具体构成描述:终端(运行应用程序)称主机或端系统。 端系统通过通信链路和交换设备连 接到一起:通信链路(光纤,铜线,电磁波)的传输速率或带宽以比特每秒(bps)度量:交换设备从 它的一条入链路接收到达的分组,并从它的一条出通信链路转发该分组,交换设备包括路由 器(通常用于网络核心)和链路层交换机(通常用于接入网)。 当端系统向另一端系统发送数 据时,发送端将数据分段,并为每段加上首部字节,形成分组, 从发送到接收端系统,一个分 组所经历的一系列通信链路和分组交换机称为通过该网络的路径. 端系统通过因特网服 务提供商(Internet Service Provider)接入因特网,每个 ISP 是由分组设备和通信链路组成的网 络.各 ISP 为终端提供了各种类型的网络接入,也为内容提供者提供因特网接入服务,将 web 站点接入因特网,每个 ISP 都是**自治**的,ISP 之间是互联的,包括不同层次,本地/地区/全球 ISP, 终端分组交换机和其他因特网部件都要运行一系列协议,这些协议规定了设备之间通信需 要遵循的规则,控制因特网中信息的发送和接收(终端与终端,终端与交换设备,交换设备与交 换设备)。一个协议定义了在两个或多个通信实体之间交换的报文格式和次序,以及报文发 送和/或接收一条报文或其他事件所采取的动作。 因特网标准由因特网工程任务组 (Internet Engineering Task Force)研发.IETF 的标准文档称为请求评论(Request For Comments),TCP 和 IP 是因特网中最核心的两个协议,因特网协议统称为 TCP/IP 协议族, 因特网定义一:由一群遵循 TCP/IP 协议的 ISP 按照松散的层次结构组织而成的网络的网络 该定义对于通信功能的实现有指导作用:ISP 内部和之间互联. 因特网的几个特点:因特网 是网络的网络 不存在严格的是次结构 没有统一的管理机构

功能描述 网络國定义 二·为分布式应用提供通信服务的基础设施 该定义对于服务接口的定 以有指导作用·有序可靠的数据交付服务&不可靠的数据交付服务 因特网分布式应用程 南涉及多台相互交换数据的端系统并不运行在网络核心的分组交换机由 网络网提供给 应用程序的服务 捺口为一组用于在因结网上发送和捺收数据的**应用程序编程捺口** (Application Programming Interface)该API 规定了运行在一个端系统上的软件请求因转网 基础设施向运行在另一个端系统上的特定目的地软件交付数据的方式。 网络边缘

端系统也称为主机,因为它们容纳(即运行)应用程序,主机有时进一步被划分成两类,客户和 服条器

※ 入回 接 入回 是指路终端连接到其边缘路由器的物理链路 边缘路由器是端系统到任何其 他远程端系统的路径上的第一台路由器. 宽带住宅接入有两种最流行的类型:数字用户线 和电缆. 数字用户线(Digital Subscriber Line) 由电话公司提供,使用已有的数字电话线(双 绞铜线,每户一条线),同时(从高频到低频)承载高速下行信道(<24Mbps,<10Mbps)、中速上行 信道(<2.5Mbps,<1Mbps)和普通的双向电话信道:电话公司为 ISP:DSL modem 调制解调器转 换模拟信号和数字信号,得到数字数据后将其转换为高频音;splitter 在用户侧合并/分离话音 和数据:DSLAM 在 ISP 侧汇聚/分离多条 DSL 线路:近距离 5~10 英里内接入。 电螺因特网 接入利用有线电视公司现有的有线电视基础设施、数据和电视信号在同一条电缆中传输。同 时应用了光纤网和同轴电缆网,称为混合光纤同轴电缆(Hybrid Fiber Coax),同时还有电缆调 制解调器(cable modem),splitter 和电缆调制解调器端接系统(CMTS)构成接入网;几百~几千 户家庭共用一条电缆:一个重要特征是共享广播媒体:下行 30Mbps 上行 2Mbps(DOCSIS 2.0 定义了下行 42.8Mbps 和 30.7Mbps 上行) 企业接入以太网(Ethernet):公司,学校以及有较 多终端的家庭使用局域网(LAN):以太网是最为流行的接入技术,使用双绞铜线;以太网交换机 及链路构成接入网:10Mbps,100Mbps,1Gbps,10Gbps 无线 接入 无线局域网 WiFi:IEEE802.11 无线 LAN 接入,公司或个人提供基站(接入点), 将移动终端连接到有线网络, 终端与基站几十米内,基站位于有线网络上,11Mbps,54Mbps,450Mbps;无线局域网是共享的, 无线接入广域无线接入由移动通信公司提供,使用现有的蜂窝电话网络:基站可

为数万米半径内的用户提供无线接入服务:3G 2Mbps.4G 下行 100Mbps 上行 20Mbps. 物理媒体:对于每个传输器-接收器对,设备通过物理媒体传播电磁波或光脉冲来发送该比特 物理媒体可以有各种形状和形式,并且一条路径上的每个传输器-接收器对之间的物理媒体 可以不同。 物理媒体划分为导引型媒体(沿固体媒体传播,铜线,光纤)和非导引型媒体(在空 间自由传播.电磁波). 双绞线:两根绝缘的铜导线,3类10Mbps,5类100Mbps-1Gbps,6类 10Gbps, 速率取决于线的租细与传输距离;常用在电话线和网线. 同轴电缆: 铜芯和网状 屏蔽层组成一对同心的导电体,有线电视电缆,能作导引型共享媒体. 光纤:能导引光脉冲 的玻璃纤维:几十~几百 Gbps:抗电磁干扰.长距离传输低衰减,难窃听.低误码率。 电磁波:无 物理线路,并具有穿透墙壁、提供与移动用户的连接以及长距离承载信号的能力;蓝牙 (2.4GHz,10m),WiFi(2.4GHz,几十米),红外(室内短距离),陆地微波(2GHz,长距离),卫星(2GHz,长 距离大范围),可见光(1-2km,50Gbps),

网络核心

网络核心是由路由器和通信链路组成的网状网络,网络核心的任务是将数据包从发送侧的 边缘路由器传送到接收侧的边缘路由器.基本问题:数据包如何在网络核心中高效地传递(分 组传输延迟小,网络吞吐量高)? 移动数据的两种基本方法:电路交换(独占)/分组交换(复用) 分组交换:主机将要传输的数据分段,并组装成一系列分组。 交换是在传输路径上,交换设 备从一条链路上接收分组,将其发送到另一条链路上。 存储转发:在交换设备在接收到完 整的分组后 才能够开始转发 内列亚亚识分组以第三该链路的最大读率传输通过通信链 路 | /hit 的分组) / R/hit/s 链路传输速率 | 种 P 个分组经付 N 各链路总括时(P+N-1)| /R 当 P 元 大干 N 时存储转发不会引入过多延迟 每个分组交换机有多条链路与之相连 对于每条 相连的链路 该分组交换机且有一个输出缓存(输出队列) 它用于存储路由器准备发往那条 链路的分组 分组还要承受输出缓存的推队延迟 这是变化的 取决于网络拥靠程度 因 为输出链路缓存大小是有限的 一个到达的分组可能发现该缓存已被其他签待传输的分组 完全充满了,在此情况下出现分组丢失(丢包),到达的分组或已经排队的分组之一将被丢弃。 当大量分组集中到达时,排队延迟和丢包较严重。 当一个分组到达网络中的路由器时,路 由器检查该分组的目的地址的一部分,并向一台相邻路由器转发该分组,每台路由器具有一 个转发表(选路 routing),用于将目的地址(或一部分)映射成输出链路,交换设备按照转发表。 将分组移动到对应的输出链路(转发 forwarding). 因特网具有一些特殊的路由选择协议。 用于自动协设置议些转发表(例如最短路径)

电路交换: 电话网采用电路交换.通话前完成两部电话机间的电路接续,童话结束后释放整 冬由路 太后是葡萄瓷源和油片瓷源 由路交换网络中在通信期间 葡萄了端系统间通信瓷源 (缓存,链路传输速率)(在分组交换网络中,报文按需使用这些资源,后果是可能不得不等待排 队)。 网络创建时预留了恒定的速率,则发送方能以确保的恒定速率向接收方传送数据 区 : 链路 link 是物理媒体或称信道,可以通过某种方式划分为若干条独立的子信道,物理媒体 中的一条子信道称为一条电路 circuit. 当两台主机要通信时该网络在两主机之间创建一 条专用的端到端连接. 多路复用 multiplex:類分复用(Frequency Division Multiplexing)的 链路为每个连接专用一个频段,频段的宽度称为带宽(bandwidth);时分复用(Time Division Multiplexing)中时间被划分为固定区间的帧。每帧被划分为固定数量的时隙,帧的每个时隙 为某个连接专门使用,TDM 电路的传输速率=帧速率+一个时隙中的 bit 数。 同样的链路容 量,分组交换允许支持更多的用户,且轻负载时,分组交换可以更快服务用户,电路交换是固定 分配/信道使用模式固定,而分组交换为按需分配/信道使用模式不固定

电路交换的缺点:静默期转用电路空闲而效率较低:创建端到端电路和预留端到端带宽是复 杂的,需要复杂的指令软件以协调沿端到端路径交换机的操作。 分组交换的缺点;可能产 生延迟/丢包,需要响应的协议解决。 分组交换不适合实时服务,因为其端到端延迟是可变 和不可预测的(可能丢包),并且不能保证服务质量(带宽)。 分组交换的优点:提供更好的带 宽共享,资源利用率高;比电路交换更简单(不需要建立电路)更有效,实现成本更低。 因特网 采用分组交换:分组交换适合突发流量:传统因特网应用具有突发通信的特点

网络的网络:网络核心的任务是将全球的本地 ISP 连接在一起。 因特网生态系统:接入 ISP:

地区 ISP:第一层 ISP:对等链路:因特网交换点(Internet eXchange Point):多个 ISP 共同对等的 地方:存在点(Point of Presence):低层 ISP 接入高层 ISP 的地方:多宿(multi-home):一个低层 ISP 可以接入多个高层 ISP:对等:相同层次的 ISP 连到一起以试图绕过上层:内容提供商网络 (Content Provider Network).

