域名系统概述

- 互联网中使用IP地址寻址主机（例如：IP数据包转发）
- 为了方便记忆，每台提供服务的主机通常会有一个或多个名字
  - 例如：访问学院网站，输入名字cc.nankai.edu.cn
    - 对应的IPv4地址：222.30.45.190
    - 对应的IPv6地址：2001:250:401:d450::190

查询命令：nslookup cc.nankai.edu.cn

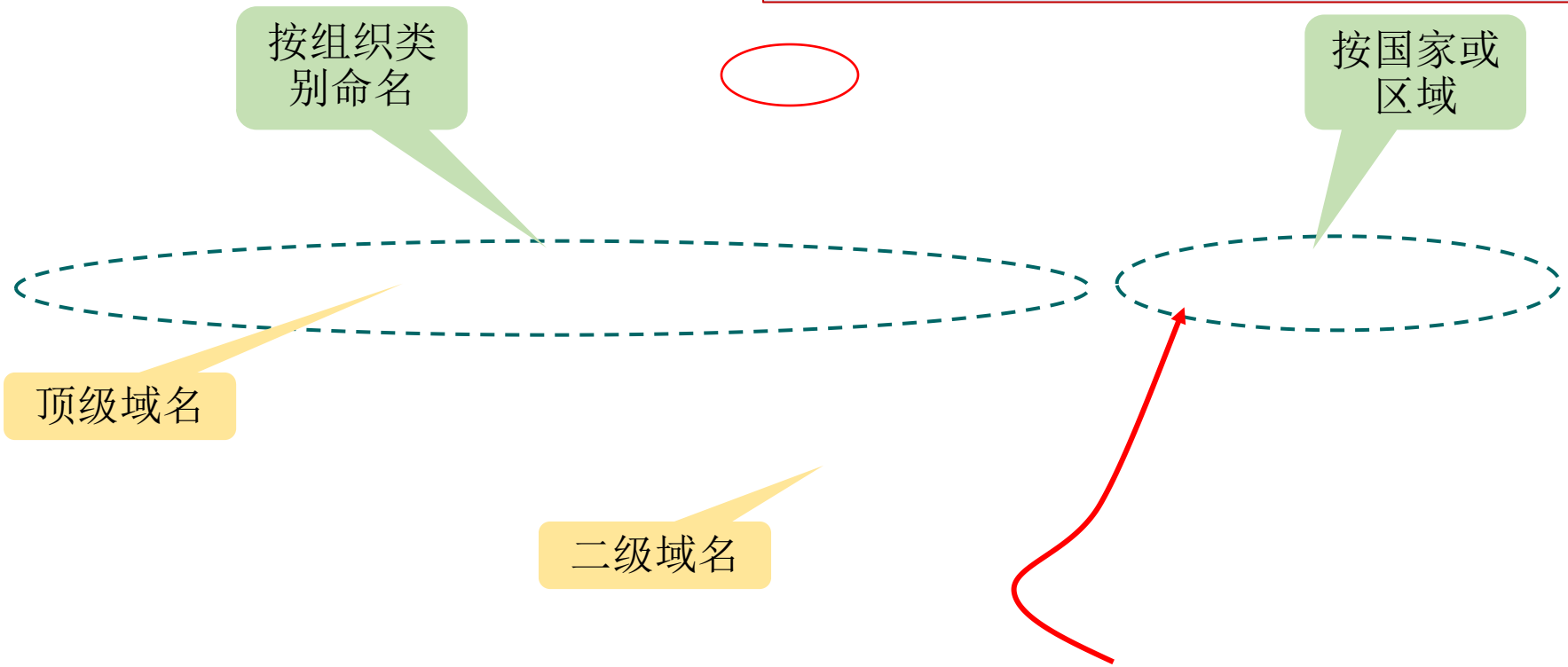
- 如何将名字映射到地址？
  - 早期的集中式管理和发布
    - 本地存储Hosts文件，实现名字到地址的静态映射
    - 可以通过FTP服务为连入Internet的主机提供域名的发布和下载

