

现代计算机网络

Ch.2 IPv6

2

- 2.1 简介
 - ▣ 2.1.1 IPv4的问题
 - ▣ 2.1.2 IPv6 设计目标和主要特征
- 2.2 IPv6 的报文结构
- 2.3 IPv6 的地址空间
- 2.4 IPv6 邻居发现协议
- 2.5 IPSec
- 2.6 IPv4到IPv6过渡



Ch.2 IPv6简介

3

2.1.1 IPv4的问题

- 互联网正在成为其自身发展的牺牲品，几乎每隔20s就有一台主机加入
- 全球 IPv4地址已经于2011年8月底耗尽
- 一开始地址分配不合理，“三只熊”问题：
- 每个地址重 1Gram
 - A类16777216: 10^4 Kg;
 - B类65536: 10^2 Kg;
 - C类256: 10^0 Kg
- CIDR推出较晚，效果有限



A= 10^4 Kg



B= 10^2 Kg

C= 10^0 Kg



Ch.2 IPv6简介

4

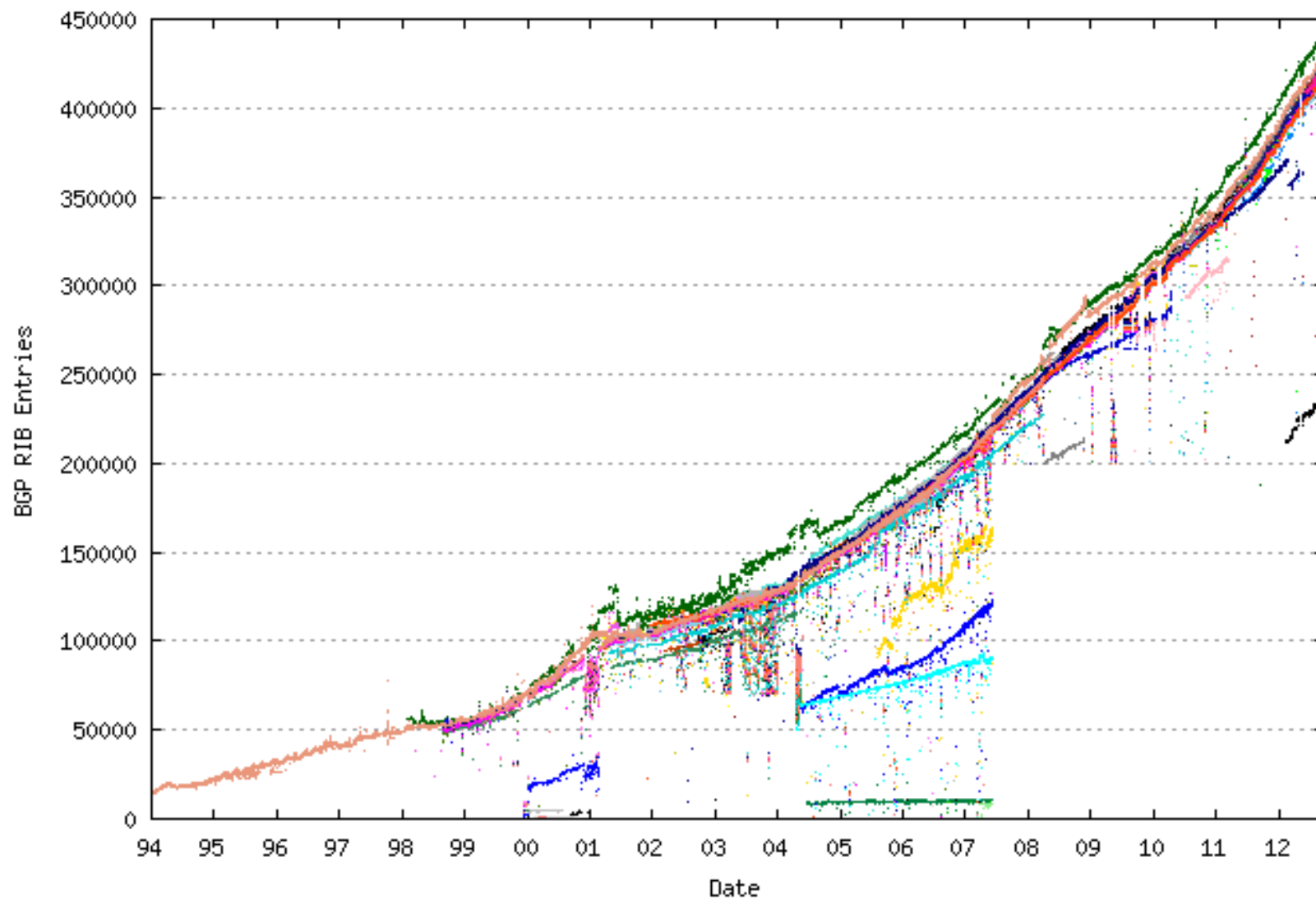
□ 2.1.1 IPv4的问题

- ▣ **Classless Inter-Domain Routing (CIDR)** is a method for allocating IP addresses and IP routing.
- ▣ The Internet Engineering Task Force introduced CIDR in 1993 to replace the previous addressing architecture of classful network design in the Internet.
- ▣ 192.168.100.0/22 represents the 1024 IPv4 addresses from 192.168.100.0 to 192.168.103.255



路由表项剧增Active BGP entries (FIB)

5



IPv4的问题（继续）

