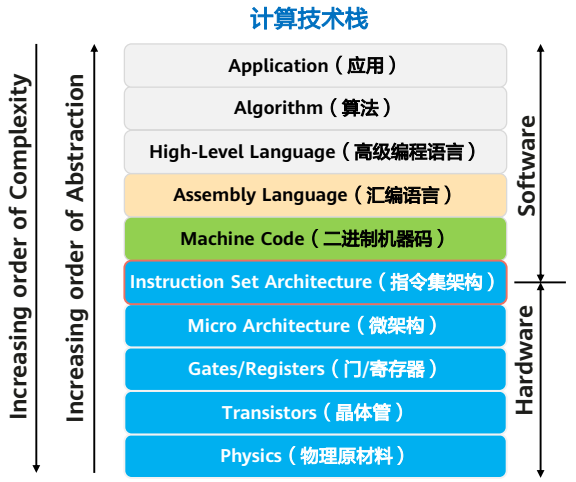
编程语言

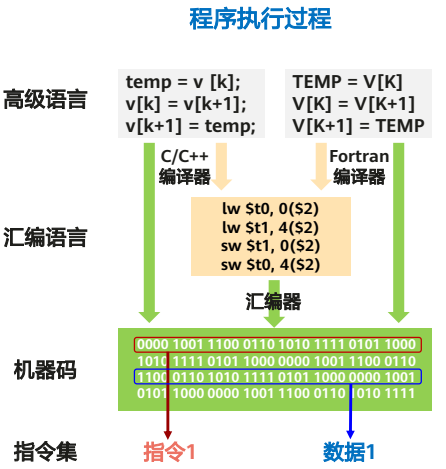
- 编程语言（Programming Language），是一种用于编写计算机程序的语言，用于控制计算机的行为。
- 按照语言在执行之前是否需要编译区分，可以将编程语言分为需要编译的编译型语言（Compiled Language），不需要编译的解释型语言（Interpreted Language）。

HUAWEI

- 计算机语言另一种分类方式（根据语言层次）是机器语言、汇编语言和高级语言。机器语言由0和1组成的指令构成，可以直接被机器识别。由于机器语言晦涩难懂，人们将0和1的硬件指令做了简单的封装，便于识别和记忆（例如MOV、ADD），这就是汇编语言。这两种语言都属于低级语言，其他语言都属于高级语言，例如C、C++、Java、Python、Pascal、Lisp、Prolog、FoxPro、Fortran等都是高级语言。高级语言编写的程序不能直接被计算机识别，必须经过转换成机器语言才能被执行。

计算技术栈与程序执行过程

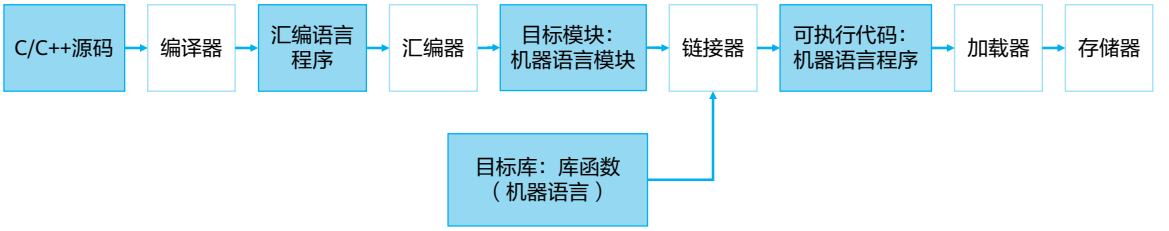




- 对于计算机的技术栈和程序执行的过程。左侧是计算的技术栈，我们可以看到硬件的最底层，是物理材料、晶体管来实现门电路和寄存器，再组成CPU的微架构。CPU的指令集是硬件和软件的接口，应用程序通过指令集中定义的指令驱动硬件完成计算。
- 应用程序通过一定的软件算法完成业务功能。程序通常使用如C/C++/Java/Go/Python等高级语言开发。高级语言需要编译成汇编语言，再由汇编器按照CPU指令集转换成二进制的机器码。
- 一个程序在磁盘上存在的形式，是一堆指令和数据所组成二进制机器码，也就是我们通常说的二进制文件。

高级编程语言 - 编译型语言

- **编译型语言**：编译型语言的程序在执行之前有一个编译过程，把程序编译成为机器语言的文件。运行时不需要重新翻译，直接使用编译的结果。典型的如C/C++/Go语言，都属于编译型语言。
- **从源码到程序的过程**：源码需要由编译器、汇编器翻译成机器指令，再通过链接器链接库函数生成机器语言程序。机器语言必须与CPU的指令集匹配，在运行时通过加载器加载到内存，由CPU执行指令。

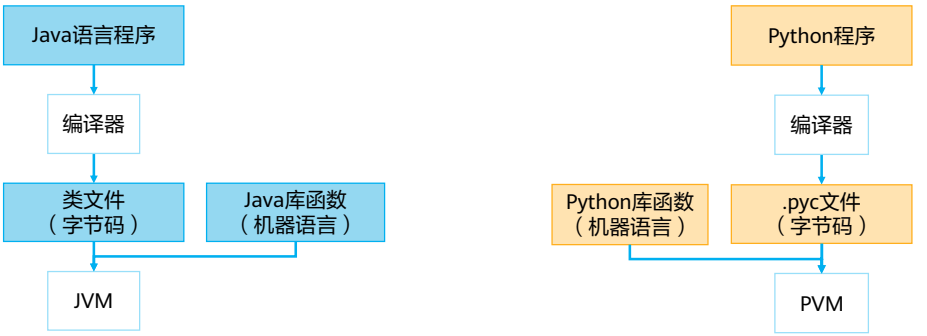


HUAWEI

- 编译型语言编译的时候直接编译成机器可以执行的格式（例如.exe .dll .ocx）。编译和执行是分开的，不能跨平台执行，例如X86程序不能在ARM架构服务器上运行。

高级编程语言 - 解释型语言

- **解释型语言**：解释型语言的程序不需要在运行前编译，在运行程序的时候才逐行翻译。典型的如Java/Python语言，都属于解释型语言。
- **从源码到程序的过程**：解释型语言的源代码由编译器生成字节码，然后再由虚拟机（JVM/PVM）解释执行。虚拟机将不同CPU指令集的差异屏蔽，因此解释型语言的可移植性相对较好。



HUAWEI

- JVM：Java虚拟机。
- PVM：Python虚拟机。

什么是Python?

- Python是一门完全开源的高级编程语言。它的作者是Guido Van Rossum。

Python的优点：
- Python拥有优雅的语法、动态类型具有解释性质。能够让学习者从语法细节的学习中抽离，专注于程序逻辑。
- Python同时支持面向过程和面向对象的编程。
- Python拥有丰富的第三方库。
- Python可以调用其他语言所写的代码，又被称为胶水语言。

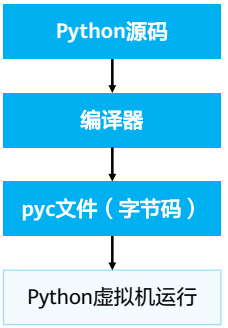
Python的缺点：
- 运行速度慢。Python是解释型语言，不需要编译即可运行。代码在运行时会逐行地翻译成CPU能理解的机器码，这个翻译过程非常耗时。

由于Python具有非常丰富的第三方库，加上Python语言本身的优点，所以Python可以在非常多的领域内使用：人工智能、数据科学、APP、自动化运维脚本等。

- Python同时也是动态类型语言。动态类型语言是指在程序运行的过程中自动决定对象的类型，不需要声明变量的类型。

Python代码执行过程

Python程序编译运行的过程



操作过程

1、在操作系统上安装Python和运行环境。

2、编写Python源码。

3、编译器运行Python源码，编译生成pyc文件（字节码）。

4、Python虚拟机将字节码转换为机器语言。

5、硬件执行机器语言。

- 对于Python而言，Python源码不需要编译成二进制代码，它可以直接从源代码运行程序。当我们运行Python代码的时候，Python解释器首先将源代码转换为字节码，然后再由Python虚拟机来执行这些字节码。
- Python虚拟机（Python VM）不是一个独立的程序，不需要独立安装。