域名系统概述（续）

- DNS（Domain Name System）
  - 自动实现名字到地址映射的系统
- DNS基本思想：
  - 名字和地址映射关系分布式存放，形成具有层次结构的分布式数据库系统（分布式管理）
  - 通过查询分布式数据库，获得名字到地址的映射，或相反
- 关键：
  - 如何组织分布式数据库？
  - 如何在分布式数据库中查找？

DNS域名体系

优点：可以支持大规模、快速扩展，不需要中心节点支持；通过部分名字空间的授权实现非集中式管理；名字到地址的映射可以通过分布方式实现



- 层级命名、逐级授权、多级管理
  - 例如：cc.nankai.edu.cn

域名已成为互联网的重要基础性资源

DNS域名体系（续）

- 顶级域名：由互联网名称与数字地址分配机构(ICANN)负责管理
  - ICANN 与域名注册商签订合同， 准许其受理顶级域名.com、.net、.org 下的域名注册

顶级域名	分配
com	商业组织
edu	教育机构
gov	政府机构
mil	军队机构
net	主要的网络支持中心
org	其他组织
Int	国际组织
国家地区代码	国家或地区

DNS域名体系（续）

- 中国互联网络信息中心（CNNIC）管理.cn域名

划分模式	我国二级域名	分配
类别域名	ac	科研机构
	com	工、商、金融等企业
	edu	教育机构
	gov	政府部门
	net	互联网络、接入网络信息中心和运行中心
	org	各种非盈利性的组织
行政区域域名	bj	北京市
	sh	上海市
	tj	天津市
	cq	重庆市
	……	

Whois查询服务（TCP 43端口）：中国万网（www.zw.cn）、站长之家（whois.chinaz.com）等

DNS域名解析

- **域名解析**：名字到地址映射（通过名字查地址）
  - 分布式：层级的服务器组织，协同实现解析
  - 有效性：大多数解析可以在本地完成，一部分会产生互联网流量
  - 可靠性：通过冗余设置，避免单点失效
- **客户-服务器模式**
  - 域名服务器：
    - 保存名字到地址映射关系（数据库）
    - 接收客户端请求，并给出响应
  - 域名解析器（客户端）：
    - 请求域名解析的客户进程
    - 向域名服务器发起解析请求，并等待服务器的响应

DNS服务器组织

- 每台域名服务器包含一个或多个区域的信息
- 父节点服务器已知子节点服务器的地址

DNS服务器组织

根域名服务器对互联网发展至关重要，除IPv4时代的13个根服务器，目前在16个国家设立了25台IPv6根服务器，**我国部署了4台**，打破了我国无根服务器的困境