- 对现有路由技术的支持不够：IPv4头长度不固定(0-40字节)**不利于ASIC处理**；没有利用包前后的相关性，每个包进行同样处理；MTU导致**分段和逐段**校验，路由处理慢
- 无法提供多样的QoS：最大努力最短时间，但不保证是否进行和何时进行，IP尽力而为的FIFO对实时多媒体信息的处理会带来延迟、间断，无法满足多媒体传输质量的要求
- 地址限制无法满足移动设备,家电,传感网络和因特网的连接：HPC/PDA将占计算机总数的50%
- 安全支持问题：源地址伪造、IP/TCP报文是明文、TCP劫持

2.1.2 IPv6 设计目标和主要特征

7

- 扩大地址空间、路由更结构层次化
 - ▣ 32bits → 128 bits
 - ▣ 全局unicast地址等价于IPv4公开地址
 - ▣ 直接使用CIDR，网络前缀取代掩码，前缀表示子网号
- 报头格式大简化，方便硬件处理
 - ▣ 基本报头固定40bytes
 - ▣ 简化路由器的操作
 - ▣ 引入结构化扩展报头，取消可选项长度限制

□ 网络管理 更加简单

- 建立一系列自动发现和自动配置功能
- 最大单元发现 (MTU discovery)
- 邻接节点发现 (neighbor discovery)
- 路由器通告 (router advertisement)
- 路由器请求 (router solicitation)
- 节点自动配置 (auto-configuration)

□ 安全性支持

- IP security ,提供IP层的安全IPSec
- 实现认证头 (Authentication Header)
- 安全载荷封装 (Encapsulated Security Payload)

□ QoS能力

- ▣ **流标号**（flow label），20比特,发送者可以要求路由器对此流进行特殊处理，路由器可以鉴别特殊流的所有报文

□ **多播**寻址

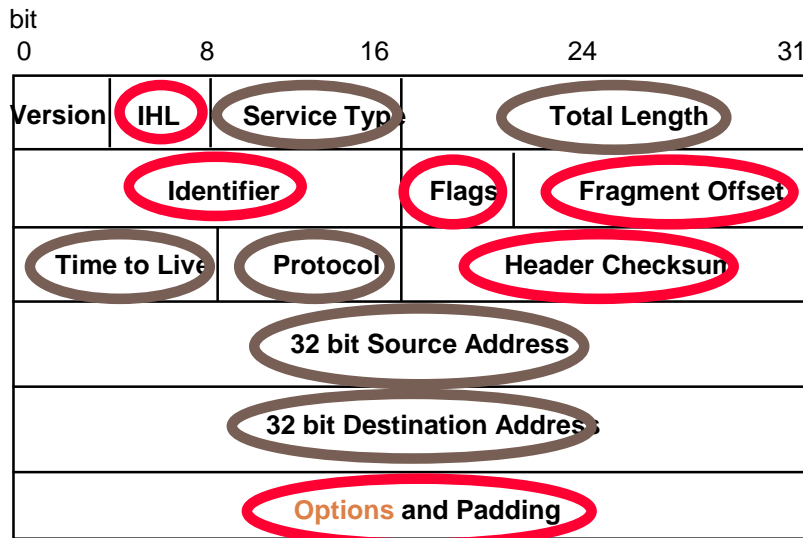
- ▣ 在multicast地址中增加了 范围 “scope”字段，允许将多播路由限定在正确的范围内
- ▣ 设置flog允许区分**永久性多播**地址和**临时性**多播地址