分组交换中的延迟、丢包和吞吐量

延迟(分组从源终端到达目的终端的时间)。丢包率(未成功交付到目的终端的分组比例)。吞吐 量(单位时间内网络成功交付的数据量)是衡量网络性能的主要指标,均与负载有关 分组延迟:当分组从一个节点沿着这条路径到后维节点,该分组在沿途的每个节点经受了几

种不同类型的延迟:节点处理延迟,排队延迟,传输延迟,传播延迟,累加起来是节点总延迟. 节点处理延迟:检查错误和确定输出链路(微秒或更低). 排队延迟:分组在输出缓存等待传 输取决于链路负载大小(差异很大,豪秒到微秒) 传输延迟;L bit 长度的分组,在 R bps 的链 路上,传输延迟是 L/R,是分组序列化时间,(毫秒到微妙,取决于链路速率) 传播延迟:从该链 路的起点到终点传播所需要的时间,速率取决于物理媒体,范围是 2/3~1 倍光速,传播延迟是 d/s,d 距离.s 速率(百豪秒到几微秒,取决于链路长度)。 总延迟 d nodal=d proc(处 理)+d queue(排队)+d trans(传输)+d prop(传播),成分作用可能差异很大(如卫星和大学), 排队延迟和丢包;不同于其他三个,排队延迟对不同的分组可能是不同的,表征排队延迟一般 使用统计量测度 a 表示分组到达队列的平均速率(分组每秒),R 传输速率 bps,L 分组大小 bit.La/R 称为流量强度。 La/R>1.排队延迟趋向无穷大,设计系统时流量强度不能大于 1:La/R->1,平均排队延迟较大,La/R~0,平均排队延迟比较小, 到达流量的性质影响排队延 识 分组 周期性到达 平均延迟低 率发一次到达多个则平均延迟高 平均推队延迟和流量 强度的定性 关系 随着流量强度 接近于 1 平均排队 延迟迅速增加 该强度 小量的增加 格异勒 延迟大得多的增加 输出队列只有有限的容量 尽管排队容量极大的依赖路中器的设计和 成本 因为该排队容量是有限的 随着流量强度接近 1 排队延迟并不实际趋向于穷大 相反 到 达的分组将发现一个满的队列 路中器 垂直该分组 即该分组将 要 金 分组 医生的 份額 随 着流量强度的增加而增加. 一个节点的性能常常不仅根据延迟来度量,而且根据分组丢失 的概率来度量。 队列太短丢包率增大,队列太长排队延迟增大(间接丢包) 端到端延迟·分组传输路径上所有节占的节占延迟之和 对端到端延迟龄成的应用·高度

些重要的延迟:作为它协议的一部分,希望向共享媒体传输分组的端系统可以有意地延迟它 的传输以与其他端系统共享媒体 媒体分组化延迟 经 IP 语音应用中 发送方在相因特网 传递分组之前必须首先用编码的数字化语音请查—个分组 这个时间为分组化延迟

计算机网络中的吞吐量: 瞬时吞吐量是给定时刻的传输速率(bos) 平均吞吐量是较长时间内 的传输速率。 吞吐量是 $\min \{R_1, R_2, ..., R_n\}$,即瓶颈链路的传输速率带宽限制了端到端 吞吐量、公共链路吞吐量是min {R_c,R_s,R/链路数}。 端到端吞吐量与瓶颈链路的速率,以 及链路上的负载有关,特别是,如果许多其他的数据流也通过这条链路流动,一条具有高传输 速率的链路仍然可能称为传输的瓶颈链路。

协议是次及其服务模型

协议要素:网络协议定义了在通信实体间交换报文的格式和次序,在发送/接收报文或其他事 件后采取的动作,掌握计算机网络知识的过程,就是理解网络协议的构成,原理和工作的过程 分层的体系结构:每个层次与其下面的层次结合在一起,实现了某些功能,服务。 系统分层 将系统按功能划分为一系列水平的层次,每一层实现一个功能(服务)。 层次间关系:每一层 的功能实现都要依赖其下各层提供的服务。 每个层次通过在该层中执行某些动作/使用 直接下层的服务提供服务。 分层的好处:系统分层易于处理复杂的系统,显示的层次结构 易于确定系统的各个部分及其相互关系:模块化简化了系统的维护和升级(改变某层服务的 实现方式对于其他层次没有影响,也就是透明)。 网络设计者以分层的方式组织协议以及 实现这些协议的网络硬件和软件,每个协议属于这些层次之一,某层向它的上一层提供的服 务,即所谓一层的服务模型。一个协议层能够用软件、硬件或二者的结合来实现(应用层和 运输层是端系统的软件实现、物理层和数据链路层是在与链路相联系的网络接口卡中实现。 网络层是硬软件混合). 一个第 n 层协议也分布在构成该网络的端系统、分组交换机和其 他组件中,也就是说第 n 层协议的不同部分常常位于这些网络组件的各部分中. 协议分层 具有概念化和结构化的优点,分层提供了一种结构化方式来讨论系统组件,模块化使更新系 统组件更为容易. 分层的缺点:一层可能冗余较低层的功能:某层的功能可能需要仅在其 他某层才出现的信息,这违反了层次分离的目标. 各层的所有协议被称为协议栈:物理层、 链路层、网络层、传输层和应用层. 应用层:网络应用程序及它们的应用层协议存留的地 方,分布在多个端系统上,在应用程序之间传输应用特定的报文(message); FTP.SMTP.HTTP 传输层:在应用程序(进程)的网络接口间传输报文段(segment);TCP,UDP. 网络层:在源主 机和目的主机(终端-终端)之间传输数据报(datagram)(PPT 称为分组 packet):IP,routing protocols. 链路层:在相邻设备(节点)之间传输帧(frame);PPP.Ethernet. 物理层:在物理媒 体上传输比特。 ISO/OSI:应用层表示层,会话层,运输层...:这两层留给应用程序开发者处理。 网络功能的分布式实现:某一层网络功能,需要该层实体(分布在不同的节点)协同完成:协同 计算要求功能实体间交互信息:载体(报文),约定(报文格式/语义规定),传输方式(封装/解封装) 封装:与端系统类似,路由器和链路层交换机以多层次的方式组织他们的网络硬件和软件,而 路由器和链路层交换机并不实现协议栈中的所有层次、链路层交换机实现了物理层和链路 层 路中器实现了物理层链路层和网络层 主机实现了所有五个层次 汶佐现了村装的概念 网络安全·恶音软件/每壶 锯中) DoS 腹探/容昕) 伪装(IP 欺骗 重故)

第2音・応用屋

网络应用是计算机网络存在的理由 应用异体设度理·

研发网络应用的核心,是编写将在不同端系统上运行的分布式程序,并能通过网络相互通信, 重要的是 你不需要也不能写在网络核心设备加路中器或链路层交换机上运行的软件 路应 用软件限制在端系统的方法极大方便了的网络应用程序的迅速研发和部署

网络应用程序体系结构: 开发应用程序之前,首先要决定采用什么网络应用架构,应用程序 体系结构即显不同于网络的体系结构 在应用程序研发者的角度者 后者是固定的 并为应用 程序提供了特定的服务集合 网络应用程序体系结构中应用研发者设计 规定了如何在各种 端系统上组织该应用程序 现代网络应用程序的两个主流体系结构·**宏户-服务器体系结** 构和对等(P2P)体系结构。 在客户-服务器体系结构中,有一台总是打开(在线)的主机上面 运行着服务器程序(server),它服务于来自许多其它称为客户的主机的请求。客户-服务器 体系结构的特点:服务器主机具有固定的、周知的地址:客户只与服务器通信,客户之间不直 接通信. 客户:用户终端上运行一个客户程序(client),需要时与服务器程序通信,请求服务. 客户机使用动态地址,通常不会总是在线. 客户-服务器架构的资源集中:客户-服务器架 构将任务合理分配到客户端和服务器端资源,复杂的计算和管理任务交给网络上的高档机 器(服务器主机),需要频繁与用户打交道的任务交给前端较简单的客户机,通常是基础设施幂 集的应用(服务提供商需提供服务器集群,高速网络等设施,投入巨大).优点:资源发现简单.缺 点:集中式计算带来的问题如服务器扩容压力,网络流量不均衡,响应延迟长. 在一个 P2P 体系结构中,没有总是在线的服务器主机,对位于数据中心的专用服务器具有最小的(或者没 有)依赖,任意一对端系统(对等方)可以直接通信,对等方使用动态地址,间歇地连接在因特网 上,每个对等方即可请求服务,也可以提供服务。 P2P 特性:自扩展性:是成本有效的,因为通 常不需要庞大的服务器基础设施和服务器带宽. P2P 应用面临的挑战:ISP 友好(大量上载 给 ISP 带来巨大压力);安全性:激励(说服用户自愿像应用提供带宽、存储和计算资源)。 P2P 架构资源分散:任何终端都可以提供资源(服务),优点:易于扩容、均衡网络流量,缺点:资源发 现困难,社会问题(版权、安全性)

进程通信:进行通信的实际上是进程而不是应用程序:进程是主机上运行的程序,在分布式应

用中,不同终端上的进程需要通信. 进程通信的方法:在两个不同端系统上的进程,通过跨 越计算机网络交换报文而相互通信。同一主机内使用 OS 提供的进程间通信机制。 一次确 定的通信会话中,对每对通信进程,主动发起请求的进程为客户,另一个在会话开始时等待联 系的(接受请求的)进程为服务器。 进程通过一个称为套接字(socket)的软件接口向网络发 送报文和从网络接收报文.套接字是同一台主机应用层与运输层的接口.也称为应用程序和 网络之间的**应用程序编程接口(API)**.背后的设计考虑为应用程序只需调用网络服务,而不需 要知道服务的实现细节. 为了标识该接收进程,需要定义两种信息:主机的地址:定义在目 的主机中的接收进程标识符,在因特网中,主机用其 IP 地址标识,接收进程由目的端口号标识. 众所周知的端口号被分配给服务器,成为服务的标识(HTTP 80,SMTP 25)

可供应用程序使用的传输服务:从四个方面对应用程序的服务要求进行分类:可靠性、吞吐量、 延迟、安全性. 可靠性:确保数据从一端正确、完全地交付给应用程序的另一端:多媒体应 用容忍丢失,文件传输要求完全可靠的数据传输。 吞吐量:运输层协议能够以某种特定的 速率提供确保的可用吞吐量;带宽敏感的应用(多媒体应用)要求最低可用带宽,弹性应用(电 子邮件,文件传输与 web 传送)适应各种可能的带宽。 **延迟:**例如发送方注入进套接字中的 每个 bit 到达接收方套接字不迟于多少时间,交互式实时应用程序(因特网电话、虚拟环境、 电话会议和多方环境)要求延迟保证,非实时的应用对延迟不敏感。 安全性:比如运输协议 能够加密发送进程传输的所有数据,然后在交付接收进程之前解密:也包括数据完整性和端 点鉴别.