根域名服务器：

- 全球13个逻辑根域名服务器（a~m），每个根服务器都有多个镜像，实际的服务器数量目前达1326个（中国：大陆13个，台湾6个，香港9个）

根服务器：*letter.root-servers.net*

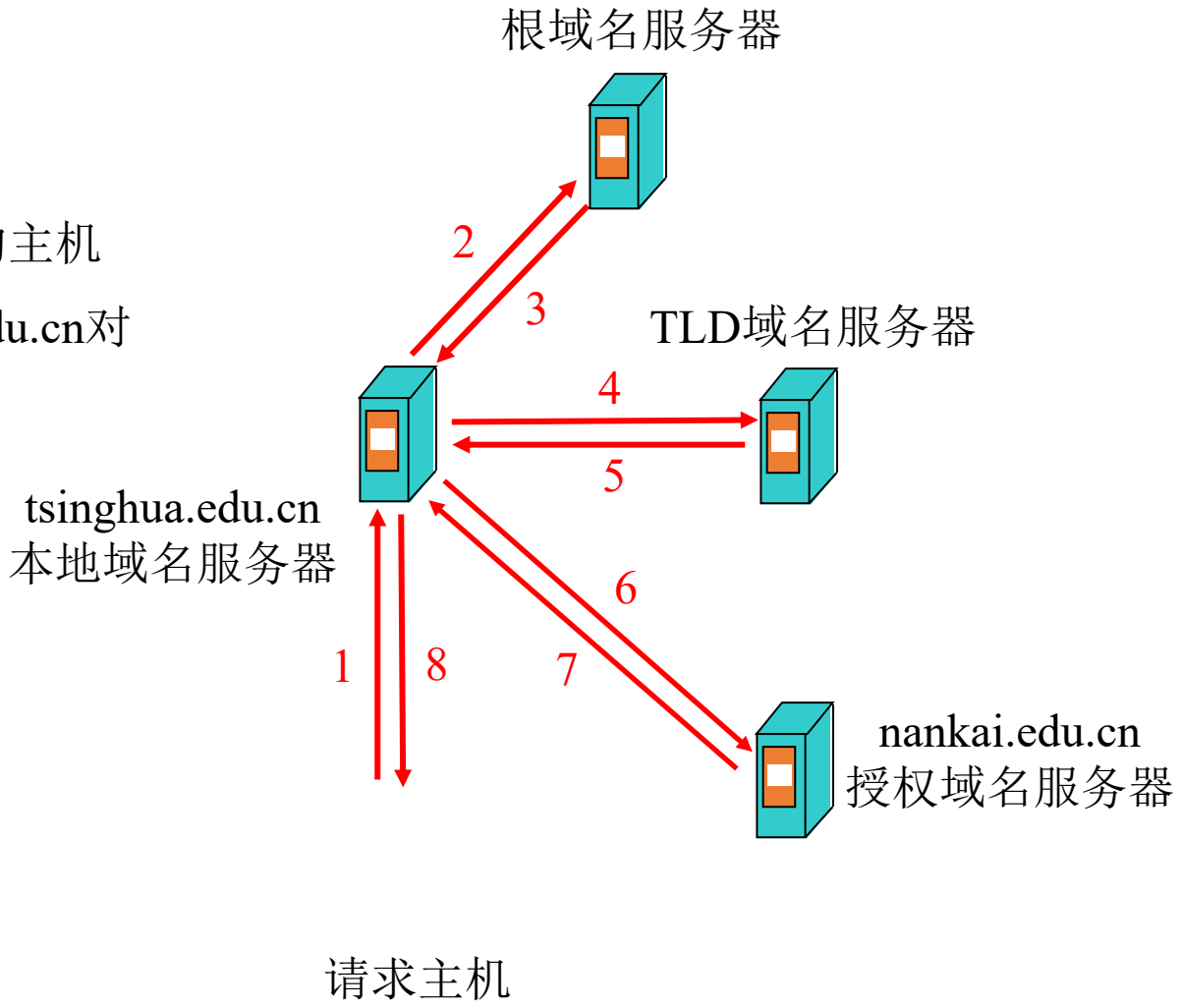
https://root-servers.org/

DNS服务器组织

- **顶级域名服务器**（Top-Level Domain, TLD）
  - 负责顶级域名的解析
- **授权域名服务器**
  - 对于名字与地址映射，保留其初始数据来源的服务器
  - 主要区分名字与地址映射是原始的还是被缓存的（非授权）
- **本地域名服务器**（或称默认域名服务器）
  - 一般每个ISP都部署有域名服务器，其用户可将该服务器设置成本地域名服务器（或默认域名服务器）
  - 当进行域名解析时，查询请求首先发送到本地域名服务器（即查询的起点）