□ 可移动性

- ▣ 信宿选项报头、路由选项报头、自动配置、安全机制、以及anycast技术，将QoS同移动节点结合，从而强化对移动的支持

2.2 IPv6的报文结构

10

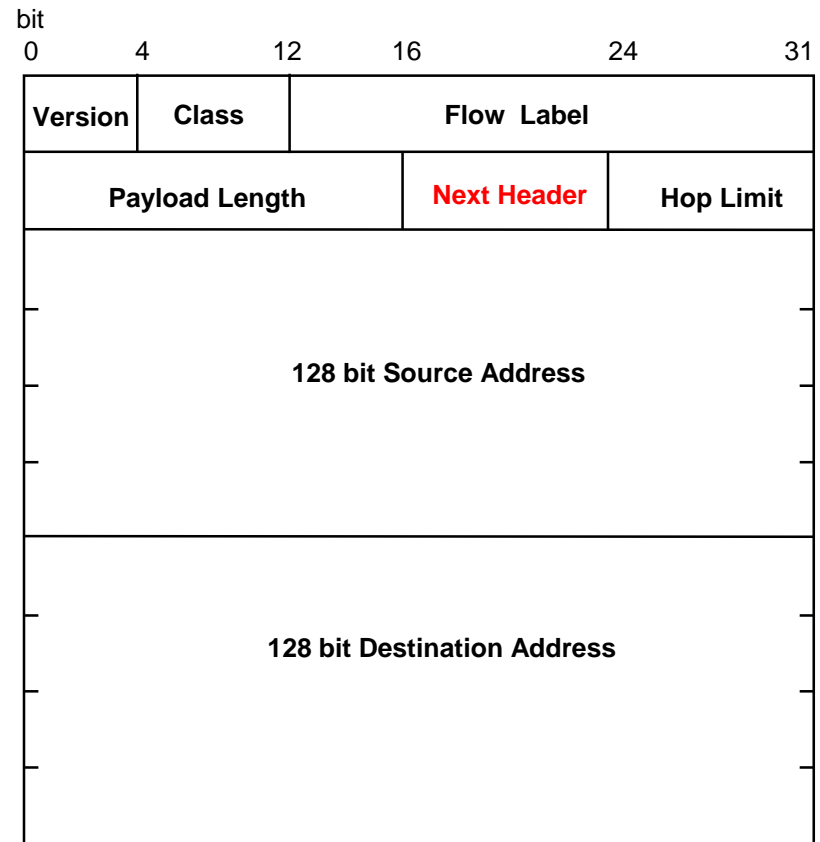


IPv4 Header

20 octets, 12 fields, including 3 flag bits, fixed max number of options

Changed

Removed



IPv6 Header

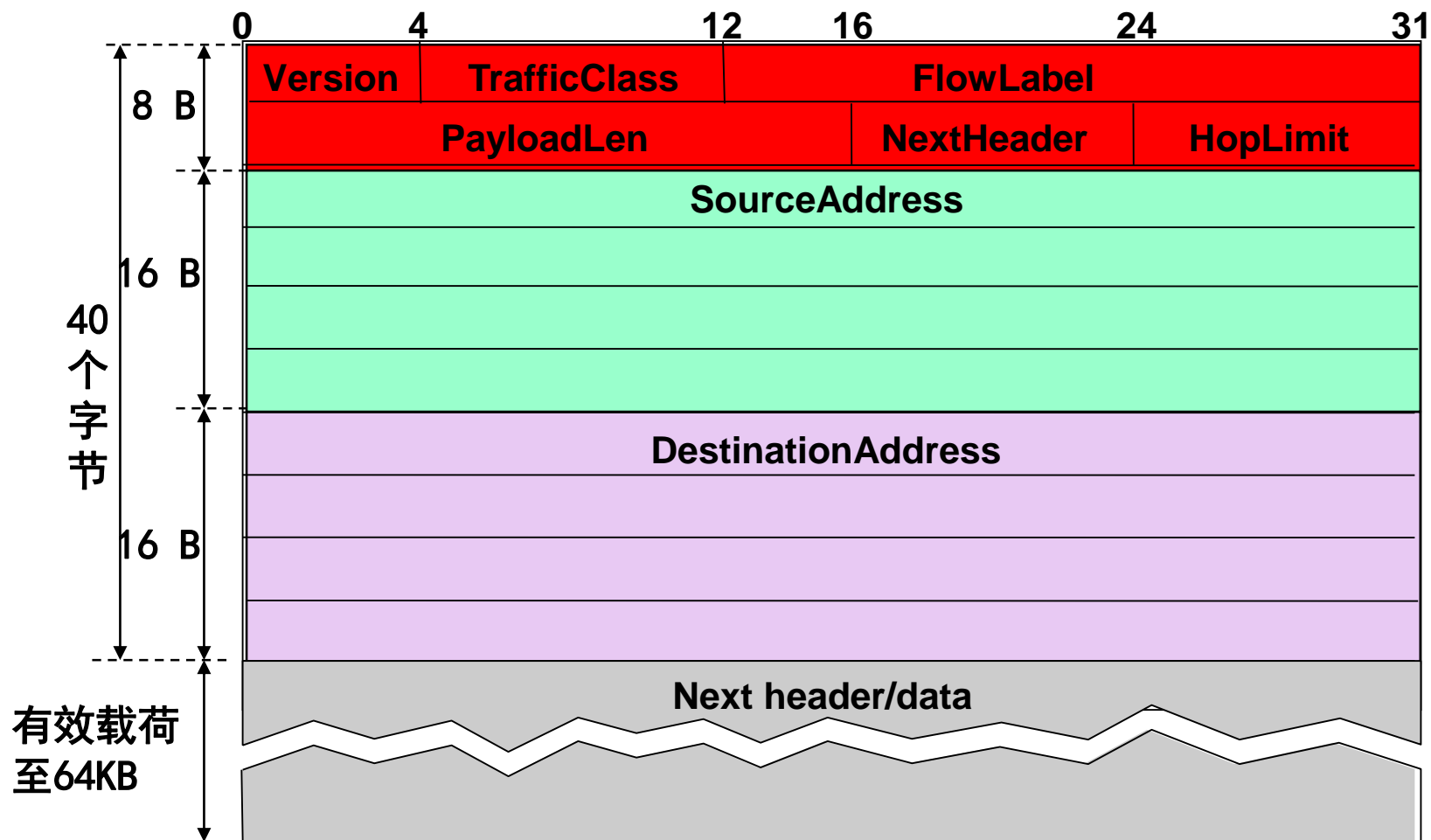
40 octets, 8 fields
+ Unlimited Chained Extension (options) Header

主要改变

- **对齐**（alignment）已经从32bit的整数倍改为64bit整数倍
($5 \times 64 \text{ bits} = 40 \times 8 \text{ B} = 320 \text{ bits}$)
- 取消了报头长度字段，基本报头长度固定40Bytes
- Total Length长度字段被Payload Length字段取代
- 源目地址字段增加到每个字段16个bytes
- 分片信息已经从基本报头的固定字段移到一个扩展报头中
- 生存时间TTL改为跳数极限hop limit字段
- 业务类型改为数据流标号flow label 字段
- 协议字段改为下一个报头字段，以指明下一个报头类型