因特网提供的传输服务:因特网(更一般的是 TCP/IP 网络)提供 UDP 和 TCP 两个传输层协议 TCP 服务模型包括面向连接服务(保证传输顺序)和可靠数据传输(不出错)服务·按手计后创 建一个 TCP 连接 连接是全双下的并且结束发送时必须拆除连接·无差错。按顺序交付所有 发送的数据 还且有流量控制(发送进程不全压烧炼收进程)和拥塞控制(招载时抑制发送进 程\机制(不一定对诵信讲程带来直接好外 伯能对因结网带来整体好外:必平共享网络带雷) 不提供·及时性、最低带宽保证 安全性 TCP 的加强版本安全套接字层(SSI)提供了 TCP 所 能提供的一切以及讲程到讲程的安全性服务(加索 数据宗整性和端占奖别) 这种强化是在应 用层实现的。 UDP 是一种不提供不必要服务的轻量级运输协议,它仅提供最小服务;无连 接,不可靠数据传输(可能无法到达或乱序到达),没有拥塞控制机制(发送进程可以以任何速 窓向网络岸注入数据\没有安全性 无论 LIDP 还是 TCP 都不提供。吞吐量和定时保证 Ø 用程序被良好设计以对付这种保证的缺乏。 能够容忍某些丢失但要求达到一定最小速率 才能有效工作的应用(如因特网电话)使用 UDP 以设法避开 TCP 的拥塞控制机制和分组开销 用具体设 皮用具体设定义了法行在不同端系统上的应用程序进程如何相互传递报文 结 别是应用屏协议定义了·交换的报文类型(请求/响应)·各种报文类型的语法(各个字段以及字 段如何描述);字段的语义(包含的信息的语义);一个进程何时以及如何发送报文,对报文进行 相应的规则。 有些应用层协议是由 RFC 文档定义的,因此他们在公共域中,一些别的应用 层协议(如 Skype)是专用的,有意不为公共层使用. 区分网络应用和应用层协议是重要的, 应用层协议只是网络应用的一部分.

HTTP 概况:web 的应用层协议是超文本传输协议(Hyper Text Transfer Protocol). HTTP 由 客户程序和服务器程序共同实现,二者分布在不同的端系统通过 HTTP 报文会话. Web 页 面(也叫文档)是由对象组成的,一个对象只是一个文件(HTML,JPEG),多数 web 页面含有一个 HTML 基本文件以及几个引用对象(一个有 1 个 HTML 基本文件和 5 个图形的 web 页面有 共 6 个对象).HTML 基本文件通过对象的 URL 地址引用页面中的其他对象.每个 URL 地址由 两部分组成;存放对象的服务器主机名(www.someSchool.edu)和对象的路径名 (/someDepartment/picture.gif).Web 浏览器实现了 HTTP 的客户端.用于请求.接收和显示 web 对象,web 服务器实现了 HTTP 的服务器端,它应客户请求发送对象,每个用 URL 寻址, HTTP 定义了 web 客户向 web 服务器请求,以及服务器向客户传送 web 页面的方式。 HTTP 使用 TCP 服务。 HTTP 服务器并不保存关于客户的任何信息.所以 HTTP 是一个无状态协议 服务器周知端口号:80 网络应用架构:客户-服务器架构。 客户发起到服务器 80 端口的 TCP 连接(客户端创建一个套接字)服务器接受来自客户的TCP连接(服务器端创建一个套接 字)浏览器和服务器交换 HTTP 报文(通过各自的套接字)关闭 TCP 连接(关闭各自的套接字) 非持续连接和持续连接:非持续连接:每个请求/响应对是经一个单独的 TCP 连接发送;持续连 接:所有的请求及其响应经相同的 TCP 连接发送.HTTP 1.0 使用非持续连接.HTTP 1.1 缺省使 用持续连接. 非持续连接例子:客户在 80 端口发起服务器 TCP 连接->客户经连接发送一 个请求报文->服务器接受请求,发送响应->服务器通知断开连接->HTTP 接收响应,连接关 闭,对每个对象重复一次步骤,每个TCP连接在服务器发送一个对象后关闭。 往返时间(RTT) 是指一个短(不考虑传输延迟)分组从客户到服务器然后再返回客户所花费的时间(包括传播。 排队和处理延迟) 非持续连接传输每个对象需要两个 RTT+传输延迟 2N 非持续连接的 缺点:为每个请求的对象建立和维护一个全新的 TCP 连接,每个都要维护缓冲和变量,为服务 器带来负担:每个对象经受 2RTT 的交付延迟;浏览器要打开多个 TCP 连接来获取一个网页。 在采用持续连接的情况下,服务器在发送响应后保持该 TCP 连接打开,后续请求和响应报文 能够通过相同的连接进行传送、特别是一个完整的页面可以用单个 TCP 连接传送:更有甚者。 同一服务器同一客户的多个页面使用一个连接,一般来说,一条连接经过一定时间间隔仍未 使用 HTTP 就把他关闭.HTTP 默认使用带流水线的持续连接。 无流水线.仅当收到前一个 响应后再发送新的请求 每请求一个对象用时一个 RTT N+2·**液水线** 每解析到一个引用对象 前发送请求可以在一个 RTT 内请求所有对象 3

HTTP 报文格式 ASCII 文本组成 请求报文每行用 crif 最后—行附加—个额外 crif 结束 管 一行为请求行·方法字段(GFT(获取) POST HEAD PLIT DELETE() IIRL 字段(请求对象的标 识) HTTP 版本(自解释):后继的行为首部行:Host(对象所在的主机) Connection close(不持续 连接),User-agent(用户代理浏览器版本),Accept-language(没有这个语言的就发送默认版 木) 额外的 crif 后是实体体(entity body) GFT 时为空 POST 时使用该部分发送用户向搜索引 整提供的关键字来提交表单(但提交表单可以放在 GFT 的 LIRI 里) HFAD 方法令让服务器 相应一个报文但不返回内容(调试用)PUT 方法分许用户上传对象·DFIFTF 分许用户删除服 务器对象, HTTP 响应报文包含初始状态行(协议版本,状态码,相应状态信息),首部行 (Connection:close:Date 服务器产生这个报文的时间:Server:Last-Modified 对象创建或最后 修改的时间;Contest-Length; Contest-Type 对象类型应该由这个而不是扩展名标识),实体体 (主要部分,包含了所请求对象的本身) HTTP 响应状态代码:200 OK:301 Moved Permanently 新的 URL 定义在响应报文的 Location:首部行;400 Bad Request 不能被服务器 理解:404 Not Found:505 HTTP Version Not Supported.

用户与服务器的交互:Cookie:HTTP 无状态简化了服务器的涉及,允许工程师开发高性能 Web 服务器.cookie 允许站点对用户跟踪,以将内容与用户身份联系,或限制用户的访问 cookie 技术的四个组件:HTTP 响应报文的一个 Cookie 首部行;在 HTTP 请求报文中的一个 cookie 首部行;在用户端系统浏览器中保留管理一个 cookie 文件;Web 站点的一个后端数据 库。 例子:A 第一次访问某网站,HTTP 请求到达网站时,网站为其创建:一个 ID,后端数据库中 为该 ID 建立一个表项 服务端:信息保存在服务端的后端数据库,返回 ID 给客户 客户端:信息 发回客户端,保存在cookie 文件中,并随请求报文发送给服务器. cookie 可以标识一个用户, 所以 cookie 可以在无状态的 HTTP 之上建立一个用户会话层,但 cookie 允许网站收集用户 的大量信息 带来隐私问题

Web 缓存:Web 缓存器也叫代理服务器:它是能代表初始 Web 服务器来满足 HTTP 请求的网 络实体,Web 缓存器有自己的磁盘存储空间,并在存储空间中保存最近请求过的对象的副本 用户设置浏览器,所有 HTTP 请求首先发往 web 缓存,浏览器将 HTTP 请求发送给 web 缓存. 对象在 web 缓存中,web 缓存返回对象;对象不在 web 缓存中,web 缓存从原始服务器(请求 中的 host 首部行)获取对象、缓存在本地、然后返回给客户. Web 缓存器既是服务器同时 也是客户,Web 缓存器通常由 ISP 提供,多级 ISP 可能形成多级 ISP 缓存, 部署 Web 缓存器 的原因:大大减少客户请求的响应时间:减少一个机构的接入链路到因特网的通信量(降低了

费用):从整体上大大减低因特网上的 Web 流量.从而改善了所有应用的性能

条件 GET 方法:存放在 Web 缓存器里的对象副本有可能是陈旧的,解决这个问题的机制是条 件 GET:请求报文使用 GET 方法.请求报文中包含一个 If-Modified-Since:首部行。 缓存器在 缓存对象时也存储了最后修改日期. 若对象无更新,条件 GET 的响应报文没有包含该对象 (否则只会浪费带宽),状态行为 304 Not Modified;若有更新,200 OK

文件传输协议:FTP

用户通过一个 FTP 用户代理与 FTP 交互、 TCP 连接、 FTP(和 HTTP 一样)是文件传输协议 FTP 使用了两个分离并行的 TCP 连接来传输文件,是控制连接(21端口)和数据连接(20端口) 使用 7bit 的 ASCII 格式在控制连接上传送命令/响应交互(不是 HTTP 那样使用报文交互),控 制连接在会话期间一直保持,为了区分连续的命令,每个命令后紧跟回车换行符,每个命令由 4 个大写 ASCII 组成,有些包含可选参数。 每一次数据连接只传送一个文件,发送方用关闭连 接表示一个文件传输结束. 将控制连接与数据连接分开:不会混淆数据与命令/响应,简化协 议设计和实现;在传输文件的过程中可以继续执行其它的操作,便于控制传输过程(如客户可 以随时终止传输)。 用关闭数据连接的方式结束文件传输:允许动态创建文件(不需预先告知 文件的大小)。 服务器在端口 21 等待客户,客户使用临时端口号建立连接,客户选择一个临 时端口号,在该端口上等待服务器的连接请求,客户在控制连接上用 PORT 命令将临时端口 号发送给服务器服务器使用端口 20 与客户机给出的端口建立连接

常见命令:USER username:PASS password:LIST(返回响应的文件列表是经过一个独立数据连 接传送的而非使用控制连接):RETR filename(传给用户文件);STOR filename(用户存储文件). 同答是一 3 位数字 后跟可选信息:331 I Isername OK Password required:125 Data connection already open; transfer starting:425 Can't open data connection:452 Error writing file.