DNS域名解析示例

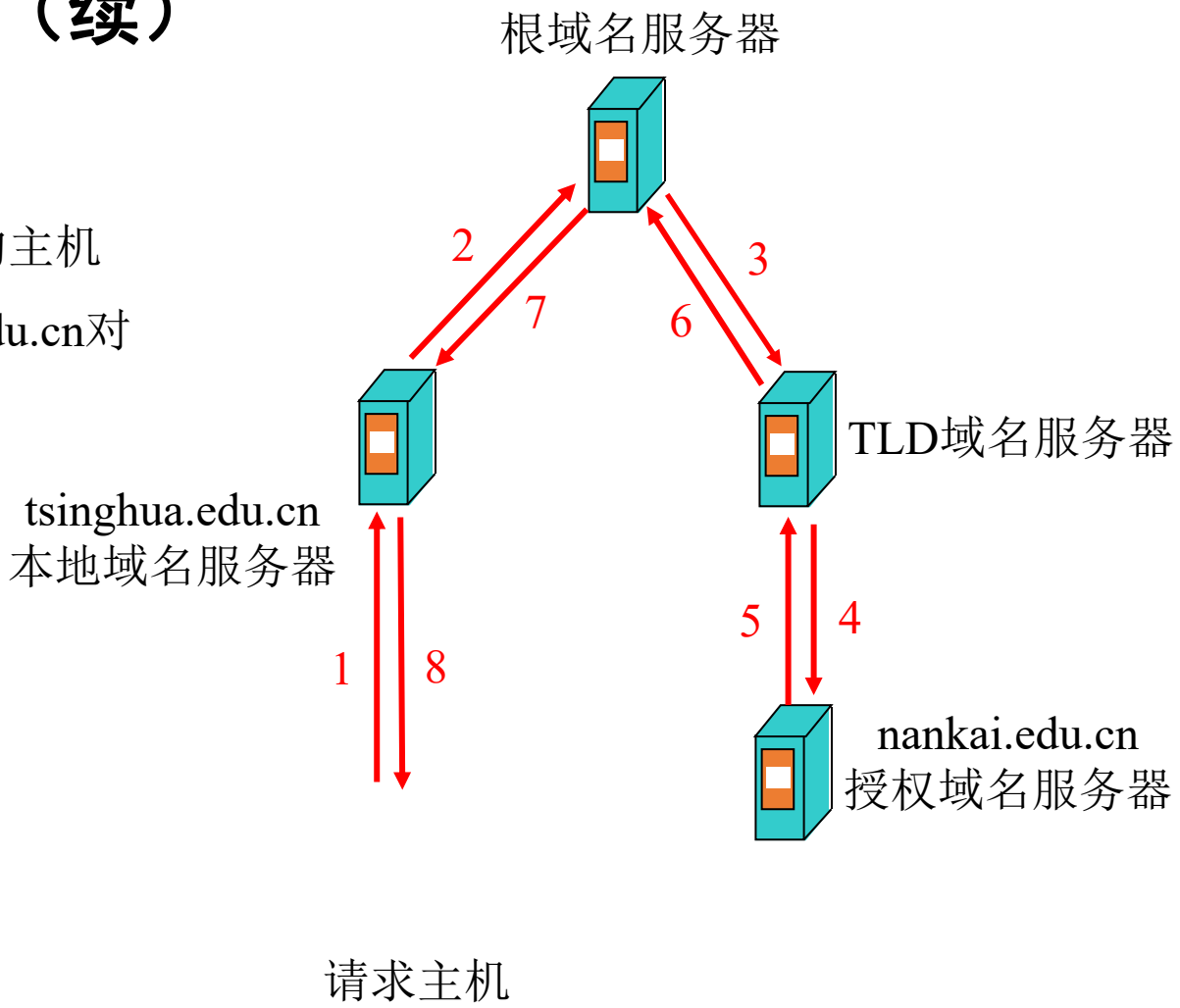
- 例如：
  - tsinghua.edu.cn域中的主机要解析www.nankai.edu.cn对应的IP地址
- 解析过程
  - 反复解析
  - 递归解析



www.nankai.edu.cn

DNS域名解析示例（续）

- 例如：
  - tsinghua.edu.cn域中的主机要解析www.nankai.edu.cn对应的IP地址
- 解析过程
  - 反复解析
  - 递归解析



www.nankai.edu.cn

DNS域名服务器缓存

- 目的：降低非本地名字查询开销及查询延时
  - 通常地址与名字的绑定变化不频繁
- 服务器缓存名字与地址映射关系
  - 服务器学习到某个名字和地址的映射关系时，便进行缓存
  - 记录名字和地址的映射从何处获取
  - 基于授权服务器中的TTL值设置超时时间，缓存的映射关系经过一定时间会超时
  - TLD服务器通常会被本地域名服务器缓存，可以有效减少根域名服务器的访问频度

DNS域名服务器缓存（续）

- 服务器使用缓存的映射关系响应客户端的请求
  - 标记为非授权（*nonauthoritative*）映射
  - 给出获取映射的服务器的域名和IP地址
- 客户机接收服务器响应
  - 映射有可能过时
  - 如果注重效率，客户端接受非授权响应
  - 如果注重准确性，客户机可以再联系授权服务器，验证映射是否仍有效

