



网络

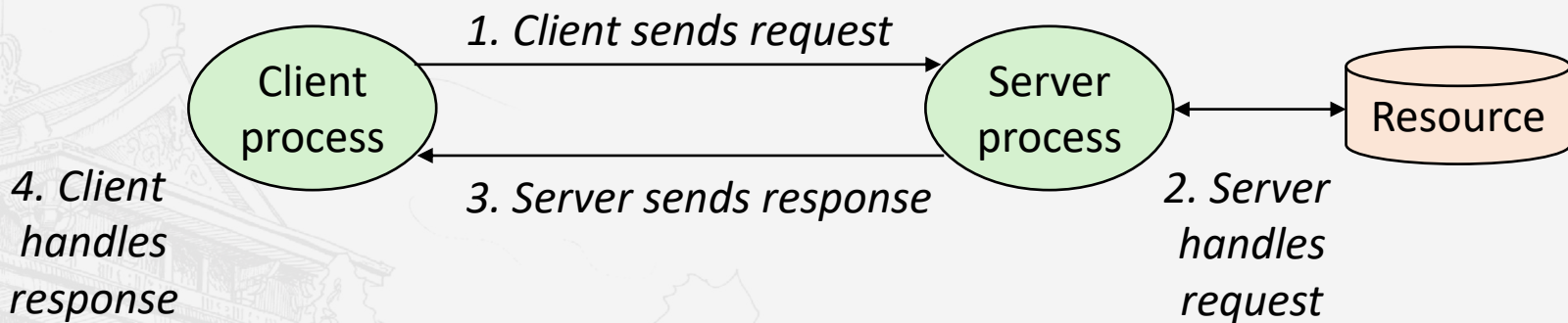
Network

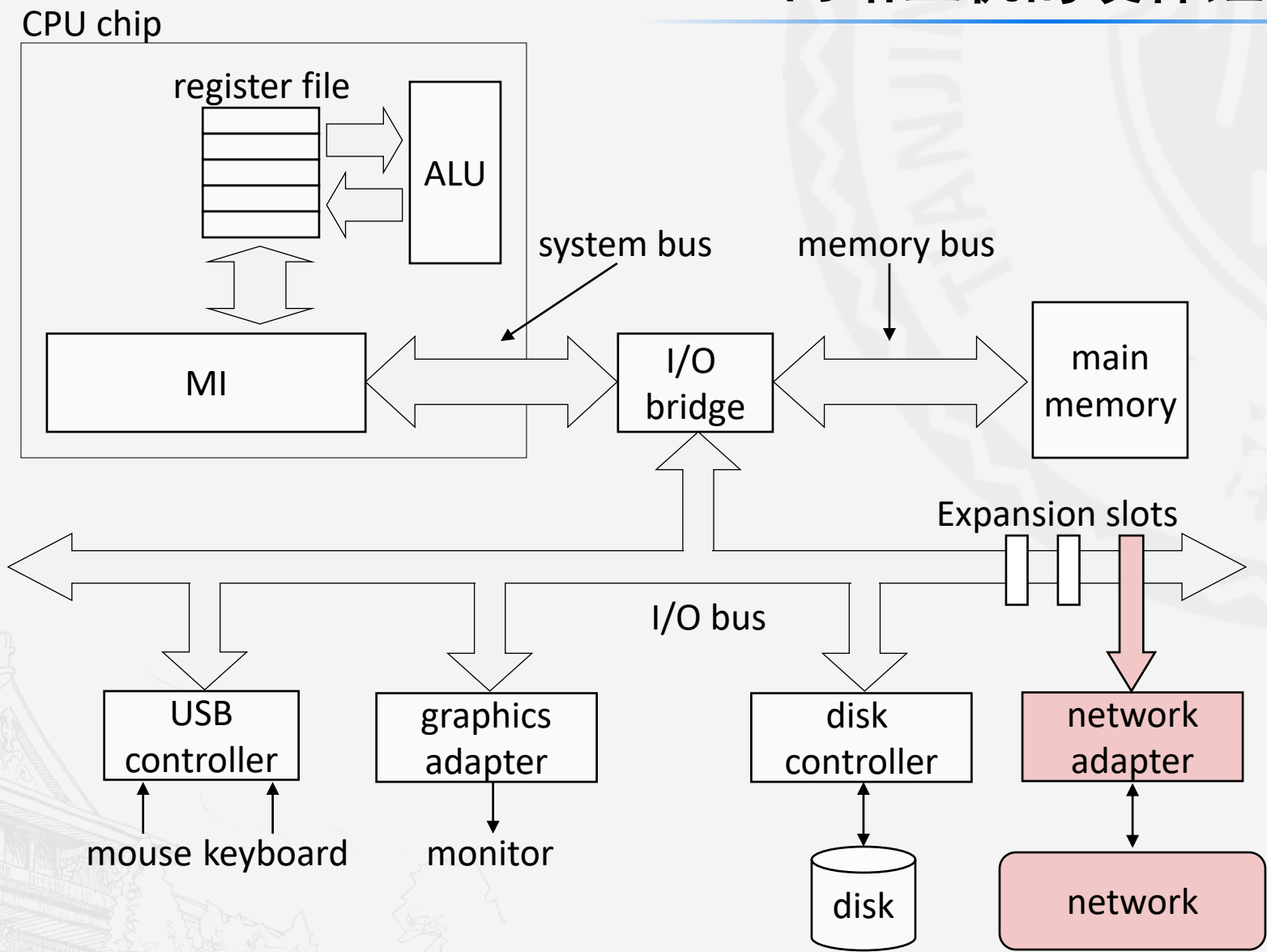
■ 大多数网络应用基于客户端-服务器模型：

- 由一个服务器进程和一个或多个客户端进程组成
- 服务器管理某些资源
- 服务器通过为客户端操作资源来提供服务
- 服务器通过客户端的请求激活服务

注意：客户端和服务端是运行在主机上的进程

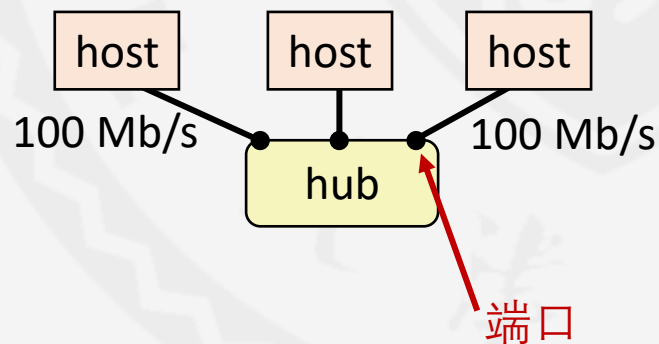
(客户端和服务端可以运行在同一个或不同的主机上)