初识Python代码 - 交互式运行

- Python有两种运行方式，交互式运行和脚本式运行。
- 交互式编程不需要创建脚本文件，是通过 Python 解释器的交互模式编写代码。

```
C:\Users\Richard>python
Python 3.7.4 (default, Aug  9 2019, 18:34:13) [MSC v.1915 64 bit (AMD64)] ::
Anaconda, Inc. on win32
Type "help", "copyright", "credits" or "license" for more information.
1. Input  -- >>> print ("hello world")
2. Output -- hello world
3. Input  -- >>> a = 1
4. Input  -- >>> b = 2
5. Input  -- >>> print ( a + b )
6. Output -- 3
>>>
```

初识Python代码 - 脚本式运行

- 脚本模式里的代码可以在各种Python编译器或者集成开发环境上运行。例如Python自带的IDLE、Atom、Visual Studio、Pycharm和Anaconda等。

demo.py文件

```
print ("hello world")
a = 1
b = 2
print ( a + b )
```

1. Input --
2. Output --
3. Output --

```
C:\Users\Richard>python demo.py
hello world
3
```

1 编写Python脚本文件(.py)

2 执行脚本文件

Python编码规范

- 编码规范是使用Python编写代码时应遵守的命名规则、代码缩进、代码和语句分割方式等。良好的编码规范有助于提高代码的可读性，便于代码的维护和修改。
- 例如分号、圆括号、空行和空格的使用规范建议如下：

分号

- Python程序允许在行尾添加分号，但是不建议使用分号隔离语句。
- 建议每条一句单独一行。

空行

- 不同函数或语句块之间可以使用空格来分隔。用以区分两段代码，提高代码可读性。

圆括号

- 圆括号可用于长语句的续行。一般不使用不必要的括号。

空格

- 不建议在括号内使用空格。
- 对于运算符，可以按照个人习惯决定是否在两侧加空格。

Python编码规范 - 标识符命名

- Python标识符用于表示常量、变量、函数以及其他对象的名称。
- 标识符通常由字母、数字和下划线组成，但不能以数字开头。标识符大小写敏感，不允许重名。如果标识符不符合规则，编译器运行代码时会输出SyntaxError语法错误。

1. 数值赋值 --	User_ID = 10	print (User_ID)
2. 数值赋值 --	user_id = 20	print (user_id)
3. 字符串赋值 --	User_Name = 'Richard'	print (User_Name)
4. 数值赋值 --	Count = 1 + 1	print (Count)
5. 错误的标识符 --	4_passwd = "Huawei"	print (4_passwd)

print ()为Python内置的函数，作用是输出括号内的内容。

问题：右侧print的运行结果是?

- Python最基本的数据类型有布尔型（True/False）、整数、浮点型、字符串型。Python里的所有数据（布尔值、整数、浮点、字符串，甚至大型数据结构、函数以及程序）都是以对象（object）的形式存在的。这使得Python语言有很强的统一性。
- 运行结果分别为10，20，Richard，2，SyntaxError（语法错误）。
- 本文不对Python语法做针对介绍，更多Python语法请参考HCIP课程。

Python编码规范 - 代码缩进

- 在Python程序中，代码缩进代表代码块的作用域。如果一个代码块包含两个或更多的语句，则这些语句必须具有相同的缩进量。对于Python而言代码缩进是一种语法规则，它使用代码缩进和冒号来区分代码之间的层次。
- 编写代码时候，建议使用4个空格来生成缩进。如果程序代码中使用了错误的缩进，则会在运行中发出IndentationError错误信息。

正确缩进 --

正确缩进 --

错误缩进 --

```
if True:
    print (“Hello”)
else:
    print (0)

a = “Python”
    print (a)
```


- if...else...是一个完整的代码块，拥有相同的缩进。
- print(a)调用参数a，并且和if...else...在一个代码块，需要有相同的缩进。

Python编码规范 - 使用注释

- 注释就是在程序中添加解释说明，能够增强程序的可读性。在Python程序中，注释分为单行注释和多行注释。
- 单行注释以 # 字符开始直到行尾结束。
- 多行注释内容可以包含多行，这些内容包含在一对三引号内（'''...'''或者"""..."""）。

单行注释 --

```
#将字符串赋值给a
a = “Python”
print (a)
```

多行注释 --

```
“””
运行输出结果为Python
“””
```

Python编码规范 - 源码文件结构

- 一个完整的Python源码文件一般包含几个组成部分：解释器和编码格式声明、文档字符串、模块导入和运行代码。
- 如果会在程序中调用标准库或其他第三方库的类时，需要先使用import或from... import语句导入相关的模块。导入语句始终在文件的顶部。在模块注释或文档字符串(docstring)之后。

```
解释器声明      --  #!/usr/bin/env python
编码格式声明    --  #-*- coding:utf-8 -*-

模块注释或文档字符串 --  """本文档的说明（docstring）

                    本文档作用是...
                    """

导入模块time    --  import time
运行代码        --  ...
```


- 解释器声明的作用是指定运行本文件的编译器的路径（非默认路径安装编译器或有多个Python编译器）。Windows操作系统上可以省略本例中第一行解释器声明。
- 编码格式声明的作用是指定本程序使用的编码类型，以指定的编码类型读取源代码。Python 2 默认使用的是 ASCII 编码 (不支持中文)，Python 3 默认支持 UTF-8 编码（支持中文)。
- 文档字符串的作用是对本程序功能的总体介绍。
- time为Python内置模块，作用是提供处理时间相关的函数。

Python的函数与模块

- 函数(Function)是组织好的、可重复使用的一段代码。它能够提高程序的模块化程度和代码利用率。函数使用关键字 def 定义。
- 模块(Module)是一个保存好的Python文件。模块可以由函数或者类组成。模块和常规Python程序之间的唯一区别是用途不同：模块用于被其他程序调用。因此，模块通常没有main函数。

demo.py文件

```
def sit():      #定义函数
    print ('A dog is now sitting')

sit()           #调用函数
```

运行结果：

```
A dog is now sitting.
```

1 编写Python文件(.py)

test.py文件

```
import demo     #导入模块

demo.sit()      #调用函数
```

运行结果：

```
A dog is now sitting.
A dog is now sitting.
```

2 调用模块

Python的类与方法

- 类(Class)是用来描述具有一类相同的属性和方法的集合。类的定义使用关键字 class。
- 被实例化的类的”函数”被称作方法(Method)。类定义方法时候必须携带 self 关键字，它表示类的实例本身。

demo.py文件

```
class Dog():      # 定义类
    def sit(self):  # 定义方法
        print("A dog is now sitting.")

Richard = Dog()    #实例化类
Richard.sit()
print (type(Richard.sit)) #实例化后类型为方法
print (type(Dog.sit))     #类型为函数
```

运行结果：

```
A dog is now sitting.
<class 'method'>
<class 'function'>
```

1 编写Python文件(.py)

test.py文件

```
import demo

demo.Dog.sit
```

运行结果：

```
A dog is now sitting.
<class 'method'>
<class 'function'>
```

2 调用模块

- 对于函数和方法的官方定义：
- 函数 Function: A series of statements which returns some value to a caller. It can also be passed zero or more arguments which may be used in the execution of the body.
- 方法 Method: A function which is defined inside a class body. If called as an attribute of an instance of that class, the method will get the instance object as its first argument (which is usually called self).
- 更多类的学习，请参考https://docs.python.org/3/tutorial/classes.html。

telnetlib介绍

- telnetlib是Python标准库中的模块。它提供了实现Telnet功能的类telnetlib.Telnet。
- 这里通过调用telnetlib.Telnet类里的不同方法实现不同功能。

```
导入telnetlib模块Telnet类  --  from telnetlib import Telnet
Telnet连接到指定服务器上  --  tn = Telnet(host=None, port=0[, timeout])
调用read_all()方法  --  tn.read_all()
                        ...
```

方法	功能
Telnet.read_until（expected, timeout=None）	读取直到给定的字符串expected或超时秒数。
Telnet.read_all ()	读取所有数据直到EOF(End Of File)。阻塞直到连接关闭。
Telnet.read_very_eager()	读取从上次IO阻断到现在所有的内容，返回字节串。连接关闭或者没有数据时触发EOFError异常。
Telnet.write(buffer)	写入数据。在套接字(Socket)上写一个字节串，加倍任何IAC(Interpret As Command)字符。
Telnet.close()	关闭连接。