IPv6的报文格式

12



- 4 bits—IP协议的版本号=6
- 8 bits—通信流类型
 - ▣ 相关应用层填充该类型值，默认值是全0
 - ▣ 某些节点可对某些比特按特定要求改变其产生、转发和接收，对不能理解的比特，节点忽略
 - ▣ 上层协议不能假定信源填充的值不变，宿端收到的值可能与源端不同
- 20 bits—数据流标号
 - ▣ 流：一条路径及其上的一些路由器，它保障一定的服务质量；或有相同源目地址的包集合，由信源给出标号
 - ▣ 支持新的机制：资源预定
 - ▣ 允许路由器将每个数据报同一个给定的资源分配相联系
 - ▣ 仍在实验中，两个例子：发收视频图象的两个应用程序之间可以建立一个数据流，其带宽和延时可得到保证；；ISP要求用户指明他所希望的QoS，然后指明一个数据流来限制某个计算机或应用程序所发送的流

◆ 8 bits - 跳数极限

- ♣ 经过路由器数量

◆ 128 bits - 源/目地址

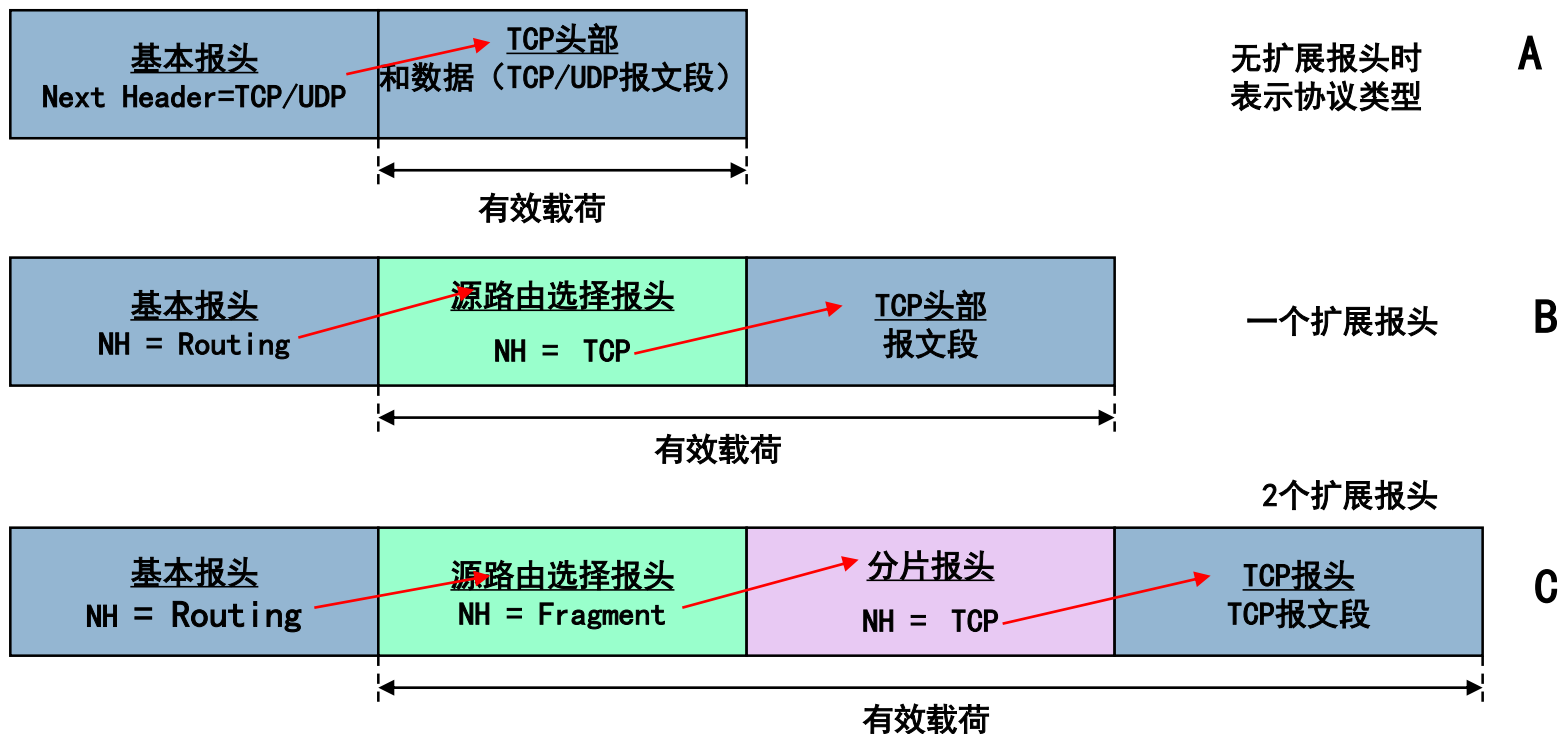
- ♣ 源/目地址分别都是 $16 \text{ Bytes} \times 8 = 128 \text{ Bits}$
- ♣ 如果扩展头中出现路由报头，宿地址可能不是最终接收站