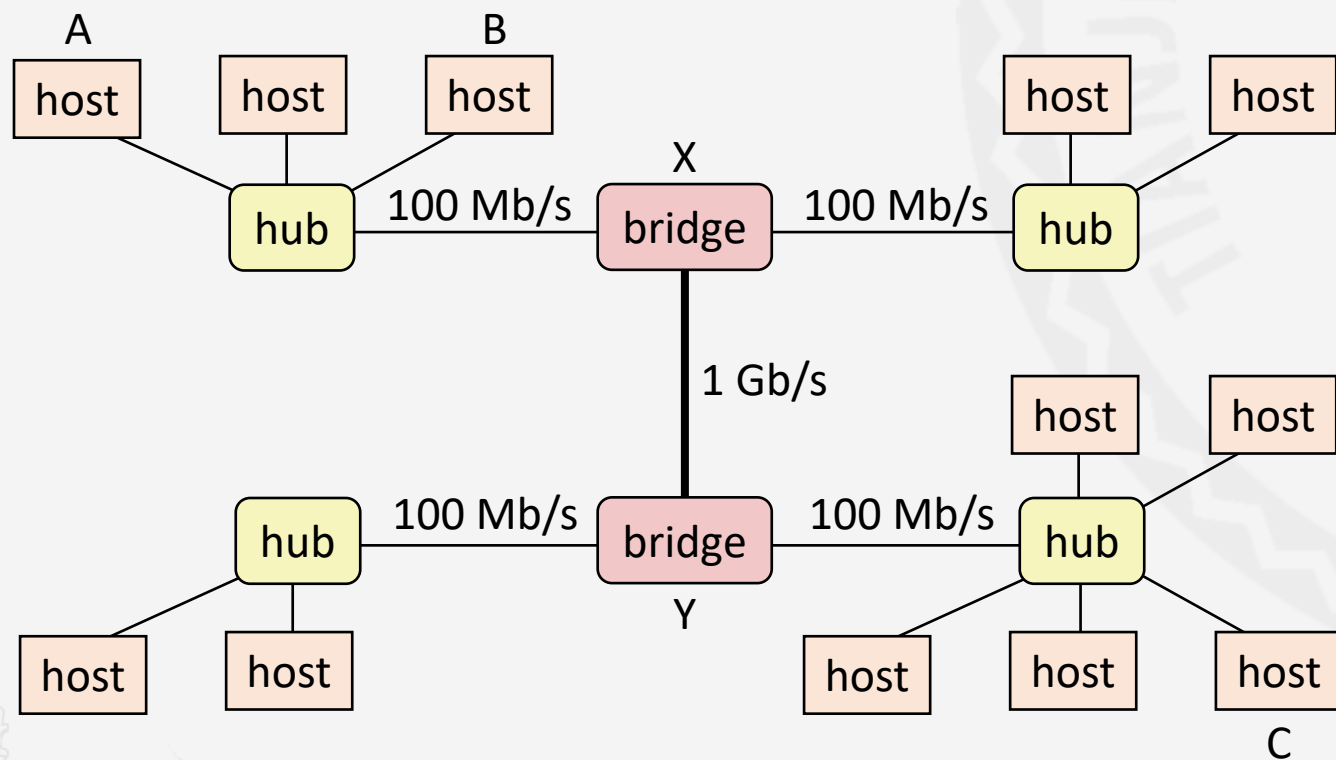


- **网络**是一种由盒子和电线按照地理位置组织的分层系统：
 - 系统区域网络（SAN, System Area Network）：覆盖集群或机房。典型的技术包括交换以太网、Quadrics QSW等。
 - 局域网（LAN, Local Area Network）：覆盖一个建筑或校园。以太网是最显著的例子。
 - 广域网（WAN, Wide Area Network）：覆盖一个国家或世界。通常使用高速点对点电话线路。
- 互联网（internet）：一个由互联的网络组成的系统
 - 全球 IP 互联网（Internet）是最著名的互联网（internet）的例子。
- 接下来，我们将从基础开始看看互联网是如何建立的。

- 以太网段由一组通过电缆（双绞线）连接到集线器的主机组成。
- 它通常覆盖建筑内的一个房间或楼层。
- 工作模式：
 - 每个以太网适配器都有一个唯一的48位地址（MAC地址）
 - 例如：00:16:ea:e3:54:e6
 - 主机将数据以块的形式发送到任何其他主机时，这些数据块被称为帧（frame）
 - 集线器（hub）会将每个端口的数据都复制到其他端口上
 - 每个主机都能看到数据
 - 集线器正在逐渐被淘汰，因为桥接器（交换机、路由器）变得足够便宜，可以替代它们。