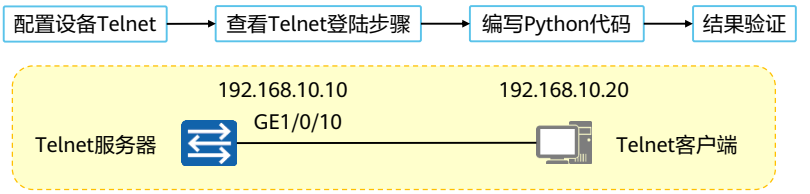
- Telnet定义了网络虚拟终端（NVT，Network Virtual Terminal）。它描述了数据和命令序列在Internet上传输的标准表示方式，以屏蔽不同平台和操作系统的差异，例如不同平台上换行的指令不一样。
- Telnet通信采用带内信令方式，即Telnet命令在数据流中传输。为了区分Telnet命令和普通数据，Telnet采用转义序列。每个转义序列由两个字节构成，前一个字节是(0xFF)叫做IAC（Interpret As Command）“解释为命令”，标识了后面一个字节是命令。EOF也是一种Telnet命令，十进制编码是236。
- 套接字（socket）是一个抽象层，应用程序通常通过"套接字"向网络发出请求或者应答网络请求。
- 更多可参考https://docs.python.org/3/library/telnetlib.html

目录

案例：使用telnetlib登陆设备

案例描述：

- 现有一台网络设备作为Telnet服务器，需要实现使用Python telnetlib作为Telnet客户端登录此设备。



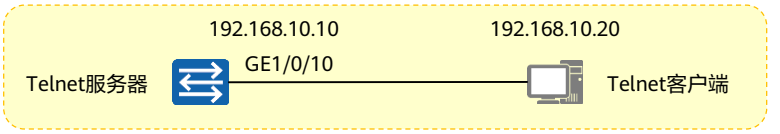
实现过程分为四个步骤：

- 配置设备Telnet服务。
- 手动验证和查看Telnet登录步骤，作为代码实现的参考。
- 编写和运行Python代码。
- 验证结果。

- 本案例出于学习角度让工程师了解代码和设备之间telnet交互过程使用telnetlib。在真实的工作场景中因为安全原因更为推荐设备开启SSH功能而关闭Telnet功能，对应的Python模块可以选择paramiko或netmiko。更多进阶内容请参考HCIP-Datacom-Network Automation Developer。

案例：使用telnetlib登陆设备

配置设备Telnet → 查看Telnet登陆步骤 → 编写Python代码 → 结果验证



配置设备接口地址：

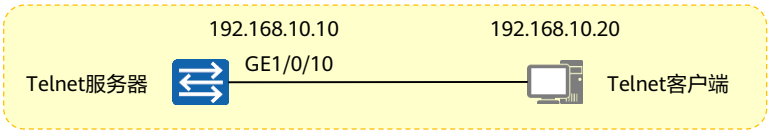
```
[Huawei] interface GE 1/0/0
[Huawei -GE1/0/0] ip add 192.168.10.10 24
[Huawei -GE1/0/0] quit
```

配置设备Telnet服务：

```
[Huawei] user-interface vty 0 4
[Huawei-ui-vty0-4] authentication-mode password
[Huawei-ui-vty0-4] set authentication password simple Huawei@123
[Huawei-ui-vty0-4] protocol inbound telnet
[Huawei-ui-vty0-4] user privilege level 15
[Huawei-ui-vty0-4] quit
[Huawei] telnet server enable
```

案例：使用telnetlib登陆设备

配置设备Telnet → **查看Telnet登陆步骤** → 编写Python代码 → 结果验证



Telnet登录操作：

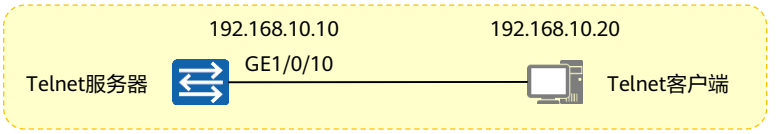
1 登录命令	--	C:\Users\Richard>telnet 192.168.10.10
回显信息	--	Login authentication
2 输入密码	--	Password:
回显信息	--	Info: The max number of VTY users is 5, and the number of current VTY users on line is 1. The current login time is 2020-01-15 21:12:57. <Huawei>

HUAWEI

- 本案例手工Telnet登录操作以windows为例：首先输入登录命令，telnet 192.168.10.10。因在前序步骤中设备配置Telnet使用密码登录，所以此处回显信息为“Password:”。此时输入密码 Huawei@123 完成验证，成功登录。

案例：使用telnetlib登陆设备

配置设备Telnet → 查看Telnet登陆步骤 → **编写Python代码** → 结果验证



导入模块	--	import telnetlib
定义登录设备IP	--	**host** = '192.168.10.10'
定义登录设备密码	--	**password** = 'Huawei@123'
Telnet登录到主机	--	tn = telnetlib.Telnet(**host**)
读取直到回显信息为”Password:”	--	tn.read_until(b’Password:’)
输入编码为ASCII的密码并换行	--	tn.write(**password**.encode('ascii') + b"\n")
输出读取直到到”<Huawei>”的信息	--	print (tn.read_until(b'<Huawei>').decode('ascii’))
关闭Telnet连接	--	tn.close()

HUAWEI

- Python中encode()和decode()函数的作用是，以指定的方式编码格式编码字符串和解码字符串。本例中，password.encode(‘ascii‘) 表示将字符串’Huawei@123’转为为ASCII。此处编码格式遵守telnetlib模块官方要求。
- Python字符串增加b，b’ str‘ 表示这是字符串是 bytes 对象。本例中，b’Password:’表示将字符串’Password:’转换为bytes类型字符串。此处编码格式遵守telnetlib模块官方要求。
- 更多Python对象描述，请参考 https://docs.python.org/3/reference/datamodel.html#objects-values-and-types。

案例：运行结果对比

配置设备Telnet → 查看Telnet登陆步骤 → 编写Python代码 → **结果验证**

手动Telnet登录结果：

```
C:\Users\Richard>telnet 192.168.10.10
Login authentication

Password:
Info: The max number of VTY users is 5, and the number of current VTY users on line is 1.
The current login time is 2020-01-15 21:12:57.
<Huawei>
```

Python代码运行结果：

```
#编译器运行Python代码
Info: The max number of VTY users is 5, and the number
      of current VTY users on line is 1.
      The current login time is 2020-01-15 22:12:57.
<Huawei>
```

思考题

1. Python属于编译型语言。（　　）
   A. 正确
   B. 错误
2. 基于本文案例，要求使用telnetlib创建VLAN 10如何实现?

HUAWEI

1. B
2. 可以使用telnetlib.write()方法，在登录设备后，首先输入命令system进入系统视图，然后输入配置命令VLAN 10进行创建。（VRP8请输入system immediately进入系统视图）

本章总结

- 网络自动化是通过工具实现网络自动化的部署、运行和运维，逐步减少对“人”的依赖。可以通过编程语言或者工具实现。
- Python是一门完全开源的高级编程语言，语法简单，容易学习。拥有丰富的标准库和第三方库，适用于网络工程领域。
- Python的telnetlib模块提供了实现Telnet功能的类telnetlib.Telnet。可以让您初窥网络编程与自动化世界！华为更多开放可编程内容请参考HCIP-Datacom！

更多信息

- 更多Python相关问题请参考https://www.python.org/


谢 谢
www.huawei.com

园区网典型组网架构及案例实践

前言

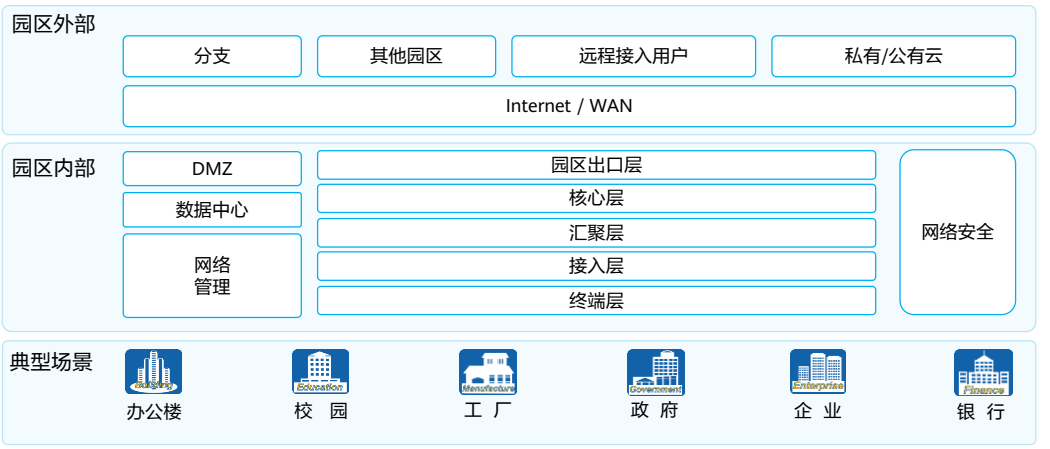
- 当您在校园学习，单位工作，商场购物时，您可能会注意到，这些场所都被网络覆盖。通过网络，您可以访问学校内部资源，可以访问公司内部打印机打印文档，也可以访问Internet浏览新闻资讯。
- 这些网络在分类上都属于园区网络，一般由企业或者机构自己搭建。园区网络不仅可以提升企业的运作效率，同时也可以对外提供网络接入服务。
- 本章我们将学习园区网络基本架构，同时了解如何搭建一张园区网。