因结园中的由子邮件:

3 个主要组成部分·**用户代理 邮件服务器 简单邮件传输协议(SMTP)** 用户代理介许用户 间速 同复 转发 保存和撰写邮件 用户代理将要外发的邮件发送到用户的邮件服务器 从 用户邮箱中取邮件 邮件服务器句令用户信箱 发送报文队列和报文传输代理 MTA 邮件放弃 郵件服务器的发送报文队列中 邮件服务器是由子邮件体系结构的核心 每个採收方在其中 的某个邮件服务器上有一个邮箱(计算机上的存储区域组成 每个信箱也被分配了唯一的由 子郎件場が 由子郎件場が包括・标识用自信箱的字符串和の之后的部件服务器的名字) 部件 服务器还包含报文传输代理 MTA:运行在服务器后台的系统守护进程,负责在邮件服务器之 间传输邮件 及将收到的邮件放入用户信箱 邮件从发送方用户代理开始 传送到发送方 邮件服务器被置于发送报文队列中邮件服务器与排放方邮件服务器建立 TCP 连接 再传输 邮件到接收方的邮件服务器 然后在议里被分发到接收方的邮箱中 接收方调用用户代理问 读邮件. 发送者的邮箱也必须能处理接收者邮件服务器的故障,如果不能交付邮件,则邮 件放在一个发送者的邮件服务器的报文队列中保持该报文并在以后尝试再次发送.

SMTP:SMTP 是因特网电子邮件中主要的应用层协议.使用 TCP.端口 25.SMTP 有两个部分 运行在发送方邮件服务器的客户端和运行在接收方邮件服务器的服务器端,(每台服务器既 是客户也是服务器)。 SMTP 限制所有邮件报文的体部分(不只是首部)只能采用简单的 7 比 特 ASCII 表示。 SMTP 不使用中间邮件服务器传送邮件,即使两个邮件服务器相聚很远,这意 味着邮件不会在中间某个邮件服务器存留。 采用命令/响应交互方式:命令为 ASCII(HELO,MAIL FROM,RCPT TO,DATA,OUIT),客户通过一个只包含一个句点的行(crif,crif) 指示 DATA 报文结束响应为状态码和可选英文解释短语. 使用持续连接:几个发往同一邮 件服务器的报文共用一个连接.那么每个报文以新的 MAIL FROM 开始,以独立句点结束,所 有都发完了才 OUIT.一个方向的报文传输结束后.可以在另一个方向上传输报文

与 HTTP 的对比:都用于从一台主机向另一台主机传送文件,都用持续连接:HTTP 主要是一个 拉协议,而 SMTP 基本上是一个推协议 SMTP 要求所有报文都是 ASCII,非字符数据要编码成 字符,HTTP 无此限制.HTTP 把每个对象封装到它自己的 HTTP 响应报文中,而 SMTP 则把所 有报文对象放在一个报文中

邮件报文格式和 MIME:电子邮件报文格式:首部行和报文体用空行(crlf)进行分割:首部行包 括:From:,To:,(可以包含 Subject:和一些其他的).这些首部行和 SMTP 命令是有区别的,现在讲 的首部行是邮件报文的一部分. Base64 编码:每 24 比特数据划分成 4 个 6 比特的单元,每 个单元编码成一个 ASCII 字符,其对应关系为:0~25 编码成'A'~'Z';26~51 编码成'a'~ 'z';52~61 编码成'0'~'9';62 和 63 分别编码成'+'和'/';若最后一组只有 8 比特或 16 比特, 分别加上'=='和'='后缀:回车和换行忽略,可以插在任何地方. quoted-printable 编码:适 用于绝大部分都是 ASCII 字符的报文实体,其编码方法是

每个 ASCII 字符保持不变;对于非 ASCII 字符(大于 127 的字符).将该字符的十六进制表示用 两个 ASCII 字符标记,前面冠以特殊字符"=" . 多用途因特网邮件扩展协议 MIME:扩展了 RFC 822,允许实体具有不同的数据类型,并规定了非 ASCII 文本信息在传输时的统一编码形 式.扩充了一些首部行,MIME version:最重要的是:Content-Transfer-Encoding:实体采用的传 输编码形式:Content-Type:实体的数据类型及子类型.

邮件访问协议:邮件访问方式:早期用户登陆到邮件服务器上,直接在服务器上运行一个邮件 阅读程序来阅读邮件:今天,用户在终端上安装用户代理,获取和阅读邮件,不能将用户信箱 放在本地终端。因为用户终端不可能一直连在因特网上。 用户代理可以用 SMTP 将邮件推 向邮件服务器.但 SMTP 是一个推协议.不能从服务器到用户代理。引入邮件访问协议将邮 件服务器上的报文传到末地 PC:第三版的邮局协议(Post Office Protocol 3) 因结网部件访问 协议(Internet Mail Access Protocol) 以及 HTTP POP3-TCP 110 端口・有三个工作阶段・特许 (代理明文传输用户名和密码)、事务处理(代理取回报文,对报文做删除标记,取消报文删除标 记、获取邮件统计信息)、更新(quit 之后服务器删除被标记的报文),命令:特许阶段 user.pass; 車条外理 list retr dele quit 同答·+○K(有时后面跟有数据) - FRR 下數并删除(其他终端不能 再看到)和下费并保留(其他终端可以再辞取) — 行单种的句点标识报文的结束 POP3 服务 器保留了一些状态信息 特别是记录哪些用户报文被标记为删除了 (H POP3 服务器并不在 POP3 全活中携带状态信息 设饰化了协议的实现 IMAP·所有邮件保存在服务器 F 分许 用户将邮件组织在文件来中 分许用户在文件来之间移动邮件 维护了会话的用户状态信息 允许用户代理获得报文的一部分。 HTTP:使用 HTTP 在用户代理和服务器之前上传和下载 邮件,而邮件服务器之间仍然是 SMTP.

DNS·因结网的目录服务

主机的一种标识方法是用它的主机名(www.yahoo.com),另一种是 IP 地址(32bit).

DNS 提供的服务域名系统(DNS)是:一个由分层的 DNS 服务器实现的分布式数据库:一个使 得主机能够查询分布式数据库的应用层协议。 DNS 运行在 UDP 上端口 53(也可以是 TCP 端口号 53),请求/响应报文交互。 DNS 在应用层上的原因:使用客户服务器模式运行在通信 的端系统之见。在通信的端系统之间通过下面的端到端运输协议来传送 DNS 报文. DNS 与 其他应用的不同之处:不是直接和用户打交道的应用,而是为因特网上的用户应用程序以及 其它软件提供一种核心功能. DNS 给使用它的应用带来了额外的延时. DNS 的服务:主机 名到 IP 地址的转换;主机别名到规范主机名的映射;迁移服务不需要修改主机名;邮件服务器 **别名**,允许使用域名作为邮件服务器的别名;负载分配,允许一个规范主机名对应一组 IP 地址, 将服务请求分配到功能相同的一组服务器上

DNS 工作机理概述:应用程序(如浏览器)调用一个本地 DNS 客户端(解析器),主机名作为参数 之一传递:解析器向网络中的 DNS 服务器发送查询报文,包含要查询的主机名:解析器收到包 含 IP 地址的响应报文:解析器将 IP 地址返回给调用者(如浏览器),对应用程序,DNS 是 供直接转换服务的黑盒子。 不使用集中式 DNS 的原因:单点故障.通信容量(单个服务器处 理全部查询),远距离的集中式数据库(较远的地方会有高延迟),维护(单个服务器维护庞大的 数据库),没有可扩展能力。 三种类型的 DNS 服务器:根 DNS 服务器(最高层.13 个),顶级域 (TLD)DNS 服务器(次高层,分为组织域(com,org),国家域(uk,fr)和反向域(域名 arpa,用来把一 个 ip 地址映射为名字)).权威 DNS 服务器(低层,每个组织机构,可以组织自己维护也可以由 ISP 维护), DNS 服务器以层次方式组织,还有本地 DNS 服务器(不属于 DNS 服务器的层次结

构),它通常邻近本主机,起着代理的作用,并将请求转发到 DNS 服务器的层次结构中。 DNS 名字树:域(domain):名字树中一个特定的节点以及该节点下所有的节点构成一个域.标记 (label):树上每一个节点都有一个标记(最多 63 个字符),树根的标记是一个空字符串,域名 (domain name):某个域的名字表示为从该域开始向上直到树根的标记序列:标记之间用句 点隔开(类似国外邮政地址的写法),域名的任一后缀也是一个域,同一个机构内的主机具有相 同的域名后缀.每个节点只需保证其孩子节点的标记不重名. 查询示例:DNS 客户查询根服 务器,得到 com 域的 TLD DNS 服务器地址;DNS 客户查询 com 域的 DNS 服务器,得到 amazon.com 域的权威 DNS 服务器地址;DNS 客户查询 amazon.com 域的 DNS 服务器.得到 www.amazon.com 的 IP 地址. TLD 服务器一般只是知道中间的某个 DNS 服务器.该中间 DNS 服务器依次才能知道用于该主机的权威 DNS 服务器。递归查询:高负载,选代查询:收 到查询报文的服务器将下一个需要查询的服务器地址返回给查询者,实践中查询通常是.从 请求主机到本地 DNS 服务器的查询是递归的,其余的查询是迭代的。整个 DNS 名字空间 被划分为一些不重叠的区域,称为 DNS zone,每个 zone 包含域名树的一部分;在管理上,每个 zone 代表一个权威域的边界。 一个物理服务器保存的信息可能涉及域名空间的若干层 它 也可以把它的域划分成若干子域,把其中的一些子域委托给其它服务器,实际的物理服务器 的层次与域名空间的逻辑层次不同。 DNS 缓存:每当收到一个响应报文,DNS 服务器将报文 中的映射信息缓存在本地,DNS 服务器首先使用缓存中的信息响应查询请求,DNS 缓存中的 映射在一定时间后被丢弃、特别地、本地 DNS 服务器通常会缓存 TLD 服务器的 IP 地址.因而 很少去访问根服务器. DNS 记录和报文·所有 DNS 服务器存储了资源记录(Resource Record) 每个 DNS 报文句念