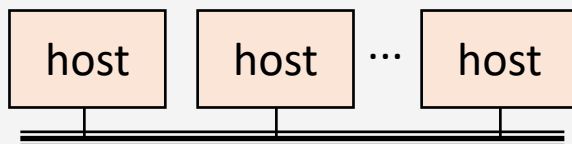


下一级别：桥接以太网段

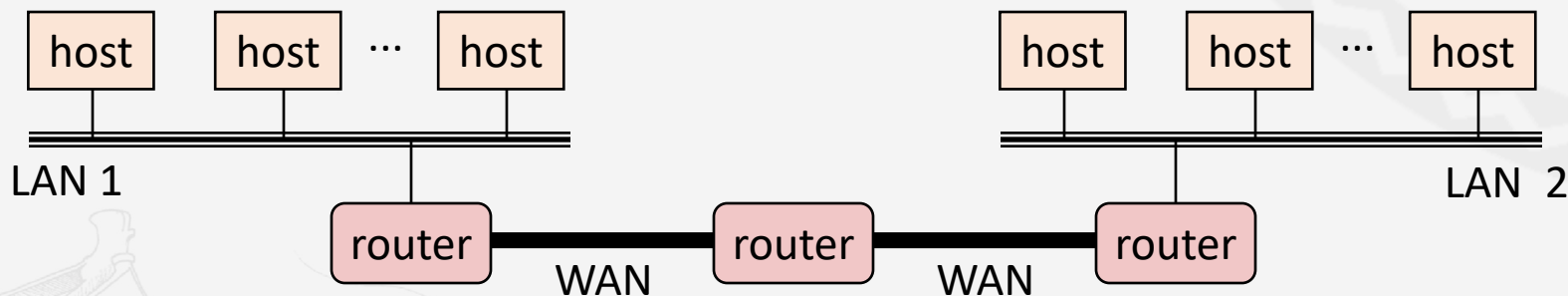


- 覆盖一个建筑或是整个校园
- 桥接器（Bridge）可以学习哪些主机可以从哪些端口访问，然后有选择地将帧从一个端口复制到另一个端口

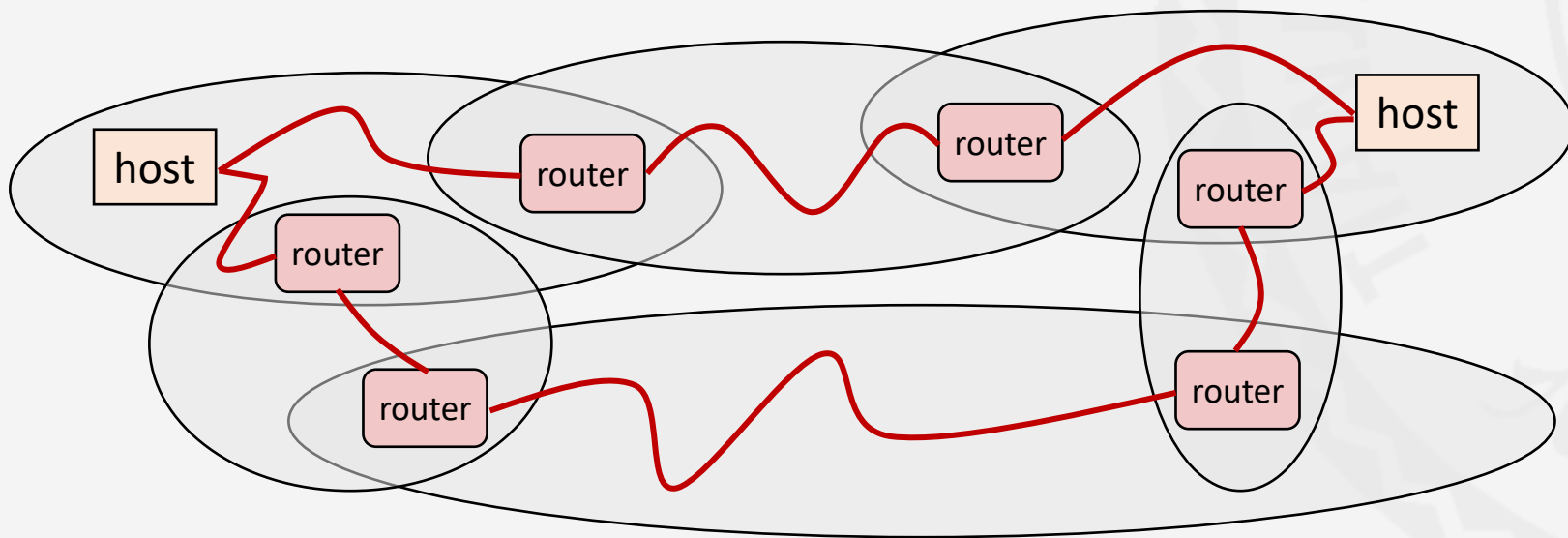
■ 简单起见，集线器、桥接器和电线通常显示为连接到一根电线的多个主机的集合：



- 多个不兼容的局域网可以通过称为**路由器** (router) 的专用计算机进行物理连接。
- 这些连接的网络被称为**互联网** (internet) 。



注意：LAN1和LAN2可能完全不同，完全不兼容（例如，以太网、光纤、802.11*、T1链路、DSL等）。



- 采用临时互连（Ad hoc）的方式组织

- 没有特定的拓扑结构