主机缓存

- 基本方法
  - 在启动时可以从本地域名服务器下载名字-地址映射数据库，并定期获取新的映射
  - 缓存最近用过的名字和地址映射
- 优点
  - 无需访问域名服务器，名字解析速度快
  - 本地服务器的故障不影响名字解析
  - 减低服务器的负载

DNS资源记录

- DNS使用区域数据库存储名字与地址映射关系，区域数据库由资源记录（Resource Records，RR）组成
- 资源记录结构
  - **名字（name）**
  - TTL（Time To Live）：有效时间，通常为86400秒（24小时）
  - 类型（Type）：SOA、NS、A、AAAA、PTR、CNAME、MX
  - 类（Class）：例如，IN类
  - **值（Value）**

DNS资源记录（续）

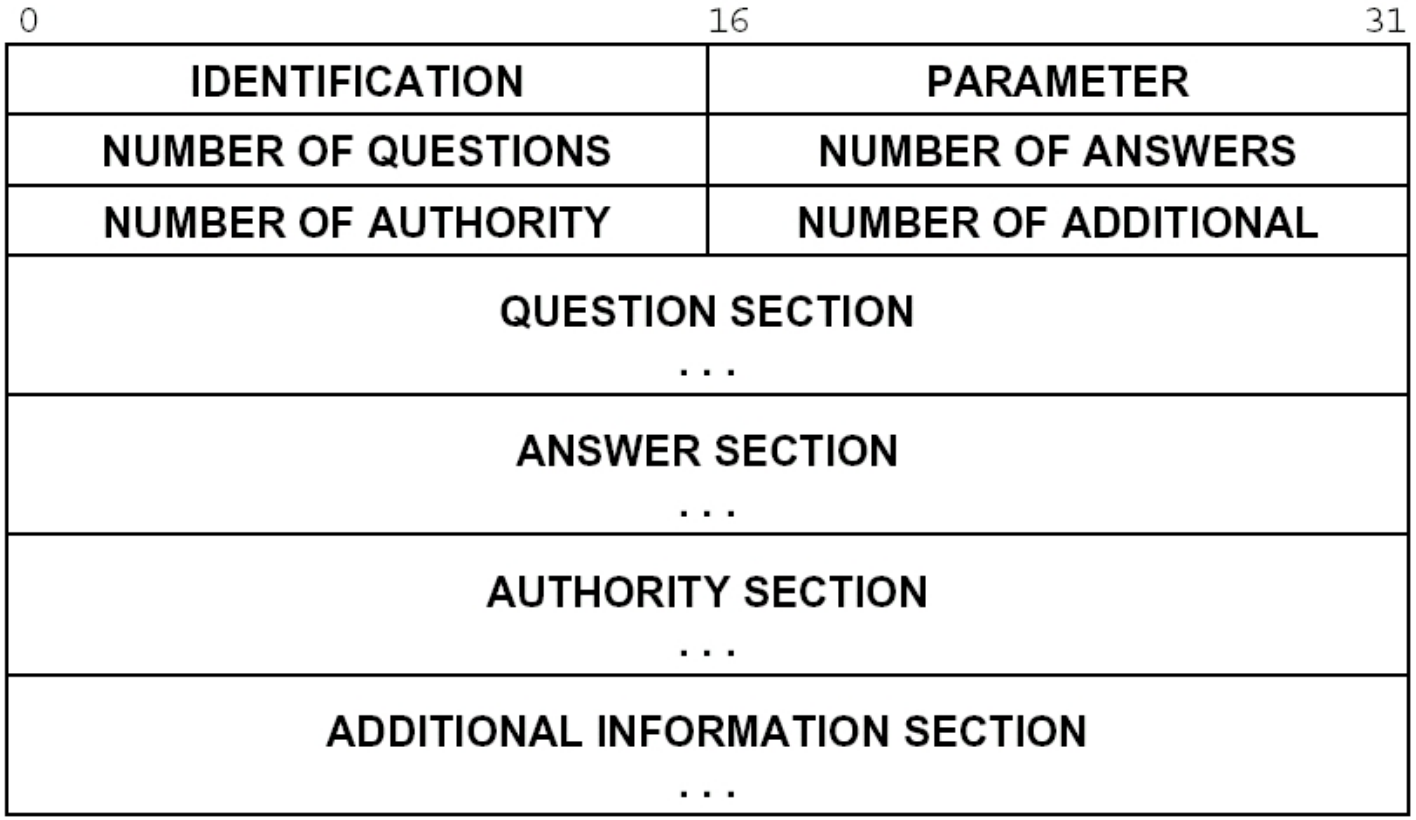
- 资源记录类型（常用）
  - SOA：区域数据库的开始，描述负责区域的域名服务器、版本信息，以及从属域名服务器备份时的一些参数等
  - NS：指定DNS服务器主机名（不使用IP地址）
  - A：将名称对应到IPv4的32位地址
  - AAAA：将名称对应到IPv6的128位地址
  - PTR：将IP 对应的名字
  - CNAME：别名，同一台主机可以有多个名字
  - MX：给出服务特定域的邮件服务器的主机名