目标

- 学完本课程后，您将能够：
  - 了解园区网的定义
  - 了解园区网的典型组网架构
  - 掌握小型园区网规划设计方法
  - 掌握小型园区网部署实施方法
  - 了解小型园区网运维概念
  - 了解小型园区网优化概念
  - 独立完成一个园区网络工程项目

目录

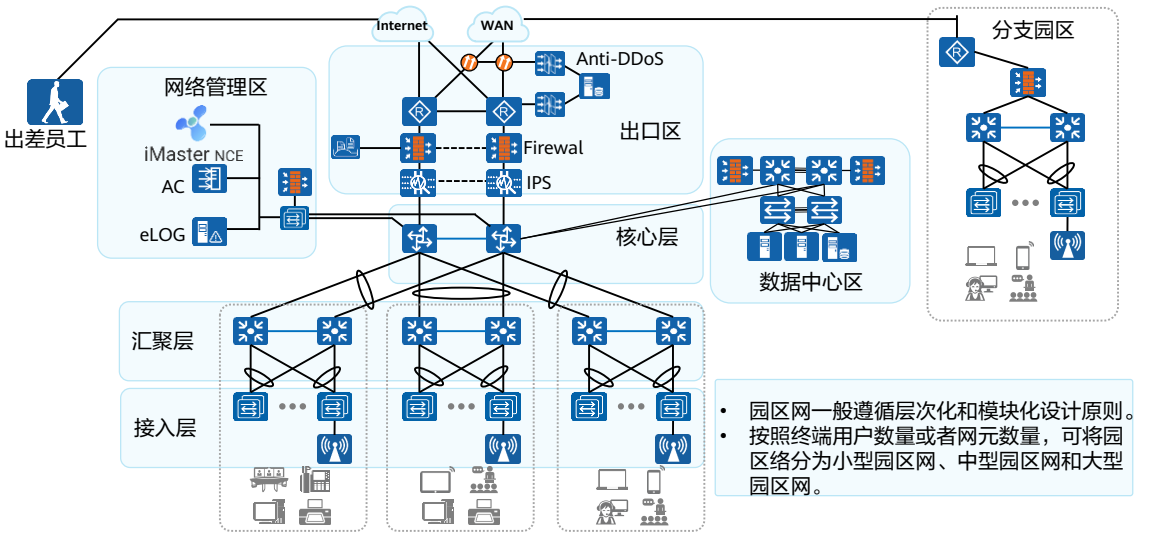
什么是园区网



园区网络是限定区域内，连接人与物的局域网络；园区网络通常只有一个管理主体；如果有多个管理主体，通常被认为为多个园区网络。

- 园区网络的规模可大可小，小到一个SOHO（Small Office Home Office，家居办公室），大到校园、企业园区、公园、购物中心等。园区的规模是有限的，一般的大型园区，例如高校园区、工业园区，规模依然被限制在几平方公里以内，在这个范围内，我们可以使用局域网技术构建网络。超过这个范围的“园区”通常被视作一个“城域”，需要使用到广域网技术，相应的网络会被视作城域网。
- 园区网络使用的典型局域网技术包括遵循IEEE（Institute of Electrical and Electronics Engineers，电气和电子工程师协会）802.3 标准的Ethernet 技术（有线）和遵循IEEE 802.11 标准的Wi-Fi 技术（无线）。

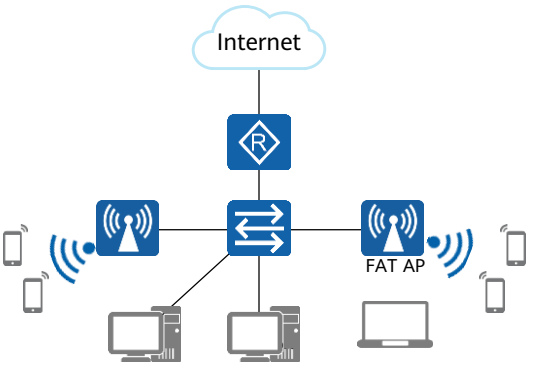
园区网络典型架构



HUAWEI

- 园区网络典型层次和区域：
  - 核心层：是园区网骨干，是园区数据交换的核心，联接园区网的各个组成部分，如数据中心、管理中心、园区出口等。
  - 汇聚层：处于园区网的中间层次，完成数据汇聚或交换的功能，可以提供一些关键的网络基本功能，如路由、 QoS、安全等。
  - 接入层：为终端用户提供园区网接入服务，是园区网的边界。
  - 出口区：园区内部网络到外部网络的边界，用于实现内部用户接入到公网，外部用户接入到内部网络。一般会在此区域中部署大量的网络安全设备来抵御外部网络的攻击，如IPS（intrusion prevention system，入侵防御系统）、Anti-DDoS设备、Firewall（防火墙）等。
  - 数据中心区：部署服务器和应用系统的区域，为企业内部和外部用户提供数据和应用服务。
  - 网络管理区：部署网络管理系统的区域，包括SDN控制器，无线控制器，eLOG（日志服务器）等，管理监控整个园区网络。

小型园区网络典型架构

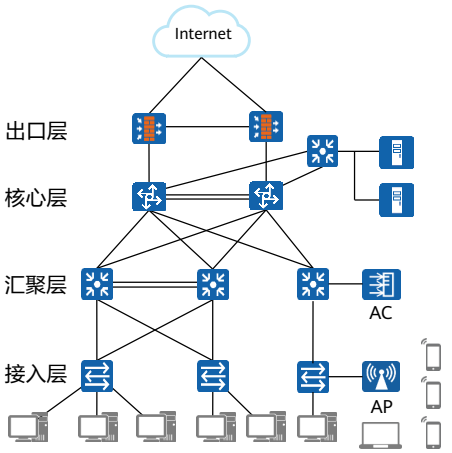


某连锁咖啡店网络拓扑

- 小型园区网络应用于接入用户数量较少的场景，一般支持几个至几十个用户。网络覆盖范围也仅限于一个地点，网络不分层次结构。网络建设的目的常常就是为了满足内部资源互访。
- 小型园区网络特点：
  - 用户数量较少
  - 仅单个地点
  - 网络无层次性
  - 网络需求简单

终端用户数（个）	<200
网元数量（个）	<25

中型园区网络典型架构

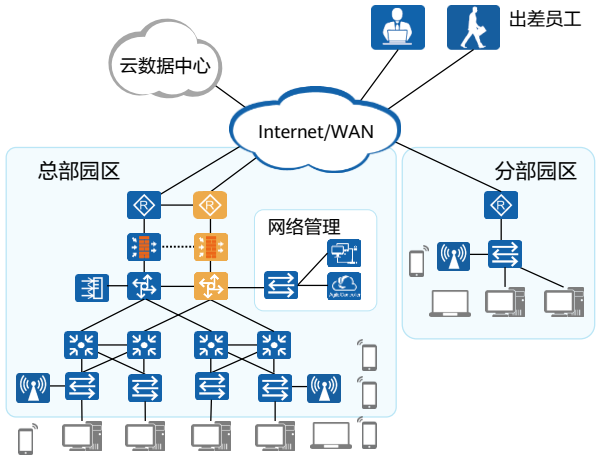


某外贸公司网络拓扑

- 中型园区网络能够支撑几百至上千用户的接入。
- 中型网络引入了按功能进行分区的理念，也就是模块化的设计思路，但功能模块相对较少。一般根据业务需要进行灵活分区。
- 中型园区网络特点：
  - 规模中等
  - 使用场合最多
  - 功能分区
  - 一般采用三层网络结构：核心、汇聚、接入