- 路由器和链路容量差异巨大

- 通过在网络之间跳跃，将数据包从源发送到目的地

- 路由器从一个网络到另一个网络形成桥接

- 不同的数据包可能会走不同的路线

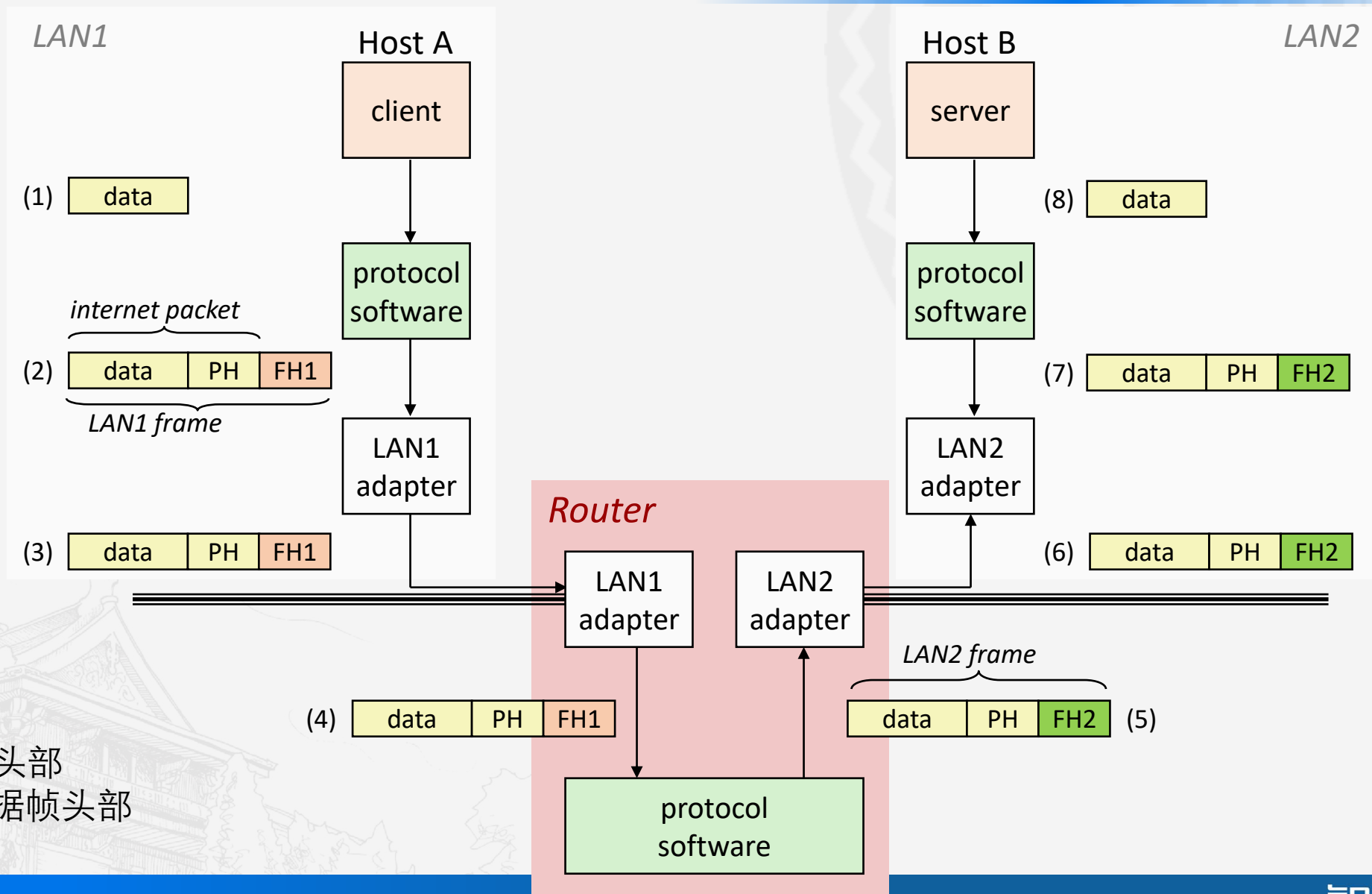
- 如何在 incompatible 的局域网（LAN）和广域网（WAN）上发送数据？
- 解决方案：在每台主机和路由器上运行**协议**软件。
 - 协议是一组规则，用于规范主机和路由器在从一个网络传输数据到另一个网络时应该如何协作。
 - 这种软件消除不同网络之间的差异。

■ 提供命名方案

- 互联网协议定义了一种统一的主机地址格式。
- 每台主机（和路由器）至少被分配一个唯一标识它的互联网地址。

■ 提供传输机制

- 互联网协议定义了一种标准的传输单元（数据包）。
- 数据包由头部和有效载荷组成：
 - 头部：包含信息，如数据包大小、源地址和目的地址。
 - 有效载荷：包含从源主机发送的数据位。