◆ 16 bits - 有效载荷长度

- ♣ 因为基本报头已固定40bytes，故其长度字段不必要
- ♣ 用16bits表示有效载荷（即不报括基本报头的40bytes，但出现的任何扩展报头都计入有效载荷长度）
- ♣ 故一个IPv6数据报最多可容纳 $2^{16} = 64 \text{ Kbytes}$

□ 8 bits — 下一个报头(相当V4的协议字段或可选字段)

- 是IPv6的重大改进;
- 当无IPv6扩展首部时, 指明基本首部后面的数据应交付给高层哪一个协议 (如, 6→tcp;17→UDP)
- 当有扩展首部时, 指明标识后面第一个扩展首部的类型



报头的扩展

16

□ 扩展报头 (等价v4首部中的选项, 7种)

Extension Header	Type	Description
<i>Hop-by-Hop Options</i>	0	Options that need to be examined by all devices on the path.
<i>Destination Options</i> (before routing header)	60	Options that need to be examined only by the destination of the packet.
<i>Routing</i>	43	Methods to specify the route for a datagram (used with Mobile IPv6).
<i>Fragment</i>	44	Contains parameters for fragmentation of datagrams.
<i>Authentication Header (AH)</i>	51	Contains information used to verify the authenticity of most parts of the packet.
<i>Encapsulating Security Payload (ESP)</i>	50	Carries encrypted data for secure communication.
<i>Destination Options</i> (before upper-layer header)	60	Options that need to be examined only by the destination of the packet.
<i>Mobility</i> (currently without upper-layer header)	135	Parameters used with Mobile IPv6 .
<i>Host Identity Protocol</i>	139	Used for Host Identity Protocol version 2 (HIPv2). ^[10]
<i>Shim6 Protocol</i>	140	Used for Shim6 . ^[11]
Reserved	253	Used for experimentation and testing. ^{[12][4]}
Reserved	254	Used for experimentation and testing. ^{[12][4]}

报头的扩展

17

- 扩展报头
- V6把选项等效功能放在扩展首部，并将此留给两端主机处理，中间路由器除逐跳扩展（Hop-By-Hop Options）、源路由选择扩展（Routing），都不处理，提高了效率
- 逐跳扩展主要是巨型荷载选项（Jumbo Payload Option），报文长度可以达到4G
- IPv6中，仅数据报的发送者可以执行分片操作。这就是说，中间路由器或着主机不再需要处理分片报文，这样会提高分片报文处理效率
- IPv6分片是否会遭到拒绝服务攻击？当然也会，
 - Generation of IPv6 Atomic Fragments Considered Harmful（RFC 8021）

2.3 IPv6 地址

18

IPv6地址比IPv4复杂，承载了更多的功能

□ How big is the 128 bits address space ?

- 128 bits = 340 trillion (10^{12}) trillion trillion addresses
- $2^{128} = (2^{10})^{12.8} > (10^3)^{12.8} = 10^{38.4}$, 准确数目是
 $340,282,366,920,938,463,463,374,607,431,768,211,456 \geq 3.4 \times 10^{38.4}$
- **655,570,793,348,866,943,898,599** addresses per m^2 of the planet's surface = $2^{128} / 511,263,971,197,990 \text{ m}^2$ (地球表面面积) $\approx 6 \times 10^{23}$
- 而IPv4则只有每平方米**4个**地址

If an IPv6 Address Weighed 1 Gram...