DNS资源记录：示例

cs.vu.nl.	86400	IN	SOA	star boss (serial, refresh, retry, expire, ttl)
cs.vu.nl.	86400	IN	TXT	“A University”
cs.vu.nl.	86400	IN	MX	1 zephyer.cs.vu.nl.
cs.vu.nl.	86400	IN	MX	2 top.cs.vu.nl.
flits.cs.vu.nl.	86400	IN	HINFO	Sun Unix
flits.cs.vu.nl.	86400	IN	A	130.37.16.112
flits.cs.vu.nl.	86400	IN	MX	1 flits.cs.vu.nl.
flits.cs.vu.nl.	86400	IN	MX	2 zephyer.cs.vu.nl.
flits.cs.vu.nl.	86400	IN	MX	3 top.cs.vu.nl.
www.cs.vu.nl.	86400	IN	CNAME	top.cs.vu.nl.
ftp.cs.vu.nl.	86400	IN	CNAME	zephyer.cs.vu.nl.
zephyer	86400	IN	A	130.37.56.201
		IN	HINFO	Sun Unix

DNS报文格式

- DNS包括*query*和*reply*两种报文



DNS报文格式（续）

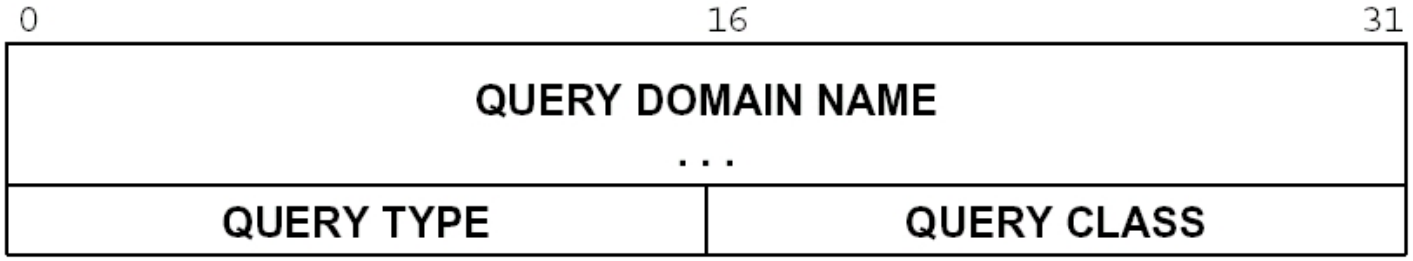
- 参数域的定义

参数域中的位	含义
0	报文类型：0-query 1-reply
1-4	查询类型： 0-标准查询 1-反向查询
5	授权响应，则置1
6	报文被截断，则置1
7	期望递归，则置1
8	支持递归，则置1
9-11	保留
12-15	应答类型： 0-无错误 1-查询中格式错 2-服务器失效 3-名字不存在