PH: 互联网包头部
FH: 局域网数据帧头部

■ 我们在上面的讨论中忽略了许多重要的问题：

■ 如果不同网络有不同的最大帧大小怎么办？（分段）

■ 路由器如何知道将帧转发到哪里？

■ 当网络拓扑结构发生变化时，路由器如何获知？

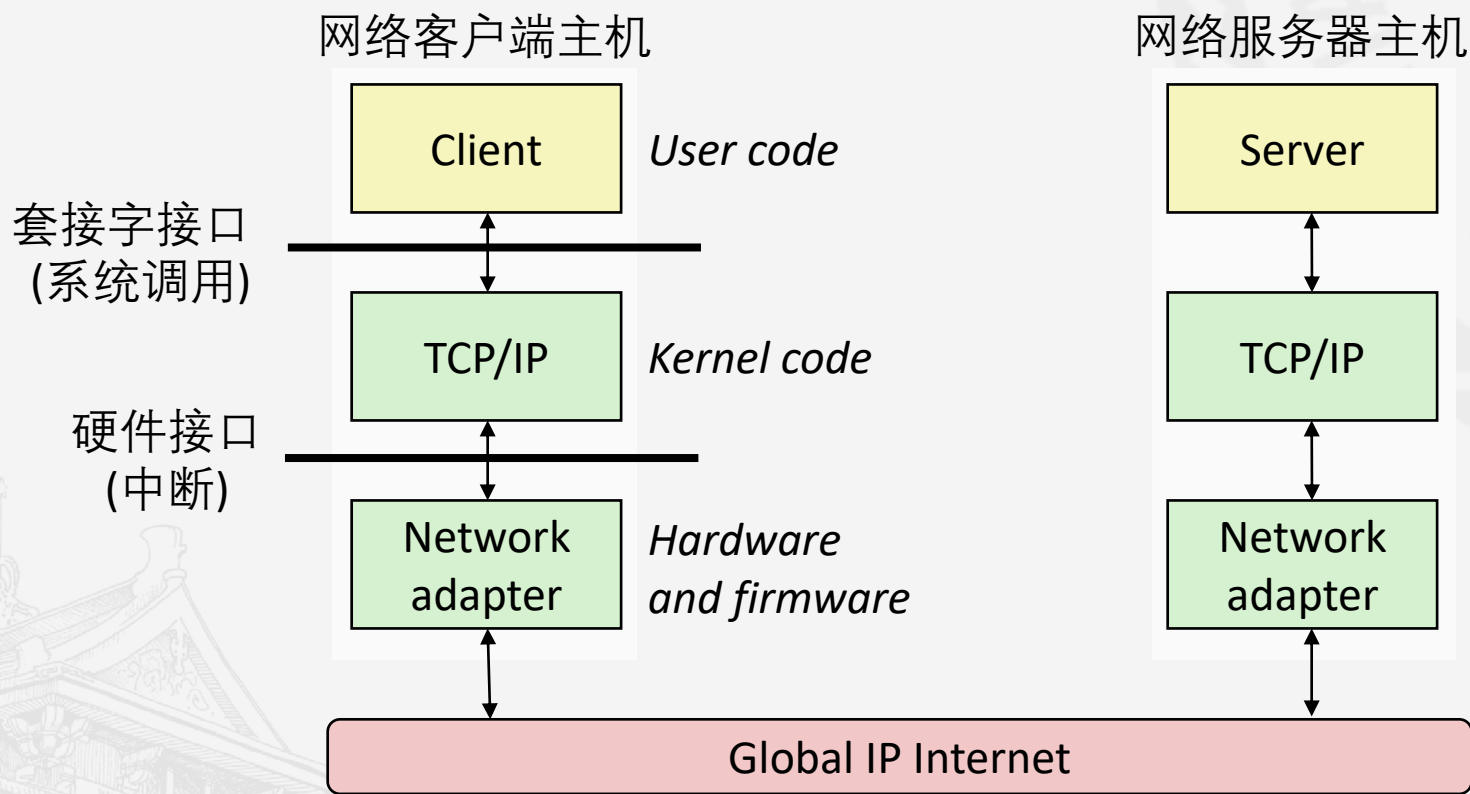
■ 如果数据包丢失了怎么办？

■ 计算机网络领域的专家们已经给出了解决方案



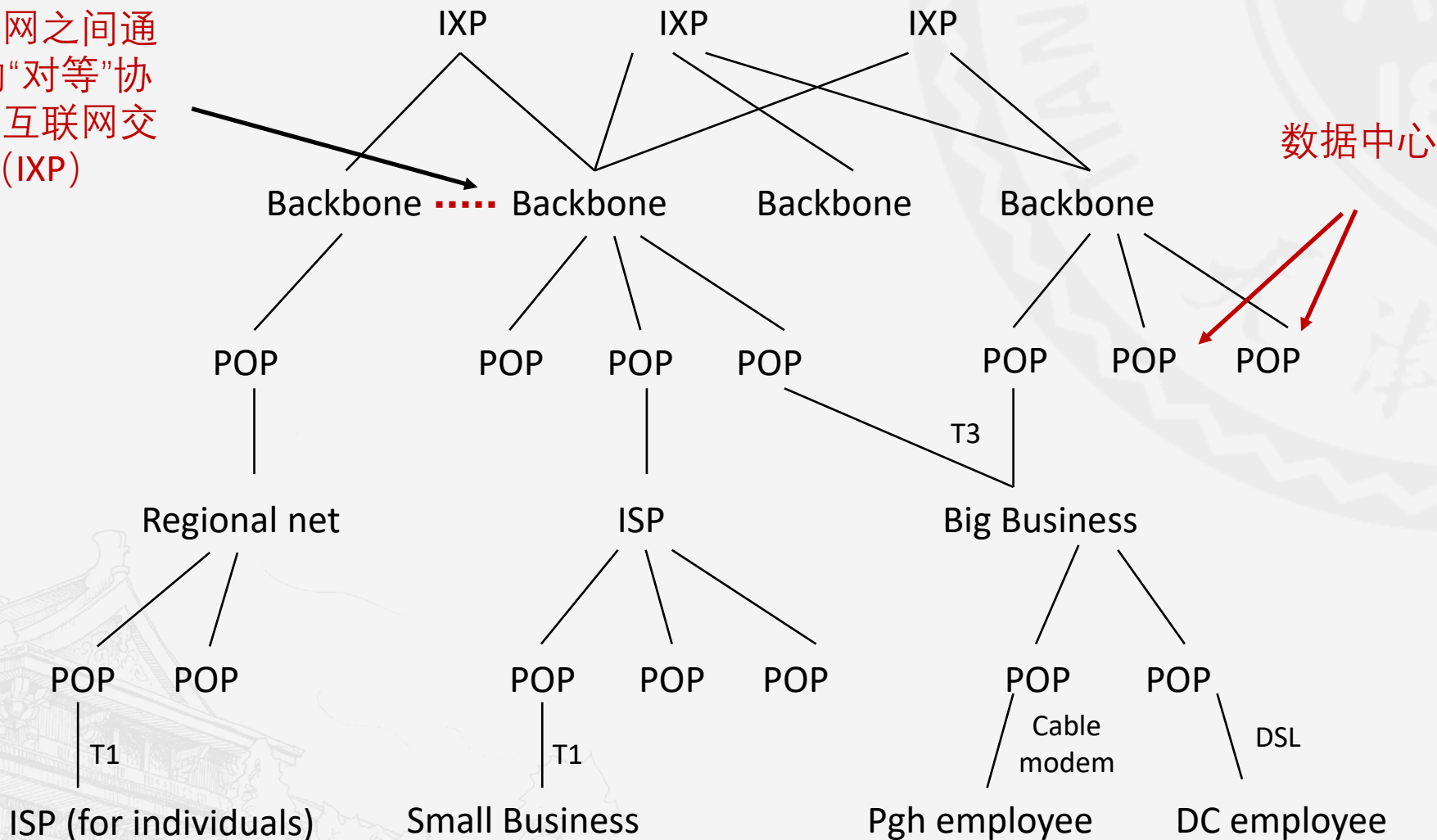
全球IP因特网（Internet）

- 最著名的互联网应用
- 基于TCP/IP协议族：
 - IP（Internet协议）：提供基本的命名方案和不可靠的从主机到主机的数据包（数据报）传输能力。
 - UDP（不可靠数据报协议）：使用IP提供从进程到进程的不可靠数据报传输。
 - TCP（传输控制协议）：使用IP提供从进程到进程的可靠字节流传输。
- 通过Unix文件I/O和套接字接口的函数混合访问互联网。



- 骨干网（Internet backbone）：
 - 由高速点对点网络连接的路由器集合（全国范围或全球范围）。
- 互联网交换点（Internet Exchange Points，IXP）：
 - 连接多个骨干网的路由器（也会采用对等互联方案）。
 - 也称为网络接入点（NAP）。
- 区域网络：覆盖较小地理区域（例如城市或省）的较小骨干网。
- 出现点（Point of presence，POP）：连接到互联网的机器。
- 互联网服务提供商（Internet Service Providers，ISP）：提供拨号或直接访问POP的服务。

两个骨干网之间通过私有的“对等”协议，绕过互联网交换点（IXP）





- 主机映射到一组32位的**IP地址**:
 - 128.2.203.179
- 这组IP地址映射到一组的标识符，这些标识符被称为**域名**:
 - 128.2.203.179映射到www.cs.cmu.edu
- 互联网上的一个主机上的进程可以通过**网络连接**与另一个主机上的进程进行通信