了一条或名条资源记录 资源记录是一个句念了下列字段的购示 组(Name Value Type TTI) TTI 是该记录的生存时间 它决定了资源记录应当从缓存中删除 的时间:Type=A Name:中机名 Value:IP 地址:Type=NS Name:據 (e.g. foo.com) Value:该據 的权威 DNS 服务器的主机名:Type=CNAME,Name;主机别名,Value;规范主机 名:Type=MX.Name:城(e.g. foo.com), Value:该域的邮件服务器名字。 一台权威 DNS 服务器 包含其所用于的主机名的 A 记录:如果不是权威 DNS.则包含一条 NS 记录和记录了在 NS 记 录的 Value 字段中的 DNS 服务器 IP 地址的 A 记录。 DNS 只有查询和回答两种报文,并且有 着相同的格式:前12字节是首部区域。包含16bit 标识符来标识该查询。16bit 标志来标记报文 是查询(0)还是回答(1)、回答者是权威(1)还是非权威(0)、客户希望递归查询(1)还是不希望、 递归可用(1)或不可用 4 个 16bit 的有关数量的字段(依次是问题数,回答 RR 数,权威 RR 数,附 加 RR 数):问题区域包括正在进行的咨询的文字字段和类型字段:回答区域包含了对最初语 求的名字的 RR(可以有名名)和威区域句念了其他权威服务器的记录·附加区域句念了其他 有帮助的记录(如 MX 的回答包含一邮件服务器的规范主机名,那么附加区域包含一个该规 范主机的 A 记录)。 DNS 报文封装时如果长度小于 512B.使用 UDP.否则使用 TCP:如果不知 道响应报文的长度.先使用 UDP.一旦响应报文超过 512B.则截断该报文.置 DNS 首部 TC 标 志为 1.客户机 DNS 解析器打开 TCP 连接并重复请求。往 DNS 数据库中插入资源记录。则向 DNS 注册登记机构注册域名提供权威 DNS 服务器(主域名服务器和辅助域名服务器)的名 字和 IP.对每个权威域名服务器,注册机构往 TLD 服务器插入 NS 记录和对应 NS 中 Value 名 字的 A 记录. 建立权威 DNS 服务器 特别是建立 www.networkuptopia.com 的 Type A 记录 和 networkuptopia.com 的 MX 记录以及相应邮件服务器的 A 记录.

P2P 应用:集中式->查询洪泛->层次结构的覆盖网

P2P 文件分发:分发时间是所有 N 个对等方得到该文件副本所需要的时间。 客户-服务器体 系结构的分发时间:D cs = max{NF/u s, F/d min} P2P 体系结构的分发时间:D P2P = max {F/u s, F/d min, NF/(u s + Sigma(i=1 to N)u i)} P2P 体系结构的应用程序是能够自扩展 的,其直接成因是:对等方除了是比特消费者外也是重新分发者。 BitTorrent:用于文件分发 的流行 P2P 协议。 Torrent(洪流):参与一个特定文件分发的对等方集合。 每个洪流有一个 追踪器(tracker)跟踪洪流中的对等方。 当对等方加入洪流时,向跟踪器注册(并周期的通知 它还在洪流中),获得一个随机的对等方子集,试图与该子集建立 TCP 连接(成功创立的称为邻 近对等方)。 文件被划分为长 256KB 的块。周期性地,对等方询问每个临近对等方拥有的数据 块集合,并向邻居请求它缺少的数据块,请求数据块的顺序遵循最稀罕优先的技术(最稀缺的 块得到更为迅速的重新分发,目标是大致均衡每个块在洪流中的副本数量). 为了决定对等 方响应谁的请求,采用对换算法(tit-for-tat,一报还一报):Alice选择当前向其发送数据最快的 4 个邻居(称为疏通),响应他们的数据块请求,每隔 10 秒,重新评估向其提供数据最快的 4 个 邻居:每隔30秒,随机选择另一个对等方(如 Bob)响应其请求(Alice 可能成为向 Bob 上载最快 的 4 个邻居之一,Bob 也可能成为向 Alice 上载最快的 4 个邻居之一),其效果是对等方能趋于 找到彼此协调的速率下载。

视频流和内容分发网络

图像编码计数利用帧内冗余和帧间冗余减少需要的比特数.视频编码速率,Constant Bit Rate(編码效率低,简单帧浪费,复杂帧不够)/Variable Bit Rate(编码效率高) 流式多媒体技 术标准(Dynamic Adaptive Streaming over HTTP)适应异构网络环境,服务器将视频文件划分 为多个块,每块不同码率编码和存储,元文件提供不同 URL,客户周期性测量网络带宽,每次请 求一个块选择最大码率,避免缓存不足,获得最好视频质量,选择最高带宽/最近节点 套接字

IIDP 报文传输服务·由于没有建立管道 应用程序发送每个报文必须绘出证程进程协业·服务 器使用一个讲程和一个套接字为所有客户服务 一次请求 - 响应完成一次服务

TCP 字节流传输服务·中于建立了管道 应用程序只要向套接字中写入字节序列 不要提出证 程讲程协业服务器为每个客户单独生成一个套换字和一个新讲程 介许双方长时间通信



第3音·伝输局

据述和运输岸服务:

运输层协议为运行在不同主机上的应用进程之间提供了逻辑通信。不同主机上的应用进程 仿佛直接连在一起, 运输层协议是运行在端系统中而不是在路由器里实现的, 在发送端, 运输层将从发送应用程序进程接收到的报文转换成运输层分组,称为运输层报文段

法输尿与网络层的关系

网络层提供了主机之间的逻辑通信.传输层依赖并增强网络层服务.运输协议能够提供的服 务常常受制于底层网络协议的服务模型(网络层没有延迟和带宽保证,则运输层也没有),但运 输层也能提供某些服务(可靠的数据传输,加密)。

因特國法論屋据述 基本服务 为終主和间交付扩展到进程间交付

因特网网络层的网际协议 IP 提供尽力而为的交付服务,但并不做任何确保(不确保交付、按 需交付和数据完整性).这是不可靠服务。 UDP 和 TCP 最基本的责任是.将两个端系统间 IP 的交付扩展为运行在端系统上的两个进程之间的交付服务,被称为运输层的多路复用和多 路分解 UDP 和 TCP 还可以在报文段首部加上差错检查字段提供完整性检查 进程到进程的 数据交付和差错检查是两种最低限度的运输层服务,也是 UDP 仅能提供的服务,UDP 是不可

多路复用与多路分解

多路复用和多路分解是所有计算机网络都需要的。 一个进程有一个或多个套接字,它相当 于从网络向进程传递数据和从进程向网络传递数据的门户. 每个运输层报文段有几个字 段,在接收端,运输层检查这些字段,标识出接收套接字,进而将报文定向到该套接字.将运输 层报文段中的数据交付到正确的套接字的工作称为多路分解。 在源主机上从不同的套接 字中收集数据块,并为每个数据块封装上首部信息,从而生成报文段,然后将报文段传递到网 络层,所有这些工作称为多路复用。 运输层多路复用要求:套接字有唯一标识符:每个报文 段有特殊字段来指示该报文段所要交付到的套接字.这些特殊字段是源端口号字段和目的 端口号字段. 端口号是 16bit 的数,大小 0~65535,0~1023 是周知端口号,是受限制的,保留 给诸如 HTTP 的周知应用层协议. UDP 的多路分解:在主机上的每个套接字分配一个端口 号,报文段到达主机时,运输层检查报文段中的目的端口号,并将其定向到相应的套接 字.UDP 套接字标识为<IP 地址.端口号>二元组.这意味着如果两个 UDP 报文段有不同的源 IP 地址和/或源端口号,但具有相同的目的 IP 地址和目的端口号,那么这两个报文段将通过 相同的目的套接字被定向到相同的的目的进程.源端口号用作"返回地址"的一部分. 通常 应用程序的客户端让运输层自动地分配端口号,而服务器端则分配一个特定的端口号 TCP 套接字是由一个四元组<源 IP 地址,源端口号,目的 IP 地址,目的端口号>标识的,两个具 有不同源 IP 地址或源端口号的到达 TCP 报文段将被定向到两个不同的套接字、除非 TCP 报 文段携带了初始创建链接的请求.