终端用户数（个）	200~2000
网元数量（个）	25~100

大型园区网络典型架构

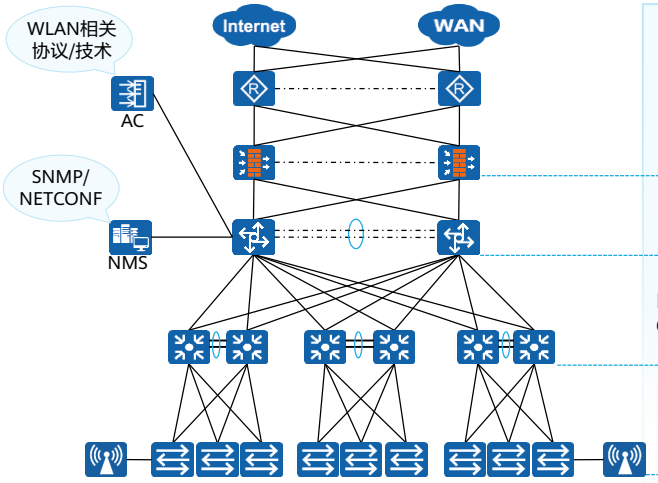


某大型企业网络拓扑

- 大型园区网络可能是覆盖多幢建筑的网络，也可能是通过WAN连接一个城市内的多个园区的网络。一般会提供接入服务，允许出差员工通过VPN等技术接入公司内部网络。
- 大型园区网络特点：
  - 覆盖范围广
  - 用户数量多
  - 网络需求复杂
  - 功能模块全
  - 网络层次丰富

终端用户数（个）	>2000
网元数量（个）	>100

园区网络主要协议/技术



常用协议/技术	
NAT、OSPF、静态路由、PPPoE等	出口区
堆叠、OSPF、静态路由、ACL等	核心层
DHCP、堆叠、链路聚合、生成树、OSPF、静态路由等	汇聚层
VLAN、生成树、链路聚合、AAA等	接入层

HUAWEI

目录

HUAWEI

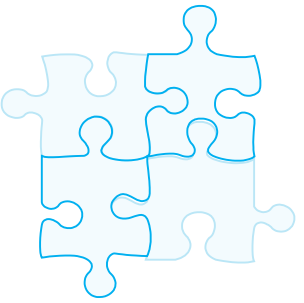
网络需求

- 某公司（规模为200人左右）因业务发展需要，准备搭建一张全新的园区网络，对网络需求如下：
  - 能够满足公司当前的业务需求
  - 网络拓扑简单，维护方便
  - 提供有线接入供员工办公使用，提供WiFi服务供访客使用
  - 做到简单的网络流量管理
  - 保证一定的安全性

园区网络项目生命周期

1 规划与设计

- 设备选型
- 物理拓扑
- 逻辑拓扑
- 使用技术与协议等



2 部署与实施

- 设备安装
- 单机调测
- 联调测试
- 割接并网等

3 网络运维

- 日常维护
- 软件与配置备份
- 集中式网管监控
- 软件升级等

4 网络优化

- 提升网络的安全性
- 软件与配置备份
- 提升网络的用户体验等

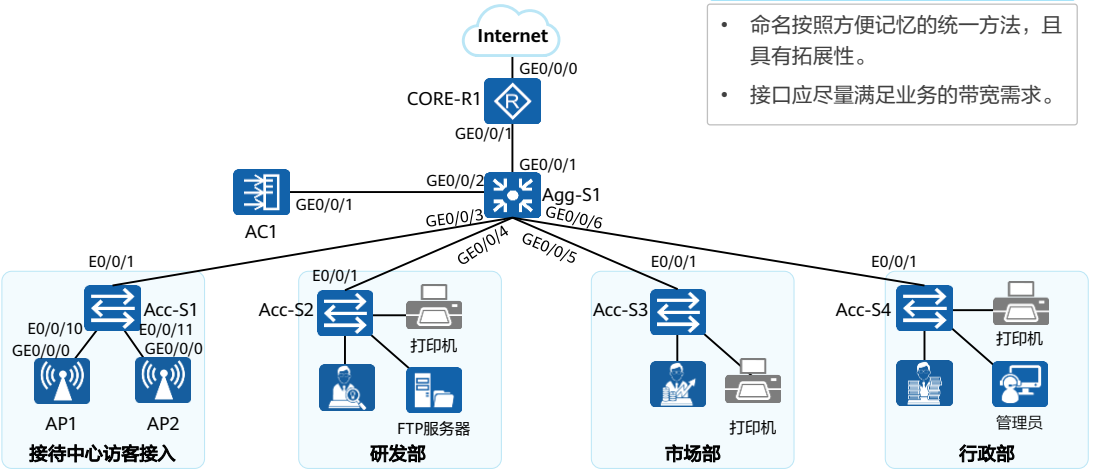
- 网络的规划与设计是一个项目的起点，完善细致的规划工作将为后续的项目具体工作打下坚实的基础。
- 项目实施是工程师交付项目的具体操作环节，系统的管理和高效的流程是确保项目实施顺利完成的基本要素。
- 要保证网络各项功能正常运行、从而支撑用户业务的顺利开展，需要对网络进行日常的维护工作和故障处理。
- 用户的业务在不断发展，因此用户对网络功能的需求也会不断变化。当现有网络不能满足业务需求，或网络在运行过程中暴露出了某些隐患时，就需要通过网络优化来解决。

小型园区网络设计

1.组网方案设计	2.网络设计	3.安全设计	4.运维管理设计
设备选型	基础业务设计	出口安全设计	基础网络管理
物理拓扑	WLAN设计	内网有线安全	智能运维
	二层环路避免设计	内网无线安全	
	网络可靠性设计		

组网方案设计

- 综合考虑预算、业务需求等因素之后，物理拓扑如下图所示：



命名及接口选取规则

- 命名按照方便记忆的统一方法，且具有拓展性。
- 接口应尽量满足业务的带宽需求。

- 整个网络采用三层架构
  - 接入层接入交换机采用S3700，为员工PC以及打印机等终端提供百兆网络接入。
  - 汇聚层采用S5700设备，作为二层网络的网关。
  - 核心&出口采用AR2240设备，作为整个园区网络的出口。
- 注：Agg为Aggregation的缩写，表示汇聚层设备。Acc为Access的缩写，表示接入层设备。

基础业务设计：VLAN设计

- VLAN编号建议连续分配，以保证VLAN资源合理利用。
- VLAN划分需要区分业务VLAN、 管理VLAN和互联VLAN。
- 最常用的划分方式是基于接口的方式。

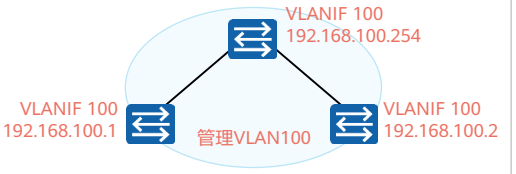
业务VLAN设计

按地理区域划分VLAN

按逻辑区域划分VLAN　　按人员结构划分VLAN

按业务类型划分VLAN

管理VLAN设计



通常，二层交换机使用VLANIF接口地址作为管理地址。建议所有属于同一二层网络的交换机使用同一管理VLAN， 管理IP地址处于同一网段。

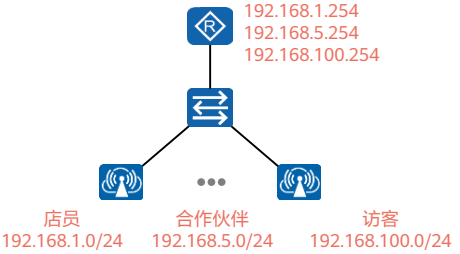
VLAN规划

- 预留二层设备的管理VLAN。
- 根据人员结构划分，分为访客VLAN，研发部VLAN，市场部VLAN，行政部VLAN。
- 考虑到三层交换机需要通过VLANIF与路由连通，所以需要预留互联VLAN。
- AP与AC之间建立CAPWAP隧道所需要的VLAN。

VLAN编号	VLAN描述
1	访客VLAN/WLAN的业务VLAN
2	研发部VLAN
3	市场部VLAN
4	行政部VLAN
100	二层设备的管理VLAN
101	WLAN的管理VLAN
102	Agg-S1与CORE-R1之间的互联VLAN

基础业务设计：IP地址设计

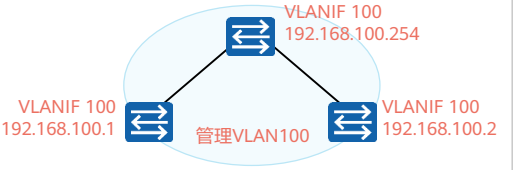
业务IP地址



业务IP地址是服务器、主机以及网关的IP地址。

- 网关IP地址推荐统一使用相同的末位数字，如.254。
- 各业务IP地址范围要清晰区分，每一类业务终端IP地址连续、可聚合。
- 建议使用掩码为24位的IP地址段。