- 原始互联网协议使用32位地址，被称为互联网协议版本4 (IPv4)。
- 1996年：互联网工程任务组 (IETF) 引入了使用128位地址的互联网协议版本6 (IPv6)，意在替代IPv4。
- 截至2015年，绝大多数互联网流量仍然由IPv4承载。
 - 只有4%的用户使用IPv6访问Google服务。
- 后面的课程中会以IPv4为主，并介绍如何编写与协议无关的网络代码。

- 32位IP地址存储在一个IP地址结构体中。in_addr
- IP地址总是以网络字节序（大端字节序）存储在内存中
- 一般来说，任何从一台机器传输到另一台机器的数据包头中的整数都是使用网络字节序
 - 例如：用于标识互联网连接的端口号。

```
/* Internet address structure */  
struct in_addr {  
    uint32_t  s_addr; /* network byte order (big-endian) */  
};
```

点分十进制表示法

- 按照约定，32位IP地址中的每个字节都以其十进制值表示，并由句点分隔。
 - 例如：0x8002C2F2 = 128.2.194.242
- 使用getaddrinfo和getnameinfo函数（稍后会介绍）在IP地址和点分十进制格式之间进行转换。

- IPv4地址空间被分为五种类型：

	0	1	2	3	8	16	24	31	
Class A	0	Net ID			Host ID				
Class B	1	0	Net ID			Host ID			
Class C	1	1	0	Net ID			Host ID		
Class D	1	1	1	0	Multicast address				
Class E	1	1	1	1	Reserved for experiments				

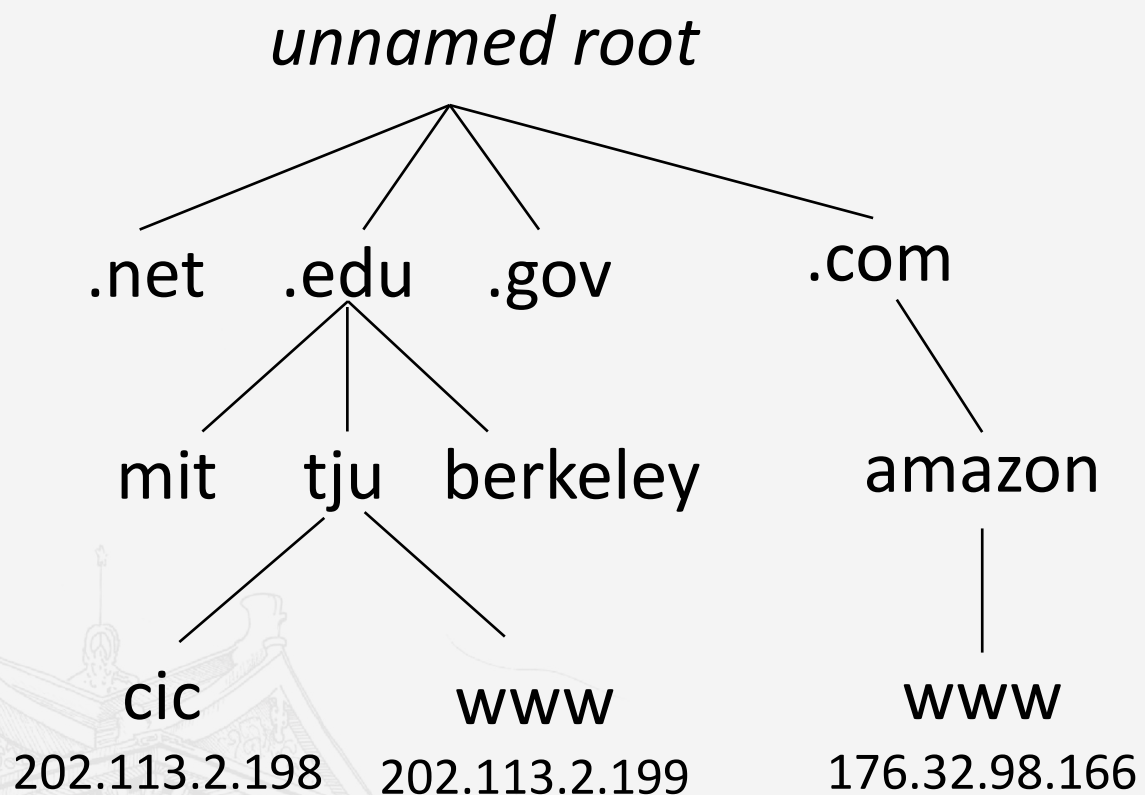
- 网络ID以w.x.y.z/n的形式表示：

- n是主机地址中的位数。

- 例如，卡耐基梅隆大学（CMU）的网络ID表示为128.2.0.0/16（B类地址）

- 私有IP地址包括：

10.0.0.0/8、172.16.0.0/12、192.168.0.0/16



一级域名

二级域名

三级域名

- 互联网在一个称为DNS的全球分布式数据库中维护着IP地址和域名之间的映射关系。
- 从概念上讲，程序员可以将DNS数据库视为由数百万个主机条目组成的集合。
 - 每个主机条目定义了域名和IP地址之间的映射关系。
 - 从数学角度看，一个主机条目是域名和IP地址的等价类。

- 可以使用nslookup来探索DNS映射的关系
 - 输出的结果经过一定程度的简化
- 每个主机都有一个本地定义的域名localhost, 它总是映射到回环地址127.0.0.1。

```
> nslookup localhost  
Address: 127.0.0.1
```

- 使用hostname命令可以确定本地主机的真实域名:

```
> hostname  
whaleshark.ics.cs.cmu.edu
```

DNS映射

```
youmeng@LYM-DESKTOP2:~$ nslookup www.baidu.com  
Server:          172.21.80.1  
Address:         172.21.80.1#53  
  
Non-authoritative answer:  
www.baidu.com    canonical name = www.a.shifen.com.  
Name:   www.a.shifen.com  
Address: 182.61.200.6  
Name:   www.a.shifen.com  
Address: 182.61.200.7  
Name:   www.a.shifen.com  
Address: 240c:4003:111:f4c0:0:ff:b07c:1608  
Name:   www.a.shifen.com  
Address: 240c:4003:111:3f1f:0:ff:b0ea:3686  
Name:   dns.baidu.com  
Address: 110.242.68.134  
Name:   ns2.baidu.com  
Address: 220.181.33.31  
Name:   ns3.baidu.com  
Address: 153.3.238.93  
Name:   ns3.baidu.com  
Address: 36.155.132.78  
Name:   ns4.baidu.com  
Address: 14.215.178.80  
Name:   ns4.baidu.com  
Address: 111.45.3.226
```