无连接运输:User Datagram Protocol 不保证可靠及顺序,不保证带宽及延迟要求

UDP提供进程到进程间的报文交付,可选的完整性检查,使用 UDP 发送报文段之前,发送方和 接收方的运输层实体之间没有握手,UDP 是无连接的。 有许多应用更适合 UDP.主要是因 为·关于何时 发送什么数据的应用控制更为结缔(终计 TCP 拥塞控制并在应用层实现 LIDP 不提供的某些服务)·无要连接建立(不会引入连接建立的延迟 这可能是 DNS 运行在 LIDP 上 的原因)于连接状态/维护连接需要额外的接收发送缓存 拥塞控制参数。需要与确认号 法 行在 IDP 上可以节省资源支持更多用户)·分组首部开销小 当分组丢句家併并日出于安 全 某些机构阻塞 LIDP 液量 LIDP 没有拥塞控制 这使得在 LIDP 上运行多媒体是有争议的 使 用 IIDP 的应用程序可以自身建立可靠性机制来达到可靠传输 还可以有应用层定现可靠性 UDP 报文段结构:UDP 首部 8 字节:源端口号,目的端口号,长度(指示 UDP 报文段的字节数 句括首部和数据\ 检验和 LIDP 数据字段

UDP 检验和:检验和计算:对 UDP 报文段(以及 IP 首部的 32 位源和目的 IP.8 位前导 0 和 8 位 protocol 字段 16 位 UDP 总长度字段)的所有 16 比特字的和进行反码运算(第一次运算 前检验和字段填入 0).求和时遇到的任何一处都被回卷,得到的结果被放在 UDP 报文段中的 检验和字段 接收方路全部 16hit 字/包括检验和\加在一起 加里没有差错则结果为全1 否则 结果有错误。但 UDP 对差错恢复无能为力。

可靠数据传输原理:Reliable Data Transfer 实现"数据可以通过一条可靠的信道进行传输,借助于可靠信道,传输数据比特就不会收到

损坏或丢失,而且所有数据都是按照其发送顺序进行交付"的服务抽象是可靠传输协议(rdt) 的责任,可靠数据传输协议的下层协议也许是不可靠的,这决定了 rdt 的复杂性



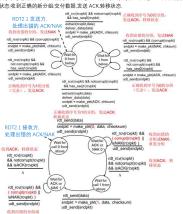
sender receiver

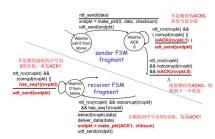
构造可靠数据传输协议·Rdt1.0: 可靠信道上的可靠传输:下层信道是完全可靠的(理想情况), 没有比特错误,没有分组丢失,发送能力≤接收能力:发送方 FSM:从上层接收数据, 封装成分组送入下层信道;接收方 FSM:从下层信道接收分组,取出数据交给上层

Rdt2.0: 可能产生比特错误的下层信道:下层信道可能使分组中的比特产生错误(比特翻转 可以通过某种检错码(如 checksum)检测比特错误;如何从错误中恢复:肯定确认(ACK 接收方 显式地告诉发送方,收到的分组正确),否定确认(NAK,接收方显式地告诉发送方,收到的分组 有错),发送方收到 NAK 后重传出错的分组;rdt2.0 中需要三种新的机制;接收方检错,接收方



Rdt 2 1·发送方·构造分组·加入序号·签结反馈-收到 NAK或者出错的反馈 重发分组-收到 ACK 转移状态 接收方:此到出错的分组:发送 NAK:此到完全的分组:发送 ACK 不交付数据 不转移 状态·此到正确的新分组·交付数据 发送 ACK 转移状态





rdt2.2: 不使用 NAK 的协议.接收方:只对正确接收的分组发送 ACK.ACK 携带所确认分组的 序号:若收到出错的分组,重发最近一次的 ACK,发送方:收到期待序号的 ACK:允许发送下一个 分组:其它情况(此时连续接收到同一分组的两个 ACK,即接收到一个冗余 ACK):重发当前分



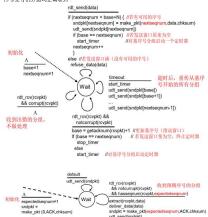
rdt3.0: 可能产生比特错误和美包的下层信道(有时称为比特交替协议):需要两项新技术:检 测丢包(包括数据及 ACK).从丢包中恢复 方法:检测丢包:若发送方在"合理的"时间内未收到 ACK认为丢句(需要定时器)从丢包中恢复:发送方面发当前分组,问题:ACK 丢失或超时设置 讨短导致的重发 会在接收端产生冗余分组

发送方:发送分组后启动一个定时器:收到期待序号的 ACK 终止定时器:转移状态:收到非期待 序号的 ACK 不做处理(为什么? 分组 1 的 ACK 在定时器超时之后到达 那么超时之后的重发 包(冗余数据分组)会被前一个 ACK 提早确认 状态转移:这时收到了重发分组 1 的 ACK 没必 要处理):定时器超时后 重发分组 接收方:数据分组丢失:接收方感知不到丢包 该事件对接收 方 FSM 无影响:ACK 丢失:过早超时:(发送方面发)接收方收到面复的数据分组:接收方利用序 号检测重复分组.重发前一次 ACK 接收方 FSM 同 Rdt2.2. Rdt3.0 是停等协议.发送方(信道) 利用塞:发送方实际上忙于将发送比特送讲信道的那部分时间与发送时间之比.停等协议的 U sender = L/R / (RTT + L/R). 解决: 允许发送方发送多个分组而无需等待确认 这种技术被 称为流水线。 流水线技术为可靠数据传输协议带来如下影响:必须增加序号范围:协议的发 送方和接收方两端也许必须缓存多个分组;所需序号的范围和对缓冲的要求取决于数据传 输协议如何处理丢失、损坏及延时过大的分组. 解决流水线的差错恢复有两种基本办法:回 退 N 步(GBN)和选择重传(SR)

将基序号(base)定义为最早的未确认分组的序号,将下一个序号(nextseqnum)定义为下一 个待发的分组序号.[0, base - 1] 已发送并被确认.[base, nextseqnum - 1] 已发送未确 认,[nextseqnum, base + N-1] 能用于缓存将要被发送的分组,大于 base + N 不可用. N 常被称为窗口长度,GBN 协议也称为滑动窗口协议。流量控制是对发送方施加限制 N 的原

回退 N 步协议的发送方最多允许 N 个已发送未确认的分组;只对于最早的已发送未确认的 分组使用一个定时器;若定时器超时,重传所有已发送未确认的分组,收到按序的 ACK,更新 base:收到失序的 ACK 不做处理.每当发送窗口从空变为不空时或者基序号更新,就重启定时

回退 N 步的接收方:收到一个按需分组时发送一个 ACK 分组 ACK 携带序号 接收方丢弃所有 失序分组(优点是接收缓存简单,缺点是浪费带宽且之后的重传也可能出错导致更多重传),并 在美车后发送一个最近按序接收的分组发送 ACK:使用累积确认: 若 ACK 包含序号 a.表明 "序号至 q 的分组均正确收到"



//预期序号加1

选择重传(SR)中发送方仅重传它认为出错(未收到 ACK)的分组,以避免不必要的重传,这种个 别的按需的重传要求接收方**逐个地**确认正确接收的分组。 SR 发送方的已发送未确认的每 个分组都要一个(逻辑)定时器:发送方收到 ACK 时ACK 在窗口内则将分组标记为已发送已 接收,如果分组序号恰为 send base,则 send base 滑动到最小序号的未确认分组处. SR 接 **收方**确认一个正确接收的分组而不论其是否按序。不使用累积确认、失序的分组将被缓存直 到所有丢失分组皆被收到为止,这时将他们一起交给上层,序号在 [rcv base , rcv base + N -1]被接收,发送一个选择 ACK;序号在 [rcv_base - N , rcv_base - 1] 内的分组被接收也 必须产生一个 ACK(如果正确收到分组然而 ACK 丢失必须确认发送方重传的分组否则发 送方窗口无法前进);其他情况忽略分组,对于哪些分组被接收了,哪些没有,发送和接受方并 不一定能看到相同的结果,SR 协议内的发送方和接收方窗口并不总是一致的。 对于 SR 协 议而言。窗口长度必须小于等于序号空间大小的一半,否则接收方能将重发的分组当作新分 组(考虑以下情形:发送方发送了一个窗口[0,N-1]的分组;接收方全都接收正确,发送了 ACK,并滑动接收窗口至[N,2N-1],但N个ACK全部丢失,发送端超时后逐个重发这N 个分组.为使接收端不会将重发的分组当成新的分组,窗口[0,N-1]和窗口[N,2N-1]不 能有重叠,N 不能大于序号空间一半) SR 特点:出错后重传代价小,接受端需要大量缓存, 面向连接的运输·TCP

TCP 连接:TCP 被认为是面向连接的,因为一个进程开始向另一个进程发送数据之前,二者必 须先发送某些预备报文段,以建立确保数据传输的参数和状态(套接字、缓存、变量),(其连接 状态完全保留在端系统中,不同于电路或虚电路)。 TCP 提供全双工服务:数据可以同时从 A 流到 R 并从 R 流到 A TCP 连接是占对占的 是在单个发送方与单个接收方之间的 发起连 接的称为**客户讲程** 另一个为服务器讲程 TCP 传送的是可靠 有序的字节液 不保留报文 边界 TCP 将客户进程通过套接字的数据存到该链接的发送编存里 在三次提手初期设立 TCP 可从缓存中取出并放入报文股份数据数量受限于最大报文股长度(MSS)(这是指报文股 里应用是数据的最大长度而非包括 TCP 首部的 TCP 报文股的最大长度) MSS 通常根据最初 确定的由太地发送主机发送的最大链路层帧长度(最大传输单元 MTII)来设置 TCP 连接的 组成 发送方和接受方所在的主机上的缓存, 变量和进程连接的套接字(路由器交换机和由线 器没有为该连接分配任何缓存或变量) 建立连接 流水式发送 流量控制

TCP 报文股结构:TCP 报文股由首部字股和一个数据字股组成数据字股句念一体应用数据 长度受限于 MSS. TCP 的首都一般 20 字节,分别是:16bit 源和目的端口号:32bit 序号:32bit 确认号:4bit 首部长度(指示以 32bit 为单位的首部长度,间接指示 TCP 选项字段的长度):6bit 保留未用·6bit 标志字段 LIRG(报文段存在由发送器的上层置为紧急的数据 实际未用)。 ACK(指示确认字段的信息 () 否 1 有效)。 PSH(接收方应立即将数据传统 上层 实际表用) RST. SYN. FIN:16bit 接收會口字段(用于流量控制 指示接收方原音接受的字节数量):16bit 因特网检验和·16bit 紧急数据指针/指向紧急数据的最后—个字节 紧急数据的置在数据数据 前端,实际未用);可选与变长的选项字段,用于发送方与接收方协商 MSS(缺省为 536 字节),或 在高速网络环境下用作窗口调节因子,或提供选择确认 SACK. TCP 把数据看成一个无结构 的、有序的字节流,故一个报文段的序号是该报文段数据首字节的字节流编号。 主机 A 填 充进报文段的确认号是主机 A 期望从主机 B 收到的下一个字节的序号。 TCP 流水式发送报 文段,提供累计确认,实践中的 TCP 接收方保留失序的字节,并等待缺少的字节填补间隔,RFC 标准没有规定。 一条 TCP 连接的双方均可以随机的选择初始序号。 对客户到服务器的数 据的确认被装在在一个承载服务器到客户的数据的报文段中,这种确认被认为是被摘带在 服务器到客户的数据报文段中的