管理IP地址



二层设备使用VLANIF地址作为管理IP地址，建议网关下的所有二层交换机使用同一网段。

网络设备互联IP地址

互联IP地址推荐使用30位掩码的IP地址，核心设备使用主机地址较小的IP地址。

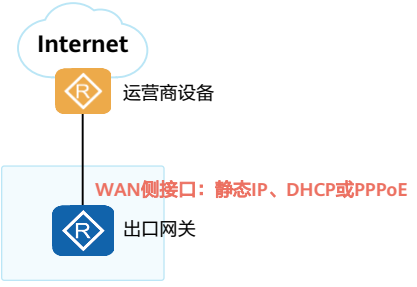
IP地址规划

- 综合考虑接入客户端个数并预留足够的IP地址，为每类业务规划网段及网关地址。
- 为管理IP划分网段。
- 为互联IP划分网段。

IP网段/掩码	网关地址	网段描述
192.168.1.0/24	192.168.1.254	无线接入访客所属网段，网关位于Agg-S1
192.168.2.0/24	192.168.2.254	研发部所属网段，网关位于Agg-S1
192.168.3.0/24	192.168.3.254	市场部所属网段，网关位于Agg-S1
192.168.4.0/24	192.168.4.254	行政部所属网段，网关位于Agg-S1
192.168.100.0/24	192.168.100.254	二层设备的管理网段，网关位于Agg-S1
192.168.101.0/24	N/A	WLAN的管理网段
192.168.102.0/30	N/A	Agg-S1与CORE-R1之间互联网段
1.1.1.1/32	N/A	CORE-R1上的Loopback接口地址，作为管理IP使用

基础业务设计：IP地址分配方式设计

出口网关

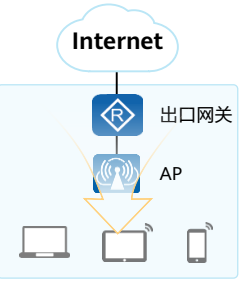


WAN侧接口的IP地址由运营商进行分配，可通过静态IP地址、 DHCP、 或者PPPoE方式分配，需与运营商沟通获取。

服务器、打印机等设备

服务器、特殊终端设备（打卡机、打印服务器、 IP视频监控设备等）建议采用静态IP地址。

终端用户



终端用户的IP地址分配，建议采用DHCP方式，由网关设备提供DHCP服务。

- IP地址分配时可以使用动态IP分配或者静态IP绑定。在中小型园区中， IP地址具体的分配原则如下：
- 出口网关设备： WAN侧接口的IP地址由运营商进行分配，可以通过静态IP地址、 DHCP、或者PPPoE方式分配，对于出口网关的IP地址需要提前与运营商沟通获取。
- 服务器、特殊终端设备（打卡机、打印服务器、 IP视频监控设备等）建议采用静态IP地址绑定方式分配。
- 用户终端：用户办公用PC、IP电话等设备建议通过在网关设备上部署DHCP Server后，统一通过DHCP方式动态分配。

IP地址分配方式规划

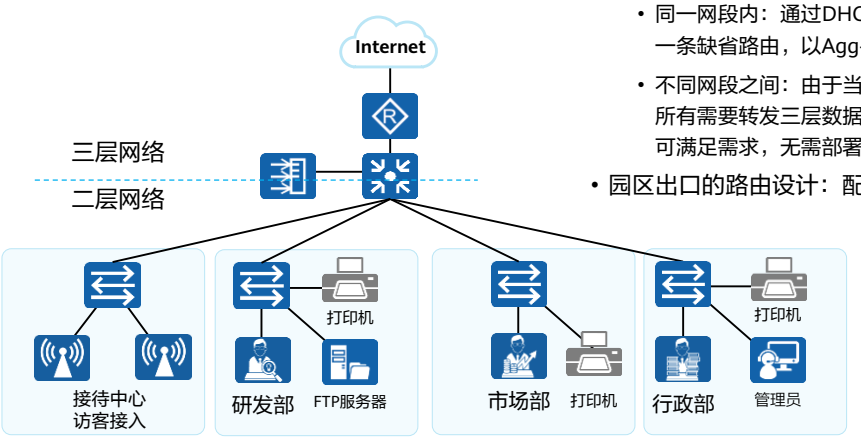
- 出口网关采用PPPoE方式获取IP地址。
- 所有终端采用DHCP方式获取IP地址，服务器及打印机分配固定的IP地址。
- 所有网络设备上的IP地址采用手工静态方式配置（AP除外）。

IP网段/接口	分配方式	分配方式描述
192.168.1.0/24 192.168.2.0/24 192.168.3.0/24 192.168.4.0/24	DHCP	由网关Agg-S1分配，还应分配给服务器及打印机等固定设备分配固定IP地址。
192.168.100.0/24	静态	设备管理IP，静态配置
192.168.101.0/24	DHCP	AC地址静态配置，AP地址由Agg-S1分配
192.168.102.0/30	静态	互联IP，静态配置
CORE-R1的GE0/0/0	PPPoE	运营商分配的IP地址

规划与设计 部署与实施 网络运维 网络优化

基础业务设计：路由设计

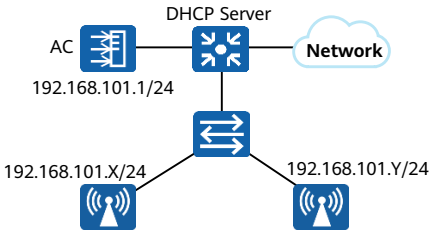
- 园区内部的路由设计：
  - 同一网段内：通过DHCP分配IP地址后默认会生成一条缺省路由，以Agg-S1作为三层网关。
  - 不同网段之间：由于当前拓扑较为简单，通过在所有需要转发三层数据的设备上部署静态路由即可满足需求，无需部署复杂的路由协议。
- 园区出口的路由设计：配置静态默认路由。



- 中小型园区网络的路由设计包括园区内部的路由设计及园区出口与Internet/广域设备之间的路由设计。
- 园区内部的路由设计：主要满足园区内部设备/终端的互通需求，并且可以与外部路由交互。由于中小型园 区的网络规模比较小，网络结构也比较简单。
  - AP设备：通过DHCP分配IP地址后默认会生成一条缺省路由。
  - 交换机、网关设备：通过静态路由即可满足需求，无需部署复杂的路由协议。
- 园区出口的路由设计：出口路由设计主要满足园区内部用户访问Internet和广域网的需求。出口设备与 Internet或者WAN连接时，建议在出口设备上配置静态路由来满足需求。

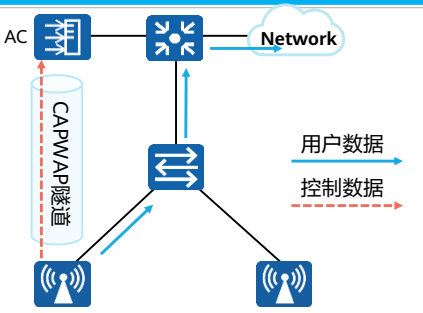
WLAN设计

WLAN组网设计



- 根据AC和AP的IP地址情况，以及数据流量是否流经AC，可将组网划分为
  - 直连二层组网
  - 旁挂二层组网
  - 直连三层组网
  - 旁挂三层组网
- 本案例采用旁挂二层组网方式