- 简单情况：域名和IP地址之间的一对一映射：

```
> nslookup whaleshark.ics.cs.cmu.edu  
Address: 128.2.210.175
```

- 多个域名映射到相同的IP地址：

```
> nslookup cs.mit.edu  
Address: 18.62.1.6  
> nslookup eecs.mit.edu  
Address: 18.62.1.6
```

■ 多个域名映射到多个IP地址:

```
> nslookup www.twitter.com
Address: 199.16.156.6
Address: 199.16.156.70
Address: 199.16.156.102
Address: 199.16.156.230

> nslookup twitter.com
Address: 199.16.156.102
Address: 199.16.156.230
Address: 199.16.156.6
Address: 199.16.156.70
```

■ 有些有效的域名不映射到任何IP地址:

```
> nslookup ics.cs.cmu.edu
*** Can't find ics.cs.cmu.edu: No answer
```

- 客户端和服务端通过在连接上发送字节流进行通信。每个连接都包含以下特征：
 - 点对点：连接一对进程。
 - 全双工：数据可以同时两个方向上流动。
 - 可靠：源发送的字节流最终将按发送顺序被目标接收。
- 一个套接字是一个连接的端点。
 - 套接字地址是IP地址和端口的组合 `ip:port`。
- 一个端口是一个16位整数，用于标识一个进程：
 - 临时端口：当客户端发出连接请求时，由客户端内核自动分配。
 - 知名端口：与服务器提供的一些服务相关联（例如，端口80与Web服务器相关联）。

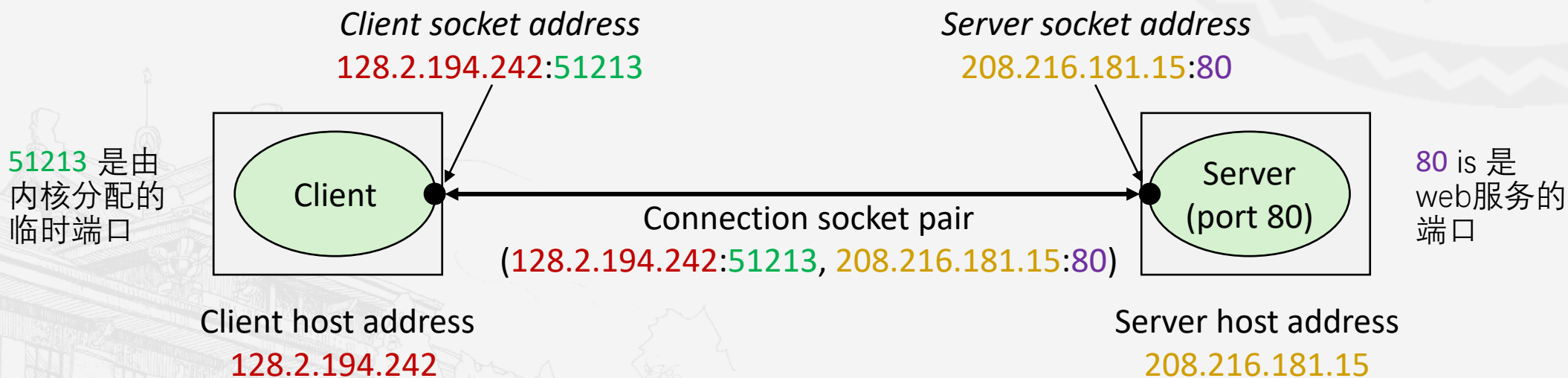


知名端口和服务名称

- 一些主流的服务有固定的端口号
 - 回显服务：7/echo
 - SSH服务：22/ssh
 - 邮件服务：25/smtp
 - Web服务：80/http
- 每台Linux上的/etc/services文件中包含知名端口和服务名称之间的映射关系。

