1. 网络安全

- **网络安全**
  - 定义：保护网络系统中的硬件、软件及数据不因偶然或恶意原因受到破坏、更改、泄露，系统能正常运行，服务不中断
  - 五要素（针对数据）：**保密性**（不泄露）、**完整性**（不被篡改）、**可用性**（能正常使用）、**可控性**（传播范围和方式）、**可审查性**（提供调查依据和手段）
- **加密技术的3种类型**
  - 哈希算法：无密钥，单向
  - 对称加密（单钥加密）：1个密钥，密文长度约等于明文长度
    - 优点：加密速度快
    - 缺点：密钥管理量大，对密钥传输信道安全性要求更高，对称加密无法实现数字签名
  - 非对称加密（公钥加密）：2个密钥（1个公钥，1个私钥）
    - 优点：加密解密能力分开，无需事先分配密钥，密钥管理量较少，解决数字签名问题，公网中实现保密通信
    - 缺点：加密解密速度慢
    - 加密过程

- **算法原理**
  - DES（Data Encryption Standard，数据加密标准）：属于对称加密，分组加密多次迭代
  - RSA：属于非对称加密，基于大数分解难题产生的不对称性
- **基于单项校验和的身份认证**
  - 数字签名：设计一个代替手记签名的方案，接收方能够验证发送方身份，发送方不能否认身份，接收方不能伪造发送方的签名
  - 单项校验和：假定有一明文报文P和具有单项性质的函数CK，可以计算出CK(P)，但几乎不能从CK(P)中找出P
  - 争议解决方法原理：CK具有单向性
  - 身份认证（数字签名认证）过程

- **主动攻击 VS 被动攻击**
  - 主动攻击：导致某些数据流的篡改和虚假数据流的产生
  - 被动攻击：不是修改数据，而是截取/窃听未经用户许可的信息
- **DDos攻击**：分布式拒绝服务攻击（Distributed Denial of Service），利用大量受控主机（僵尸网络）向目标发送海量请求，耗尽目标资源导致正常服务中断

---

2. 网络测量

- **流类型**
  - 大象流：数量少但字节数巨大，持续时间长
  - 老鼠流：数量多但字节数极小（如网页请求）
  - 乌龟流：速率极低但持续很久
  - 蜻蜓流：突发性极强，瞬间产生大量数据
- **网络测量定义及应用**
  - 定义：通过一定方法与技术，度量网络运行状态、流量特征、拓扑结构等数据
  - 应用：检测网络故障、测试协议行为、刻画流量特征、评估网络性能
- **测量分类标准与具体类别**
  - 按测量方式：主动测量、被动测量
  - 按测量内容：拓扑测量、性能测量、流量测量
- **主动 vs 被动测量**
  - 主动测量：主动发起探测，将探测分组注入网络，根据测量数据流的传递情况来分析
    - 优点：简单方便（只需要发送测试包在本地观察网络响应接即可），对用户而言很安全（不涉及用户的网络信息）
    - 缺点：增加网络负载，消耗较多计算资源
  - 被动测量：在链路或设备节点上抓包分析
    - 优点：测量的是网络上真正的流量，对所观察点网络的行为有比较完整详细的了解
    - 缺点：容易泄露用户信息，测量范围有限
- **带宽概念**
  - 可用带宽：**端到端**之间未被使用的剩余能带宽（总带宽 - 已用带宽）
  - 瓶颈带宽：**端到端**之间最小的链路带宽（瓶颈链路）所能达到的最大数据传输速率
  - **链路**带宽：**链路上**数据报文的最大传输速率
- **测量方法**：
  - 可用带宽：基于**探测报文间隔模型**的测量方法。假设路径上窄链路和紧链路为同一条且带宽C已知，通过发送间隔 Δ_{in} 和接收间隔 Δ_{out} 之间的差值来推算可用带宽。计算公式：