往返时间的估计与超时:超时间隔必须大于 RTT,否则造成不必要的重传。 报文段的**样本** RTT(SampleRTT)是从某报文段被发出到对该报文段的确认被收到之间的时间量.TCP 不对 重传的报文段测量 RTT(TCP 是对接收到的数据而不是对携带数据的报文段进行确认、因此 TCP 的确认是有二义性的,对重传报文段的 RTT 估计不准确,解决方法是,忽略有二义性的确 认、只对一次发送成功的报文段测量 SampleRTT ,并据此更新 EstimtedRTT,当 TCP 重传一个 段时,停止测量 SampleRTT),在任意时刻也仅为一个已发送但目前尚未被确认的报文段估计 SampleRTT,不是每个都估计。 EstimatedRTT = (1 - α)*EstimatedRTT + α*SampleRTT,α 参考值 0.125.是 SampleRTT 的指数加权移动平均(EWMA). RTT 偏差 DevRTT 用于估算 SampleRTT 一般会偏离 EstimatedRTT 的程度:DevRTT = (1 - β)*DevRTT + β*| SampleRTT - EstimatedRTT |,β推荐值 0.25. 设置超时值 TimeoutInterval = EstimatedRTT + 4*DevRTT, 初始 TimeoutInterval 推荐 1s,当出现超时后,TimeoutInterval 加倍(而非上述公式的估计值); 一旦报文段 ACK 被收到并更新 EstimatedRTT 后,或者收到上层的应用数据之 后,TimeoutInterval 就又用上述公式计算了.这种 TimeoutInterval 修改提供了一个形式受限

可靠数据传输:TCP 在不可靠的 IP 服务上建立可靠的数据传输:定时器管理过程仅使用单一 的重传定时器,定时第一个发送未确认的报文段,像 SR 那样每个维护一个定时器需要相当大 的开销。 定时器超时,仅重传最早未确认的报文段,并重启定时器。 当收到 ACK 时,只要确 认序号是大于基序号的,就推进发送窗口。 超时重传问题之一是:超时周期可能相对较长,报 文段丢失时这种长周期迫使发送方延迟重传丢失的分组,因而增加了端到端延迟。 冗余 ACK 就是再次确认某个报文段的 ACK.而发送方先前已经收到对它的确认.一旦收到三个冗 余 ACK,TCP 执行快速重传,在定时器过期之前重传报文段。 为减小通信量,TCP 允许接收端 推迟确认 接收端可以在收到差于个报文股后 发送一个累积确认的报文股 缺点是差征识太 大会导致不必要的重传日推迟确认法成 RTT 估计不准确 TCP 提定推迟确认的时间最多为 500ms 日当此到—个报文股时 若採此方的前—个按序报文股签结传输 ACK 则立刻同时确 认两个报文段 当能部分或完全填充接收数据间隔的报文段到达 债差该报文段起始于间 屬的低端 则立即发送 ACK 当生序报文段或重复报文段到法 则立即发送完全 ACK TCP 的修改版太添加 SACK 機獅斗 绘出收到的非连续数据性的上下边界 TCP 的差错恢复机 制可以表成是 GRN 和 SR 的混合体:TCP 在减小定时器开锁和重传开锁方面要像于 GRN 和

流量控制 TCP 提供**流量控制服备**(GRN 和 SR 均不全出现接收端缀存溢出 因为正确按序到 达分组被立即交付,释放空间:UDP 不需要保证)以消除发送方使接收方缓存溢出的可能性。 TCP 发送方维护一个接收窗口来提供流量控制,这个窗口指示发送方该接收方还有多少可用 的缓存空间 定义变量 LastRyteRead(主机 R 的应用进程从缓存进出的数据流的最后一个 字节的编号)和 LastByteRcvd(主机 B 接收缓存放入的最后一个字节的编号).LastByteRcvd -LastByteRead <= RcvBuffer.接收窗口 rwnd = RcvBuffer - [LastByteRcvd - LastByteRead]. 主机 A 满足 LaskByteSent - LastByteAcked <= rwnd. 该方法的一个小问题:对零窗口的 处理:当接收窗口为0时,发送方必须停止发送:当接收窗口变为非0时,接收方应通告增大的 接收窗口.TCP 协议规定:发送方收到"零窗口通告"后,可以发送只有一个字节数据的"零窗口 探测"报文段,以触发一个包含接收窗口的响应报文段。 零窗口探测的实现:发送端收到零窗 口通告时,启动一个坚持定时器;定时器超时后,发送端发送一个零窗口探测报文段(序号为上 一个段中最后一个字节的序号):接收端在响应的报文段中通告当前接收窗口的大小:若发送 端仍收到零窗口通告,重新启动坚持定时器. 糊涂窗口综合症:当发送速度很快、消费速度 很慢时,零窗口探测的简单实现带来以下问题:接收方不断发送微小窗口通告;发送方不断发 送很小的数据分组;大量带宽被浪费:接收方启发式策略:通告零窗口之后,仅当窗口大小显著 增加(窗口大小达到缓存空间的一半或者一个 MSS,取两者的较小值)之后才发送更新的窗口 通告.TCP 执行该策略的做法:当窗口大小不满足以上策略时,推迟发送确认(但最多推迟 500ms,且至少每隔一个报文段使用正常方式进行确认), 寄希望于推迟间隔内有更多数据被 消费:仅当窗口大小满足以上策略时,再通告新的窗口大小.发送方启发式策略:发送方应积 聚足够多的数据再发送,以防止发送太短的报文段,发送方应等待多少时间? Nagle 算法:在 新建连接上,当应用数据到来时,组成一个 TCP 段发送(那怕只有一个字节);在收到确认之前, 后续到来的数据放在发送缓存中:当数据量达到一个 MSS 或上一次传输的确认到来(取两者 的较小时间),用一个 TCP 段将缓存的字节全部发走.Nagle 算法的优点;适应网络延时、MSS

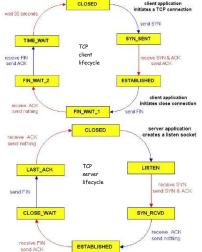
长度及应用速度的各种组合:常规情况下不会降低网络的吞吐量. UDP 不提供流量控制,接 收缓存可能溢出并丢失报文段.

TCP 连接管理:在网络中 2 次握手总是可行的吗? 在一个不可靠的网络中,会有一些意外发 生.比如:包传输延 迟变化很大:存在重 传的报文段:存在报 文重排序 TCP 连接方式:第一 步·玄户端 TCP 向服

冬器发送— SYN 署 1、不包含应用数据、 随机选择一个初始 序号 Seq=x 并将编 号放置在序号字段 中的 SYN 报文段. 第二步:服务器为 TCP 连接分配 TCP

缓存和变量。向客户发送一个允许连接的 SYNACK 报文段(SYN 和 ACKbit 置 1.确认号字段为 ACKnum=x+1.服务器选择自己的Sea=v放入序号字段).第三步:客户为自己分配缓存和变量 向服务器发送另一报文股/将 ACKnum=v+1 放入确认字股 ACKhit = 1 SYN = 0 据文股中可 以角数数据) 此后每个报文股中 SYN 都为 0 选择的起始序是必须避免新用连接上的原 · 生香桑·其干时钟 取较小的AT 递增时钟计数器/确保发送序号的增长速度不会超过起始 序号的增长速度) 取计数器低 32 位(确保序号回络的时间远大于分组在网络中的最长去命) TCP 美丽连接·客户应用讲程发出一个美丽连接命令 这会引起客户 TCP 向服务器讲程发送 一个特殊的 TCP 报文股(FINhit 署 1 sen = x) 服务器此到后向发送方同送一个确认报文股 (ACKbit=1,ACKnum=x+1),然后服务器发送它自己的终止报文段(FINbit=1,seg=v),最后客户 对这个服务器的终止报文段进行确认(ACKbit=1.ACKnum=v+1),此时连接的所有资源都被 释放

SYN 洪泛攻击:服务器收到 SYN 报文段、发送 SYNACK 段后就分配了资源,而等待一段时间 (30s~120s)才丢弃未完成的连接:攻击者伪造大量 SYN 段而不发送 ACK 段:耗尽服务器资源 防御系统 SVN cookie 当此到的 TCP 报文段与太主机 FP 有的套接字不匹配时 主机发送 报 端口扫描·扫描程序利用与目标机器建立 TCP 连接过程中获得的确应消息来收集信息 在典型的 TCP 端口扫描过程中,发送端向目标端口发送 SYN 报文段,若收到 SYNACK 段,表



明该目标端口上有服务在运行:若收到 RST 股表明该目标端口上没有服务在运行:若什么也 没有收到表明路径上有防火塘(有些防火塘会丢弃来自外网的 SYN 报文段) FIN 扫描过程如 下:发送端向目标端口发送 FIN 报文段:若收到 ACK=1、RST=1 的 TCP 段,表明目标端口上没 有服务在监听;若没有响应,表明有服务在监听(RFC 973 的规定).缺点:有些系统的实现不符 合 RFC 973 规定,如在 Microsoft 的 TCP 实现中,总是返回 ACK=1、RST=1 的 TCP 段

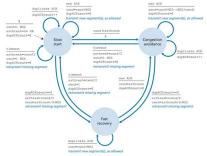
拥塞控制原理: 拥塞原因与代价:拥塞造成:丢包(缓存溢出),分组延迟增大(链路接近满载,排队延迟急剧增大);

大量网络资源用于:重传丢失的分组:(不必要地)重传延迟过大的分组:转发最终被丢弃的分 组.结果:网络负载很重.但网络吞吐量很低.必要与不必要的重传 吞吐量不可能超过链路容量,负载接近链路容量时延迟增大,丢包重传导致有效吞吐量下降