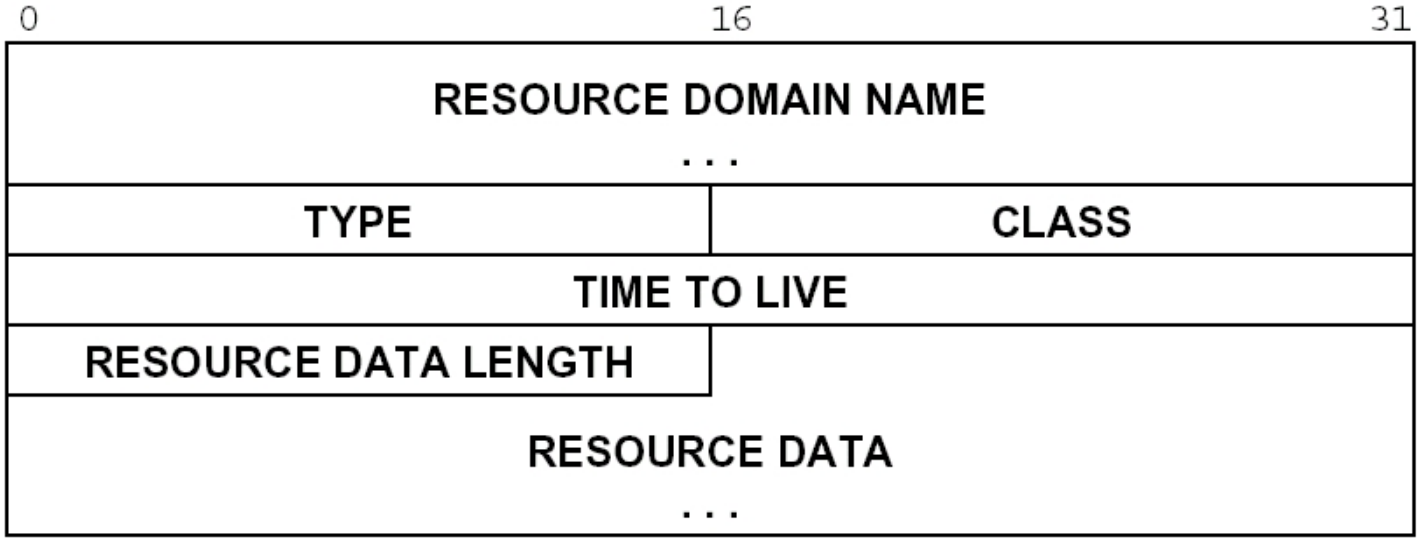
DNS报文格式（续）

- Question: 携带查询的名字和其他参数
- Answer: 携带直接响应查询的资源记录
- Authority: 携带描述其他域名服务器的资源记录
  - 在应答中可以选择携带授权数据的SOA资源记录
- Additional:携带附加的资源记录

DNS报文格式（续）



Question
格式



Answer
Authority
Additional
格式

域名格式压缩

- 报文中域名格式
  - 每个段第一个字节指定长度 (00xxxxxx=n)，后跟n个字节
  - 长度为0的段，表明域名结束
  - 例如：www.nankai.edu.cn，nankai.edu.cn
- 压缩格式
  - 域名的后缀部分经常重复，可以进行适当压缩
  - 指针：如果前两位为11，则段前两个字节的后14位为指针 (11xxxxxxxxxxxxxx)

域名格式压缩：示例

- 例如：需要查询名字F.ISI.ARPA, FOO.F.ISI.ARPA, ARPA, 和根，忽略其他域，这些名字可以表示为：

偏移量

20	1	F
22	3	I
24	S	I
26	4	A
28	R	P
30	A	0

40	3	F
42	O	O
44	11	20
64	11	26
92	0	

报文格式：示例

- 例如：一个邮件的客户端想向Mockapetris@ISI.EDU发送邮件，需要对域名ISI.EDU进行解析
  - Query: question section

    QNAME=ISI.EDU, QTYPE=MX, QCLASS=IN
  - Reply: answer section

    ISI.EDU. MX 10 VENERA.ISI.EDU.