WLAN数据转发方式设计



- WLAN中的数据包括控制报文和数据报文
  - 控制报文通过CAPWAP隧道转发
  - 用户数据报文分为隧道转发、直接转发
- 本案例采用直接转发方式

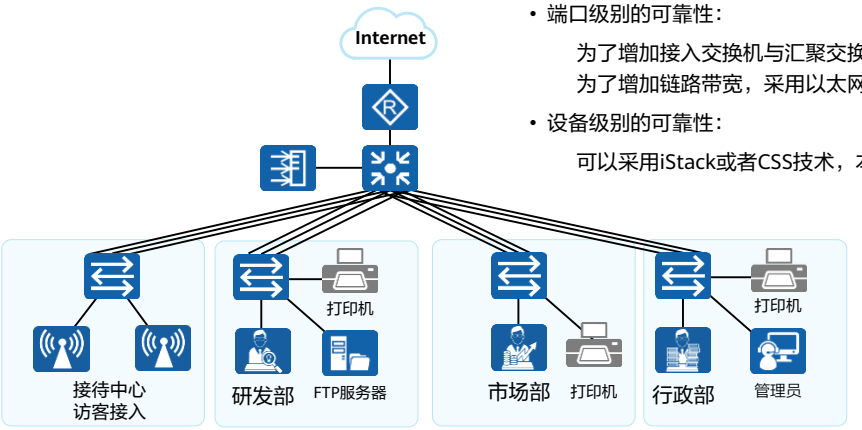
HUAWEI

- 除了要规划组网和数据转发方式外，仍需进行：
  - 网络覆盖设计：针对无线网络覆盖的区域设计规划，保证区域覆盖范围内的信号强度能满足用户的要求，并且解决相邻AP间的同频干扰问题。
  - 网络容量设计：根据无线终端的带宽要求、终端数目、并发率、单AP性能等数据来设计部署网络所需的AP数量，确保无线网络性能可以满足所有终端的上网业务需求。
  - AP布放设计：在网络覆盖设计的基础上，根据实际情况对AP的实际布放位置、布放方式和供电走线原则进行修正确认。
  - 此外还需进行WLAN安全设计、漫游设计等，本课程不再一一列举。

WLAN数据规划

配置项	配置内容
AP管理VLAN	VLAN101
STA业务VLAN	VLAN1
DHCP服务器	Agg-S1作为DHCP服务器为AP和STA分配地址，STA的默认网关为192.168.1.254
AP的IP地址池	192.168.101.2~192.168.101.253/24
STA的IP地址池	192.168.1.1~192.168.1.253/24
AC的源接口IP地址	VLANIF101：192.168.101.1/24
AP组	名称：ap-group1 引用模板：VAP模板WLAN-Guest、域管理模板default
域管理模板	名称：default 国家码：CN
SSID模板	名称：WLAN-Guest SSID名称：WLAN-Guest
安全模板	名称：WLAN-Guest 安全策略：WPA-WPA2+PSK+AES 密码：WLAN@Guest123
VAP模板	名称：WLAN-Guest 转发模式：直接转发 业务VLAN：VLAN1 引用模板：SSID模板WLAN-Guest、安全模板WLAN-Guest

可靠性设计



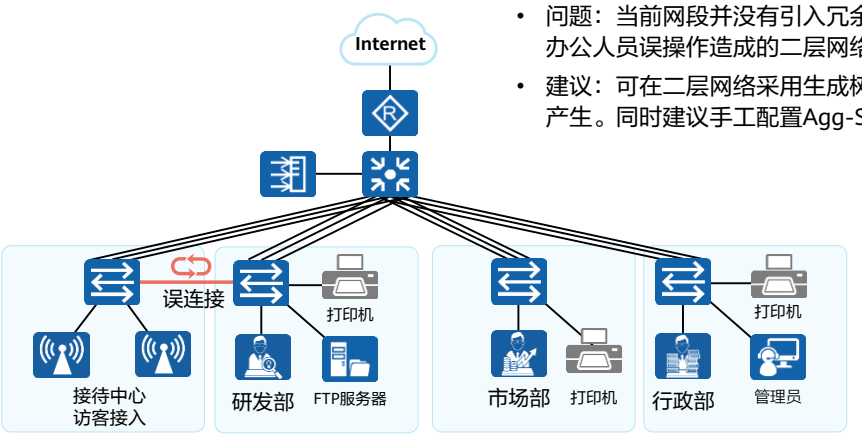
- 端口级别的可靠性：

  为了增加接入交换机与汇聚交换机之间的可靠性，同时为了增加链路带宽，采用以太网链路聚合技术。

- 设备级别的可靠性：

  可以采用iStack或者CSS技术，本组网不涉及。

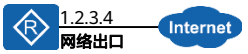
二层环路避免



- 问题：当前网段并没有引入冗余链路，如何防止办公人员误操作造成的二层网络环路呢？
- 建议：可在二层网络采用生成树技术，防止环路产生。同时建议手工配置Agg-S1为根桥。

出口NAT设计

静态NAT

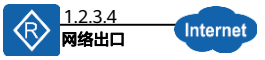


NAT映射表

私网地址	公有地址
192.168.1.1	1.2.3.1
192.168.1.2	1.2.3.2

- 静态NAT适用于有较多静态IP且有客户端需要使用固定IP地址的场景。

动态NAT

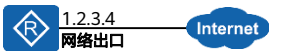


NAT地址池

1.2.3.1	Not Use
1.2.3.2	Not Use
1.3.3.3	Not Use

- 动态NAT拥有地址池概念，取地址池中可用地址给客户端访问Internet使用。

NAPT与Easy IP



NAT映射表

私网地址：端口	公有地址：端口
192.168.1.10:80	1.2.3.4:10335

- NAPT在动态NAT的基础上对端口也进行转换，提高公网地址利用率。
- Easy IP 适用于网络出接口地址动态场景。

NAT Server

NAT映射表

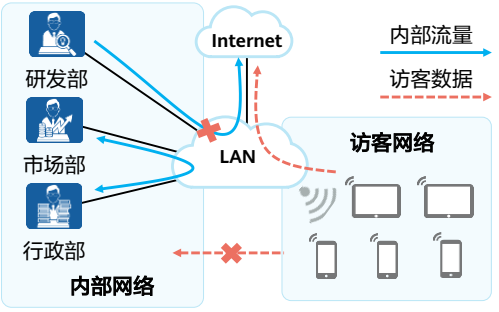
私网地址：端口	公有地址：端口
192.168.1.1:10321	1.2.3.4:1025
192.168.1.2:17087	1.2.3.4:1026

NAT Server适合内网有服务器需要向外部提供服务的场景。

HUAWEI

安全设计

流量管控



- 允许部门之间的流量互访，但不允许访问Internet
- 访客可以访问Internet但不能访问园区内部网络。
- 可以用过traffic-policy、traffic-filter等技术完成内外网隔离，通过NAT控制内部网络访问Internet。

DHCP安全

- 在园区网中，经常会出现有员工私接带DHCP的无线路由器，导致内网地址混乱，出现地址冲突、无法上网等情况。
- 一般会在接入交换机使能DHCP Snooping防止这种情况。

网络管理安全

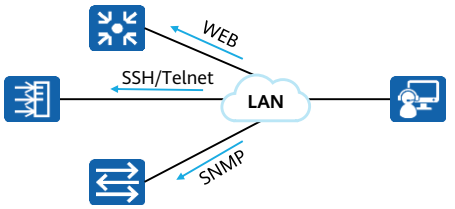
- 当使用Telnet或WEB等方式对设备进行网络管理时，可以通过ACL技术，仅允许固定的用户（IP）登录管理。
- 对于集中式网管，SNMPv3增加了身份验证和加密处理，可以大大提高网管的安全性。

HUAWEI

- 说明：本案例的安全设计仅依靠路由器或交换机设备实现。

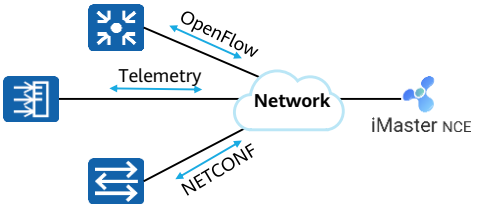
运维管理设计

传统设备管理



- 保证网络管理员与设备IP可达的情况下，可以通过Telnet、WEB以及SSH等方式对设备进行管理。
- 当网络中设备较多时，可部署基于SNMP协议的统一网管软件进行网络的运维与管理。

基于iMaster NCE平台管理



- 除了基于SNMP的传统网管之外，也可采用华为iMaster NCE平台进行网络的管理和运维，实现网络“自动驾驶”。

小型园区网络部署与实施

- 项目的部署与实施需要按照一定流程进行，内容包括：
  - 方案制定
  - 设备安装
  - 网络调试
  - 割接并网
  - 转维培训
  - 项目验收
- 具体流程按照项目实际情况进行确定。