不必要的重传进一步降低有效吞吐量、下游被丢弃的包浪费上游的传输容量和缓存空间 拥塞控制方法:网络辅助的拥塞控制:路由器向端系统提供反馈(拥塞指示比特,发送速率指 示).ATM 采用此类方法。 端到端拥塞控制:网络层不向端系统提供反馈,端系统通过观察丢 包和延迟推断拥塞的发生,TCP采用此类方法(因为IP层不提供显式反馈) TCP 拥塞控制·

运行在发送方的 TCP 拥塞控制跟踪变量拥塞窗口(congestion window,cwnd),控制发送方 速率、特别是 LastByteSent - LastByteAcked <= min { cwnd , rwnd },只考虑 cwnd 时粗略的 讲,在每个 RTT 的起始点,限制条件允许发送方发送 cwnd 个字节的数据,在该 RTT 结束时发 送方接收对数据的确认报文、发送方的速率大概是 CongWin/RTT Bytes/sec. 感知拥塞:丢 包或分组延迟过大对于发送端而言都是丢包, 丢包事件(或者超时或者收到三个冗余 ACK):TCP 使用确认的到达作为一切正常的指示.使用确认来增加窗口的长度(进而增加传输 速率). 加性增.乘性减(AIMD.拥塞避免模式):乘性减:检测到丢包后.将 CongWin 的大小减 半(迅速减小),但不能小于一个 MSS.加性增:若没有丢包,每经过一个 RTT 将 CongWin 增大一 个 MSS.直到检测到丢包(缓慢增大)。 慢启动:连接刚建立时 ConaWin = 1 MSS(发送速度= MSS/RTT):每收到一个 ACK 段,CongWin 增加一个 MSS(这一过程每过一个 RTT 就把 cwnd 翻一倍)(以一个很低的读家开始 按指数增大发送读家)直到出现手包事件(此时终止慢启动 付程) **收到3个元全 ACK**·CongWin 降为—坐 妥用 AIMD 调节: 超財・CongWin = 1MSS・使 用慢启动增大 CongWin,至超时发生时 CongWin 的一半.使用 AIMD 调节. 实现:发送方维

护一个变量 Threshold(ssthresh):发生丢包事件时,Threshold 设置为当前 CongWin 的一半 (整除 2).若事件是收到 3 次冗余 ACK,CongWin 从原 CongWin 的一半+3 开始,执行 AIMD;若 发生的是超时,CongWin 从 1MSS 开始,执行慢启动,慢启动翻倍后值大于 ssthresh 则将窗口 大小设置为等于ssthresh.Threshold 是从慢启动转为拥塞避免的分水岭:CongWin 低于门限 慢启动;CongWin 高于门限:拥塞避免(AIMD). TCP Tahoe:只要是丢包事件就将 cwnd 变成 1MSS 并减半 ssthresh TCP 的吞吐量的宏观描述(忽略慢启动):令 W=发生丢包时的 CongWin, 此时 throughput = W/RTT; 发生丢包后调整 CongWin=W/2, 此时 throughput=W/2RTT:假设在 TCP 连接的生命期内,RTT 和 W 几乎不变,则:Average throughout=0.75 W/RTT. 经高带宽路径的 TCP:吞吐量与丢包率 L 的关系 = 1.22*MSS / (RTT+Sqrt[L]);针对高速网络需要新的 TCP 拥塞控制算法. 快速恢复不是 TCP 必需的部件. 公平性:公平性的目标: 如果 K 条 TCP 连接共享某条带宽为 R 的瓶颈链路.每条连接具有平 均速度 R/K. TCP 是公平的. 若相互竞争的 TCP 连接具有不同的参数(RTT、MSS 等),不能 保证公平性:若应用(如 web)可以建立多条并行 TCP 连接,不能保证带宽在应用之间公平分配 (比如,一条速率为 R 的链路上有 9 条连接;若新应用建立一条 TCP 连接,获得速率 R/10,若新 应用建立 11 条 TCP,可以获得速率 R/2) 开发一种因特网中的拥塞控制机制,用于阻止 UDP 流量不断压制(TCP)直至中断因特网吞吐量。



小測題和易错点:

易错点:khos = 1000bps. 平均传输次数=平均重传次数+1

习题课:若偶尔发送数据,只用 NAK 的协议会导致丢失的包在很久之后的下一个包被接收时 在能发现. 美失的包不能及时重传延迟大大提升. 若要发送大量数据日美包率低.则上述情况. 不会发生。由于美包率低所需发送的 NAK 远少于 ACK使用仅 NAK 的方法可以降低网络负 担并日发送方可以不用等待 ACK 持续发送降低了传输证识。

在网络传输中和序是很常见的情况,若收到第一个冗余 ACK 就重传,对每个和序都要重传-次冗余的分组,极大降低了吞叶量,收到 3 个冗余 ACK 后重传可以很大程度减轻乱序重传,同 时兼顾快速重传

小测答案: 常用的通信传输介质有哪些? 它们之间的主要区别? (1)有线:双绞线, 同轴电缆, 光纸:无线 (2)区别:带京、误码率、传输距离、价格、频谱及复用方式、是否支持移动通信等。无连接分 组交换与面向连接(虚电路)分组交换的区别? (1)分组格式:前者完全源、目的地址:后者虚电 路号 (2)路由表:前者面向整个网络拓扑.转发时顺序查找路由表:后者面向特定路径或源路 由 转发基于索引查找路由表 (3)可靠性、顺序性:前者无:后者有 (4)建立、维护连接:前者无: 后者有 假定要传送的报文共有 x/单位 hit) 从源节点到目的节点共有 k 跳链路 每条链路的 传播延迟为 d(单位 s) 链路带宽为 b(单位 bit/s):电路交换(包括连接建立与拆除)使用的控制 帧(或信令)长度、在各节点的排队延迟忽略不计;分组交换使用的分组头、分组长度分别为 h, p(单位 bit)分组在各节点的排队延迟 p(单位 s),过分析在何种条件下电路交换的总延迟 要小于分组交换的总延迟? 电路交换总延迟 D(c):连接建立时间:kd 连接拆除时间:kd 数据 传输时间:x/b 数据传播时间:kd D(c)=3kd+x/b 分组交换总延迟 D(p): 单个分组传输时 间:(p+h)/b 第 1 跳传输时间:(x/p).((p+h)/b) (x/p 为分组个数) 传输时间每 1 跳增加 1 个分 组的传输时间,总的传输时间为 x/p*(p+h)/b+(k-1)*(p+h)/b 排队时间:kq 传播时间:kdD(p)=x/p*(p+h)/b+(k-1)*(p+h)/b+kd+kq 若D(c)<D(p),则 ··· 若使用一个256-kbps 的无差 错卫星信道(往返传播延迟为 512-msec)一个方向上发送 512-byte 数据帧,而在另一个方向 上返回很短的确认帧.则对于窗口大小为 1, 15, 127 的最大吞吐量是多少? 512*8/256k=16ms (1)k=1,16/(16+512)*256=7.75 (2)k=15,7.75*15=116.36 (3)k=127,256 滑动窗口协议中,退后 N 帧与选择性重传利用链路缓冲能力连续发送多个帧,令帧的传输时间(tral time)=1(归一化)、传播时间(propagation time)=a,则链路的缓冲能力为?a(单向)或2a(双向) TCP 协议中 ACK 的作用. (1)建立连接、拆除连接 (2)差错控制(或可靠传送) (3)流量控制 (4) 拥塞控制 TCP 连接的目标 (1)实现进程间通信 (2)实现可靠传送 (3)实现按序传送 (4)进行 流量控制 (5)进行拥塞控制 在 TCP 连接中,客户端的初始号 215.客户打开连接,只发送一个 携带有 100 字节数据的报文段,然后关闭连接,试问下面从客户端发送的各个报文段的序号 分别是多少? (1)SYN 报文段:(2)数据报文段:3)FIN 报文段(1)215:(2)216:(3)316 在一条新建 的 TCP 连接上发送一个长度为 32KB 的文件 发送端每次都发送一个最大长度的段 (MSS) MSS 的长度为 1KB.接收端下确收到一个 TCP 股后立即给予确认发送端的初始拥塞 窗口门踢设为 16KB.假设发送端尽可能快地传输数据.即只要发送窗口允许发送端就发送-MSS(1)已知发生第一次超时后发送端将拥塞窗口门限调整为 4KB 请问发生超时的时候 发送端的拥塞窗口是多大?此时发送端井发送了多少数据?其中有多少数据被成功确认了 (2)发送端从未被确认的数据开始使用慢启动进行重传。假设此后未再发生招时,当文件全部 发送完毕时,发送端的拥塞窗口是多大? 答:(1) 第一次超时发生时,发送端拥塞窗口大小 = 4KB+2 = 8KB 在新建立的 TCP 连接上,发送端采用慢启动开始发送,因此当第一次超时发生

时,发送端已发送的数据量 = 1KB + 2KB + 4KB + 8KB = 15KB.此时,除最后一批 8 个 TCP 段 未获确认外,之前发送的 TCP 段都被确认,因此成功确认的数据量为 7KB(2) 发送端采用慢 启动重新开始发送,在拥塞窗口达到 4KB 时发送数据量=1KB+2KB+ 4KB=7KB,然后进入拥塞 辦免阶段:在收到全部 4 个 MSS 的确认后 拥塞窗口增至 5KB 相应地发送端发送了 5KB 数据: 收到全部 5 个 MSS 的确认后 拥塞窗口增至 6KB:收到全部 6 个 MSS 的确认后 拥塞窗口增 至 7KB:此时则好发完 因此,文件发送结束时,发送端的拥塞窗口大小为 7KB. 由题, 在从 $W/(2 \times RTT)$ 到W/RTT 的周期内, 一共有 $\sum_i (i=0)^i (W/2)(W+i)=3/i$

8 W^2 + 3/4 W 个报文段被传输。而这个周期内只丢失了一个包、也就是丢包率为1/(3/ $8W^2 + 3/4W$

 b. 因为W^2相比W数量级高。可以忽略后者影响。则L≈8/(3W^2)。即 W≈√8/3L 而 又 有 平 均 读 率 = MSS×(3/8W^2+3/4W)/(W/2)×RTT≈\ $frac{0.75W^2 MSS}{RTT} \approx frac{1.22MSS}{RTT}$

正常 W/RTT.W/2RTT.平均 0.75W/RTT