    MX 10 VAXA.ISI.EDU.
  - Reply: additional section

    VAXA.ISI.EDU. A 10.2.0.27

    A 128.9.0.33

    VENERA.ISI.EDU. A 10.1.0.52

    A 128.9.0.32

报文格式：示例

- 例如：一个查询 (QNAME=BRL.MIL, QTYPE=A)发送到C.ISI.EDU，则应答可能为：
  - answer section

    <empty>
  - authority section

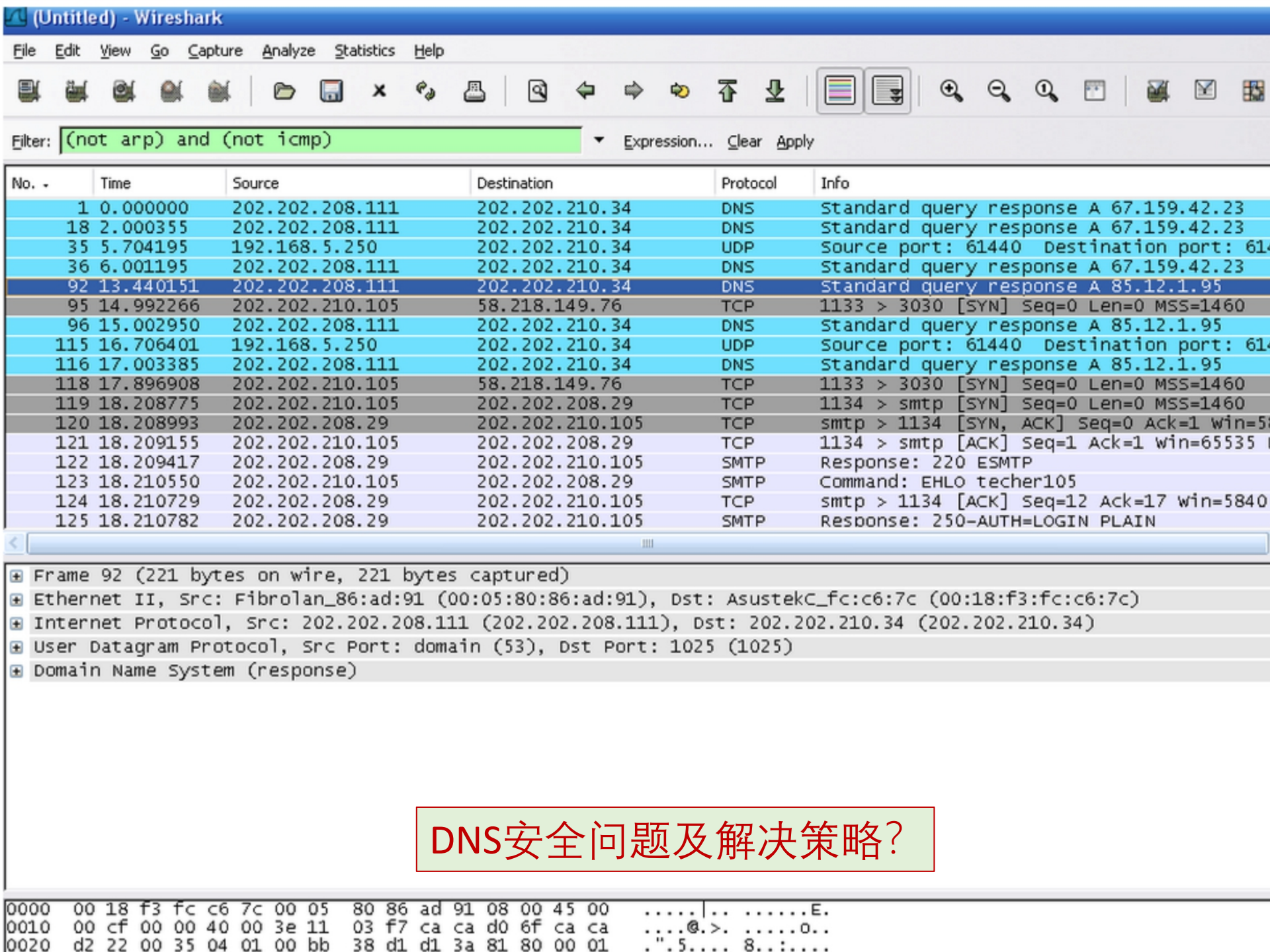
    MIL. 86400 IN NS SRI-NIC.ARPA.

    86400 IN NS A.ISI.EDU.
  - additional section

    A.ISI.EDU. A 26.3.0.103

    SRI-NIC.ARPA. A 26.0.0.73

    A 10.0.0.51



DNS安全问题及解决策略?