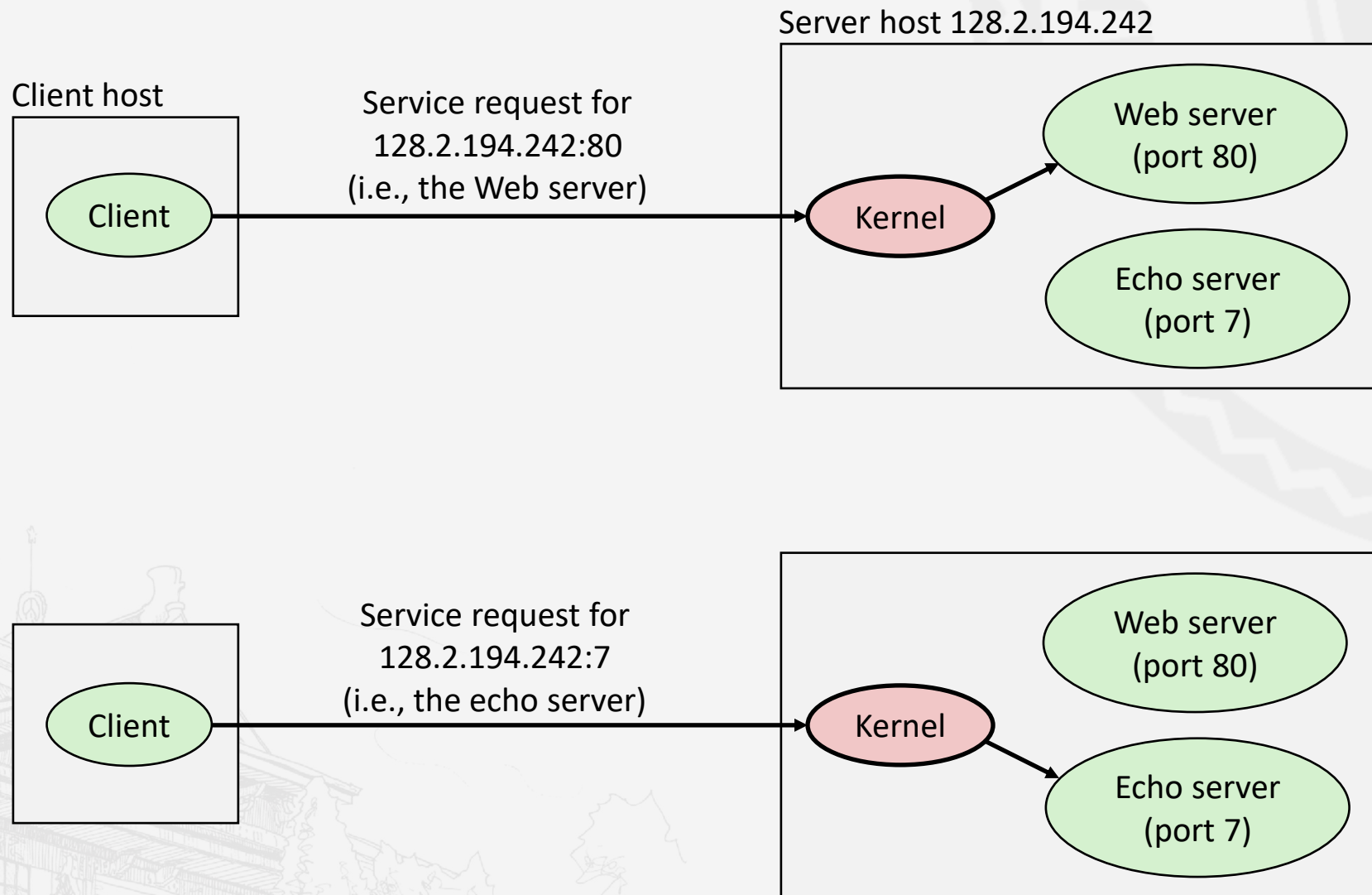
连接的结构

- 一个连接通过其端点 (endpoint) 的套接字地址 (套接字对) 唯一标识:
 - (cliaddr:cliport, servaddr:servport)
 - endpoint 指的是连接的端点, 即数据传输开始或结束的地方





使用端口识别服务





■ 最初想法：

■ 每个节点都有唯一的 IP 地址

- 每个节点都可以直接与其他节点通信。

■ 没有保密或身份验证

- 消息对于同一局域网（LAN）上的路由器和主机都是可见的
- 数据包头中的源字段可能被伪造。

■ 局限性：

- **IP 地址不足：**随着互联网的发展，最初设计的 IPv4 地址空间已经不能满足全球范围内的需求。

- **隐私问题：**不希望每个人都能访问或了解所有其他主机的信息

- **安全问题：**需要确保数据传输的保密性和身份验证，以防止恶意攻击和数据泄露。



■ 动态地址分配:

- 大多数主机不需要已知地址: 只有那些作为服务器运行的主机需要已知地址。
- DHCP (动态主机配置协议): 本地互联网服务提供商 (ISP) 为分配的临时地址。

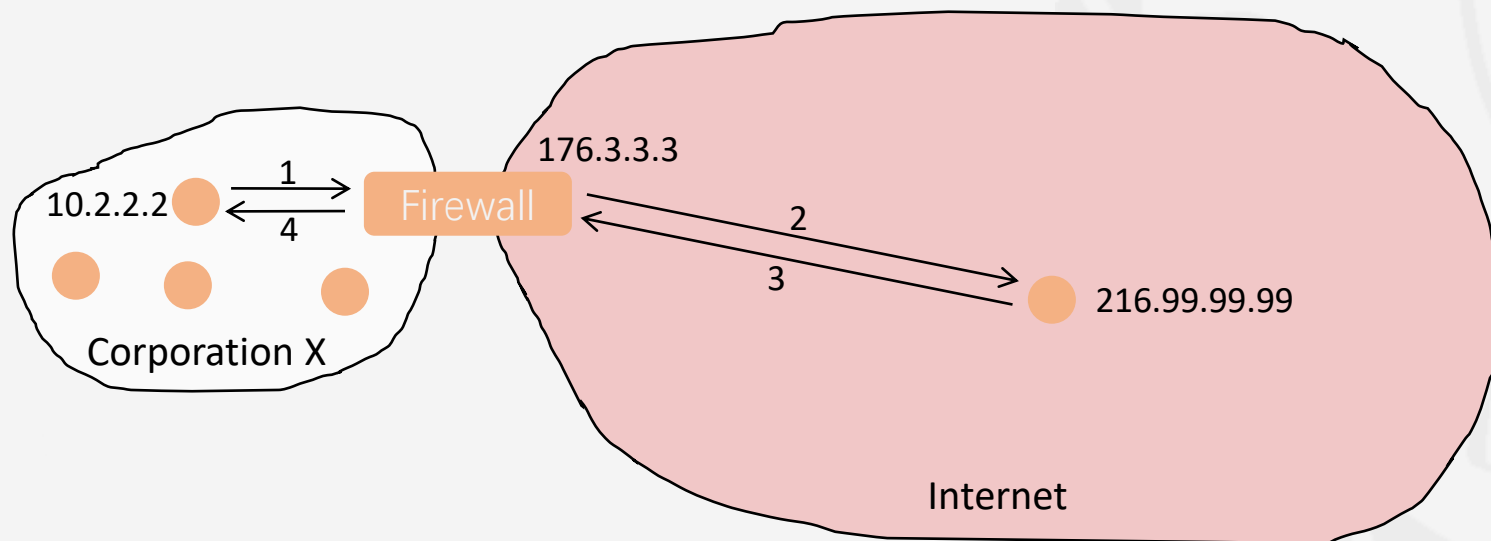
■ 示例:

■ 智能与计算学部主页:

- IP 地址为 202.113.2.198 (cic.tju.edu.cn) 。
- 静态分配。

■ 家中的笔记本电脑:

- IP 地址为 192.168.1.5。
- 仅在家庭网络中有效。



- 将组织的节点隐藏起来
- 在组织内使用局部 IP 地址
- 为外部服务提供代理服务

1. 客户端请求：源地址=10.2.2.2，目的地址=216.99.99.99。
2. 防火墙转发请求：源地址=176.3.3.3，目的地址=216.99.99.99。
3. 服务器响应：源地址=216.99.99.99，目的地址=176.3.3.3。
4. 防火墙转发响应：源地址=216.99.99.99，目的地址=10.2.2.2。