$A = C(1 - \frac{\Delta_{out} - \Delta_{in}}{\Delta_{in}})$

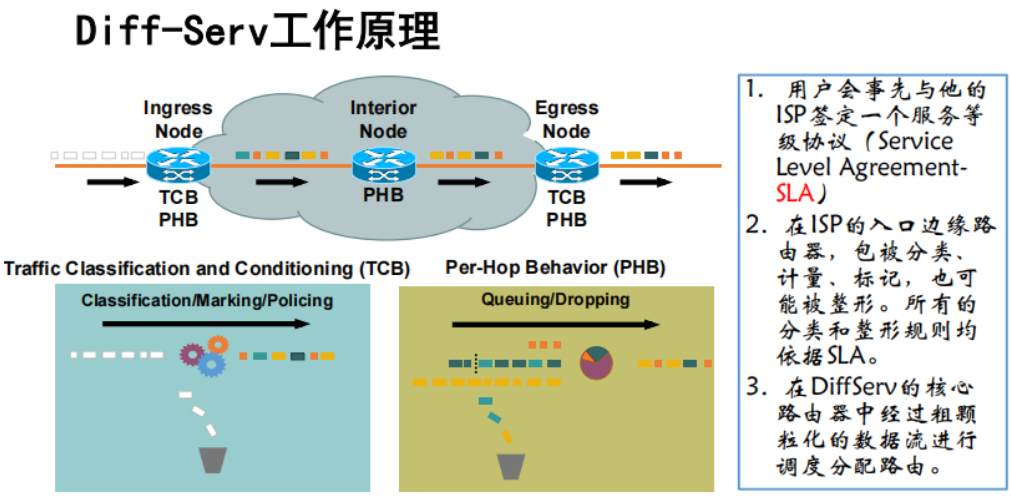
- 瓶颈带宽：基于**包对模型**的测量方法。发送两个足够近的数据报，使其在瓶颈链路处相邻，根据第二个包的大小 L_2 和接收端测量到的时间间隔 t，通过公式 $B_bottleneck = L_2/t$ 计算瓶颈带宽

- **别名解析 (Alias Resolution)**
  - 定义：监测哪些IP属于同一个路由器，从而把同一个路由器的IP地址聚合起来
  - 方法
    - 使用DNS域名反向查询，假定同一路由器的多个接口有相同域名，则对IP地址做反向域名查询可以实现别名解析
    - 使用IP数据报的ID字段，由相同路由器发出的IP数据包中的ID字段值通常是唯一且连续的
    - 基于ICMP消息：TTL超时，ICMP消息源地址为路由器上探测报文的入口地址；端口不可达，ICMP消息源地址为路由器上探测报文的出口地址

---

3. 服务质量

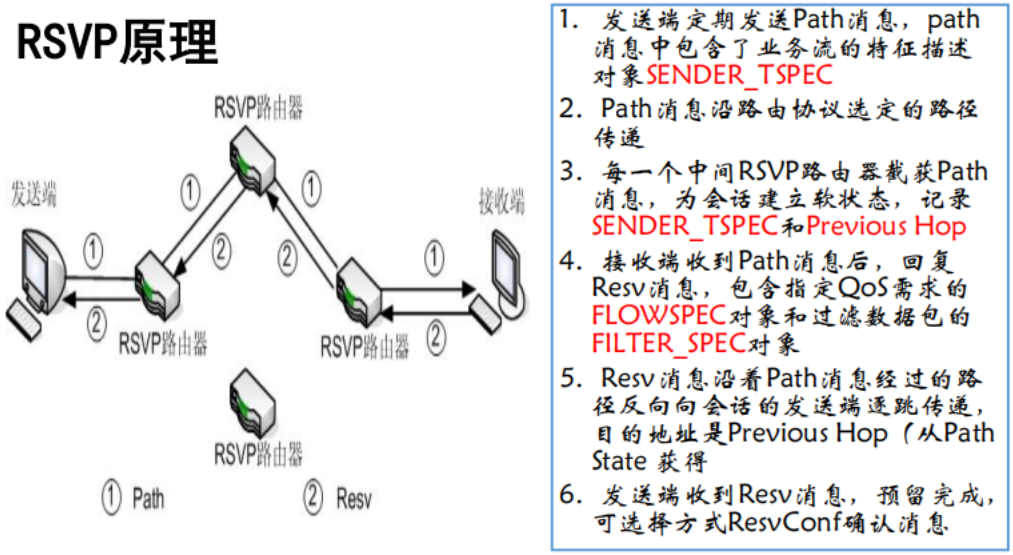
- **QoS服务质量的概念**：网络满足用户服务要求（延迟、抖动、带宽等）的能力，QoS不能创造带宽，但是可以有效地进行网络资源管理
- **IP Qos 的目标**：
  - 避免并管理IP网络拥塞
  - 减少IP报文的丢失率
  - 调控IP网络的流量
  - 为特定用户或特定业务提供专用带宽
  - 支撑IP网络上的实时业务
- **端到端时延包含哪些部分?其中每一部分的具体含义?**
  - 传播时延：数据包从线路上被发送和传播所花费的时间
  - 处理时延：数据包被路由器从入端口接受处理并放到出端口队列所花费的时间
  - 排队时延：数据包在出端口队列到被发送之前排队停留的时间
- **三种 Qos 模型及相互比较?**
  - Best-Effort service 尽力而为服务模型
    - 单一的简单服务模型：网络尽最大的可能性来发送报文，但对时延、可靠性等性能不提供任何保证
    - 网络的缺省服务模型：通过FIFO的队列实现，适用于最大多数网络应用
    - 不属于QoS的范畴：在转发尽力而为的通信时，并未提供任何服务或者传送保证
  - Integrated service 综合服务模型（Int-Serv）
    - 在发送报文前，**通过RSVP信令**向网络申请特定的服务。应用程序首先通知网络它自己的流量参数和需要的特定服务质量请求（带宽、时延等），应用程序一般在收到网络的确认信息，即网络已经为这个应用程序的报文预留了资源后，发送报文
    - 优点：能够提供绝对有保证的QoS
    - 缺点：可扩展性差、对路由器要求较高、不适合短生存期的流
  - Differentiated service 区分服务模型（Diff-Serv）
    - 基本原理