配置方案（1）

1. 网络设备之间物理线路连接，配置链路聚合，同时添加接口描述，详细内容如下：

设备	接口	配置内容
Acc-S1	Eth-trunk 1	mode:LACP-static Trunkport :GE0/0/1、GE0/0/2、GE0/0/3 description:to Agg-S1's eth-trunk 1
	E0/0/10	Description:to AP1
	E0/0/11	Description:to AP2
Acc-S2	Eth-trunk 1	mode:LACP-static Trunkport :GE0/0/1、GE0/0/2、GE0/0/3 description:to Agg-S1's eth-trunk 2
Acc-S3	Eth-trunk 1	mode:LACP-static Trunkport :GE0/0/1、GE0/0/2、GE0/0/3 description:to Agg-S1's eth-trunk 3
Acc-S4	Eth-trunk 1	mode:LACP-static Trunkport :GE0/0/1、GE0/0/2、GE0/0/3 description:to Agg-S1's eth-trunk 4
AC1	GE0/0/1	Description:to Agg-S1’s GE0/0/2
CORE-R1	GE0/0/1	Description:to Agg-S1’s GE0/0/1

设备	接口	配置内容
Agg-S1	Eth-trunk 1	mode:LACP-static Trunkport :GE0/0/3、GE0/0/7、GE0/0/8 description:to Acc-S1's eth-trunk 1
	Eth-trunk 2	mode:LACP-static Trunkport :GE0/0/4、GE0/0/9、GE0/0/10 description:to Acc-S2's eth-trunk 1
	Eth-trunk 3	mode:LACP-static Trunkport :GE0/0/5、GE0/0/11、GE0/0/12 description:to Acc-S3's eth-trunk 1
	Eth-trunk 4	mode:LACP-static Trunkport :GE0/0/6、GE0/0/13、GE0/0/14 description:to Acc-S4's eth-trunk 1
	GE0/0/1	Description:to CORE-R1’s GE0/0/1
	GE0/0/2	Description:to AC1’s GE0/0/1

配置方案 (2)

2. 基础业务-VLAN配置，采用基于端口的划分方式，详细内容如下:

设备	接口	类型	配置内容
Acc-S1	Eth-trunk 1	Trunk	PVID:100 Allow-pass VLAN 1、100、101
	E0/0/10		PVID:101 Allow-pass VLAN 1、101
	E0/0/11		
Acc-S2	Eth-trunk 1	Trunk	PVID:100 Allow pass VLAN 2、100
	其他接口	Access	Default VLAN 2
Acc-S3	Eth-trunk 1	Trunk	PVID:100 Allow pass VLAN 3、100
	其他接口	Access	Default VLAN 3
Acc-S4	Eth-trunk 1	Trunk	PVID:100 Allow pass VLAN 4 、100
	其他接口	Access	Default VLAN 4

设备	接口	类型	配置内容
Agg-S1	Eth-trunk 1	Trunk	PVID:100 Allow-pass VLAN 1、100、101
	Eth-trunk 2	Trunk	PVID:100 Allow pass VLAN 2、100
	Eth-trunk 3	Trunk	PVID:100 Allow pass VLAN 3、100
	Eth-trunk 4	Trunk	PVID:100 Allow pass VLAN 4 、100
	GE0/0/2	Access	Default VLAN 101
	GE0/0/1	Access	Default VLAN 102
AC1	GE0/0/1	Access	Default VLAN 101

配置方案 (3)

3. 基础业务-IP地址配置，终端与AP采用DHCP方式，设备采用静态配置，详细内容如下：

设备	接口	地址/掩码
Agg-S1	VLANif1	192.168.1.254/24
	VLANif2	192.168.2.254/24
	VLANif3	192.168.3.254/24
	VLANif4	192.168.4.254/24
	VLANif100	192.168.100.254/24
	VLANif101	192.168.101.254/24
	VLANif102	192.168.102.2/30
CORE-R1	GE0/0/1	192.168.102.1/30
	GE0/0/0	PPPoE自动获取
	Loopback0	1.1.1.1/32

设备	接口	地址/掩码
Acc-S1	VLANif100	192.168.100.1/24
Acc-S2	VLANif100	192.168.100.2/24
Acc-S3	VLANif100	192.168.100.3/24
Acc-S4	VLANif100	192.168.100.4/24
AC1	VLANif101	192.168.1.101/24

配置方案 (4)

4. 基础业务-IP地址分配方式配置，关于DHCP的详细内容如下：

网段	其他参数	备注
192.168.1.0/24	Gateway:192.168.1.254 DNS:192.168.1.254	Agg-S1为DHCP Server
192.168.2.0/24	Gateway:192.168.2.254 DNS:192.168.2.254	Agg-S1为DHCP Server 给打印机（1）以及FTP分配固定IP地址
192.168.3.0/24	Gateway:192.168.3.254 DNS:192.168.3.254	Agg-S1为DHCP Server 给打印机（2）分配固定IP地址
192.168.3.0/24	Gateway:192.168.4.254 DNS:192.168.4.254	Agg-S1为DHCP Server 给打印机（3）及网络管理员分配固定IP地址
192.168.101.0/24	N/A	Agg-S1为DHCP Server 不分配AC所占用的地址（192.168.101.1）

配置方案 (5)

5. 基础业务-路由配置，由于网络规模较小且网元数量较少，采用静态路由方式，详细内容如下：

设备	路由配置	备注
Acc-S1	0.0.0.0 0 192.168.100.254	为了让网络管理员可以跨网段访问二层交换机。
Acc-S2		
Acc-S3		
Acc-S4		
AC1	0.0.0.0 0 192.168.101.254	为了让管理员可以跨网段访问AC1。
Agg-S1	0.0.0.0 0 192.168.102.1	访问Internet的流量所匹配的路由。
CORE-R1	192.168.0.0 20 192.168.102.2	核心路由器访问内网，该路由为聚合后的路由。
	默认路由	指向外网接口。

配置方案 (6)

6. 网络管理配置，采用Telnet远程管理，认证方式为AAA，详细内容如下：

设备	管理方式	认证方式	备注
Acc-S1	Telnet	本地AAA	用户名和密码应该足够复杂且不一致，同时需要做好记录工作。
Acc-S2			
Acc-S3			
Acc-S4			
Agg-S1			
CORE-R1			
AC1			
AP1&AP2	AC集中控制和管理	N/A	N/A

7. 网络出口配置

设备	接口	接入方式	NAT方式	备注
CORE-R1	GE0/0/0	PPPoE	Easy IP	用户名：PPPoEUser123 密码：Huawei@123

配置方案 (7)

8. WLAN配置，按照WLAN规划内容进行配置即可。

9. 安全相关配置，详细内容如下:

模块	相关技术	配置内容
流量监控	Traffic-Policy 、NAT、ACL	1.配置高级ACL，阻止源为192.168.1.0/24，目的为内网业务网段的流量，放通其他流量。配置Traffic-filter引用此ACL，并在接口上应用。 2.配置基本ACL，仅放通源为192.168.1.0/24的流量，并引用到网络出接口的NAT功能上。
网络管理安全	AAA、ACL	配置基本ACL，仅放通源为管理员的IP地址，反掩码为0，并引用到所有被管理设备的VTY接口下。
DHCP安全	DHCP Snooping	在所有接入交换机上开启DHCP Snooping功能，同时配置上行接口为Trusted接口。

小型园区网络调试

1.联通性测试	2.高可靠性能力调试	3.业务性能测试
基础链路对接测试	防环功能测试	业务流量测试
二层互通测试	路径切换测试	访问控制测试
三层互通测试	双机热备测试	

小型园区网络运维

- 项目上线运行之后，就进入到了运维阶段，常见的运维手段包括：
  - 设备环境检查
  - 设备基本信息检查
  - 设备运行状态检查
  - 业务检查
  - 告警处理
- 当网络达到一定规模，可以采用网络管理软件进行管理和运维，提升效率。

小型园区网络优化

- 通过网络优化，能够整体提升网络的可靠性、健壮性，更好的支撑企业业务的发展。常见的优化方案包括但不限于：
  - 设备性能优化，如升级硬件设备、更新设备软件版本等。
  - 网络基础优化，如网络架构优化、路由协议调整等。
  - 业务质量优化，如针对语音、视频业务的优先转发等。
- 应从网络需求出发，结合实际情况制定适合的优化方案。

思考题

1. 园区网的完整生命周期是什么？
2. 管理IP地址的作用是什么？

1. 规划与设计、部署与实施、网络运维、网络优化。
2. 网络管理员管理设备时所使用的IP地址。

本章总结

- 本章介绍了园区网络的概念、类型以及常见技术等。
- 了解园区网络生命周期：
  - 规划与设计
  - 部署与实施
  - 网络运维
  - 网络优化
- 结合之前课程内容，着重介绍了园区网络的规划设计与部署实施，完成一张小型园区网络的搭建。


谢 谢
www.huawei.com