- 优点：扩展性好，简单可实现、具有层次化结构、不影响路由
- 缺点：仅实现粗略的分等级服务、只是是相对优先而不能保证端到端QoS性能、组件分散需要管理

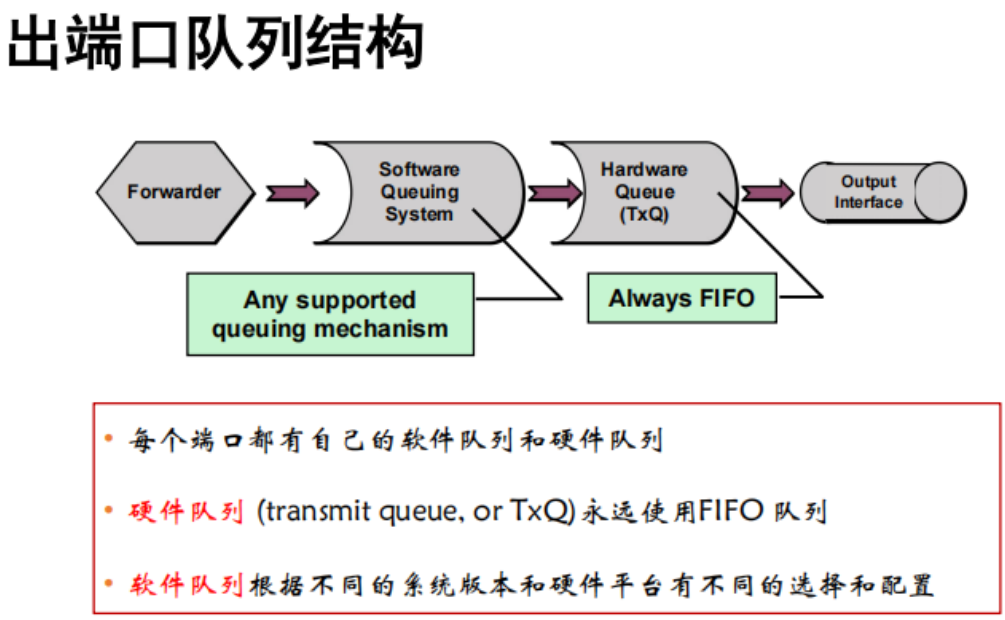
- **Int-Serv 与 DiffServ 对比**

- **RSVP协议（资源预留协议）**：运行在从源端到目的端的每个设备上，可以监视每个流，以防止其消耗资源过多。能够明确区分并保证每一个业务流的服务质量，为网络提供最细粒度化的服务质量区分
  - 工作原理



  - 协议特点
    - 工作在IP协议之上，属于传输层

        - 是一个网络控制协议，本身并不处理传输层数据
        - 可以在单播、组播网络通信应用中进行资源预留
        - 是一个单向的资源预留协议
        - 面向接收端的资源预留协议，由会话的接收端发起资源预留请求
        - 对不支持RSVP的路由器透明
- **什么是队列调度?路由器出端口队列结构?**
    - 队列调度：不同等级的分组放入不同的队列中，路由器按照一定的队列调度算法，决定从哪个队列中取出数据分组进行服务
    - 队列机制的三个组成部分：分类、排队策略、服务策略
    - 路由器出端口队列结构：



- **常见队列调度算法拥塞控制机制的原理?及各自的优缺点?**
    - FIFO先进先出：先进入队列的数据包先出
        - 优点：算法简单快速，平台支持性好
        - 缺点：流之间分配不公平，不分优先级无法保障重要业务
    - PQ优先级队列：按照优先级发送数据包
        - 优点：算法简单，可以保障高优先级的QoS
        - 缺点：算法欠公平，高优先级队列可能会饿死其他队列
    - CQ定制队列：提供16个FIFO队列供用户自定义，使用轮询算法，每次调度每个队列只能发送一定量的数据
        - 优点：可以保障每种等级的带宽，防止不同等级的业务被饿死
        - 缺点：在单一队列内仍是FIFO，需要在每一跳手动配置分类，调度时容易产生比较高的时延抖动，16个级别分类比较粗无法实现精细化控制，轮询调度无法严格保证最高优先级业务的QoS
    - WFQ加权公平队列：引入权重，数据包排列为不同的流进行调度，权重高的流会得到相应比例的带宽
        - 优点：配置非常简单，兼顾所有流的带宽，优先保障高优先级流的QoS
        - 缺点：不能对流进行客户化定制，不能提供固定的带宽保障
- **整形与限速**：
    - **整形**：将突发流量缓存起来，等流量下降后在发送，因此流量曲线更平滑
    - **限速**：直接丢弃或降级转发超过带宽限制的突发数据包
    - **令牌桶**（判断当前流量是否超过了带宽限制）：只有桶里有令牌才能发数据，否则视为超过额定带宽。会按照一定速度往桶里添加令牌，添加速度控制了用户流量带宽

4. 区块链

- **区块**
  - 定义：记录所有经数字签名后的共享信息（账本中的纸）
  - 作用：以hash链的方式保存业务记录集合，保存前序块的哈希值
- **Hash链**
  - 定义：逻辑上由前序块的哈希值串联起来的链条，将前序块的hash值作为下一个区块内容的一部分
  - 作用：保证时序性和不可篡改性
- **区块链**
  - 定义：用密码技术将共识确认的区块按顺序追加形成的分布式账本，一种分布式交易验证和数据共享技术
  - 特征：去中心化、透明性、不可篡改、可追溯
  - 核心价值：信任模型：信任人->信任技术、数据保护：产生者->拥有者、治理模式：人治->规则治
  - 基本工作原理：单点出块、广播传输、交叉验证、共同存储
- **区块链与分布式数据库的区别**
  - 分布式数据库：由中心管理、可增删改查、节点间相互信任、节点间数据一般不相同
  - 区块链：去中心、多方写入、不可删改、各个节点并不相互信任、容错性更好、各个节点数据内容保持一致
- **生活中常见的区块链应用案例**
  - 供应链金融，解决虚假贸易
  - 医疗健康，共享电子档案防篡改
  - 版权保护，保护数字资产版权

---

5. Web、CDN、P2P

- zipf定律：互联网上20%的内容吸引了80%的流量（局部访问性），排名第i位的内容访问次数 y_i ~ $1/i^s$，在log-log坐标中是一条直线
- Web缓存
  - 工作原理：被动缓存，将常访问的内容放在离用户近的地方（代理服务器），后续请求不需要再到源服务器
  - 缺点：存在版权问题、很多类型的内容不能缓存（动态数据、加密数据）
- CDN内容分发网络
  - 工作原理：主动发布内容，把内容拷贝到不同地域的多台服务器，减轻服务器负载，提升用户感知质量（利用 **DNS 重定向**实现服务器选择；IPv6任播就近访问服务器）
  - CDN服务器选择的常见策略：负载轻的、性能高的、距离客户端近的、成本低的
- 缓存容量受限时该如何管理(缓存内容选择和替换策略)缓存
  - 部分缓存，如之缓存视频的开头部分
  - 缓存替换：缓存请求大于两次的视频（LRU淘汰最近最少使用、LFU淘汰最不经常使用）
- P2P 网络：去中心化，对等网络，每个节点既是客户端也是服务器，从而与网络中其他节点直接交换数据
- BitTorrent 协议
  - 特点：文件分块传输，一个节点可以从多个节点下载文件，节点直接交换块的信息
  - 原理：需要多客户端来分担负载，所以采用以牙还牙（tit for tat）的激励机制鼓励分享，使用Rarest First块选择方式
- 基于DHT的P2P网络（结构化P2P）：给定文件ID，定位出保存该ID对应的文件的服务器

- 特点：Chord搜索，根据资源ID高效搜索服务节点
- DHT：分布式哈希表，把Hash表按一定的规律分布到节点上
- Chord 原理：将节点和键映射到一个逻辑环上，通过finger table实现 O(logN) 的查找效率