还是假设一个IP地址重1克：

IPv6 address space = 56.7 billion个地球重量



Earth = 6.00e+24 kg*

* <http://www.howstuffworks.com/question30.htm>

$$\frac{2^{128}}{6.00e+27} = 56,713,727,820 > 567 \text{ 亿个地球!}$$

IPv6的地址结构

20

- A.地址的三种文本表示，以方便怎样阅读、输入和操作
 - ▣ 点分十进制
 - 104.230.140.100.255.255.255.255.0.17.128.150.10.255.255
 - ▣ 冒分16进制，共8个，相同字间距。上面地址为
 - 68E6:8C64:FFFF:FFFF:0:1180:96A:FFFF
 - ▣ 0压缩::表示，对连续长串0用::代替，一个地址中仅出现一次，如：
 - 2080:0:0:0:8:800:200C:417A→2080::8: ... ； unicast address
 - FF01:0:0:0:0:0:0:101→ FF01::101 ； multicast address
 - 0:0:0:0:0:0:0:1 → ::1 ； loopback address
 - 0:0:0:0:0:0:0:0 → :: ； undefined address
 - ▣ 混合表示，x: x: x: x: x: x: d.d.d.d, x :表示16进制(16 Bits), d.表示10进制(8 Bits)
 - 0:0:0:0:0:0:13.1.168.3 或 ::13.1.168.3
 - 0:0:0:0:0:FFFF:129.144.52.38或 :: FFFF:129.144.52.38

B.地址结构前缀的表示

21

□ CIDR形式

- IPv6地址/前缀长度，长度是10进制，表明地址最左端连续比特个数

- 正确表示12AB00000000CD3的60bits前缀是：

- 12AB:0000:0000:CD30:0000:0000:0000:0000/60

- 12AB::CD30:0:0:0:0/60

- 12AB:0:0:CD30::/60

不正确的表示为（没有准确体现前缀=60个bit后面为0）：

- 12AB:0:0:CD3/60 ； 可理解为0CD3

- 12AB::CD30/60 ； 可理解为12AB:0:0:0:0:0:0:CD30

□ HUST的IPV6地址= $2^{81} = 2^{80} \times 2$

- 2001:0250:4000::/48

- 2001:0DA8:3000::/48

IPv6的地址模式

22

- 地址分配到接口：
 - ▣ 这同v4一样，没有变化
 - ▣ 一个接口可有多多个地址
- 地址有范围之分
 - ▣ Link Local
 - ▣ Site Local（RFC 3879解释了取消站点局部地址）
 - ▣ Global
- 地址有寿命
 - ▣ 有效的
 - ▣ 永久的
- 地址结构
 - ▣ 前缀+接口ID

IPv6的寻址

23

- 地址类型：
 - ▣ Unicast: One to One(Global,Link local,Site local, Compatible)
 - ▣ Anycast:One to Nearest(Allocated from Unicast)
 - ▣ Multicast:One to Many
 - ▣ Reserved
- 单个接口可能分配有任何单播、任播和组播地址
- 广播由组播代替

IPv6的寻址

24

以太网适配器 以太网:

```
连接特定的 DNS 后缀 . . . . . :  
IPv6 地址 . . . . . : 2001:250:4000:4160:d474:a4e6:48cc:918b  
临时 IPv6 地址. . . . . : 2001:250:4000:4160:d027:f14a:756c:f036  
本地链接 IPv6 地址. . . . . : fe80::d474:a4e6:48cc:918b%6  
IPv4 地址 . . . . . : 202.114.23.2  
子网掩码 . . . . . : 255.255.255.0  
默认网关. . . . . : fe80::1614:4bff:fe7d:4cbd%6  
202.114.23.254
```

- 全局IPv6地址（别人可以访问）
- 临时IPv6地址（访问别人，基于微软RFC 3041“Privacy Extensions for Stateless Address Autoconfiguration in IPv6”）
- 本地链接IPv6（fe开头访问同一局域网内IP，%6表示接口）

IPv6的地址类型

25

- 地址分类：1998 RFC2460 对IPv6的地址类型分为三类
 - ▣ 单播—unicast：目的地址指明一个单一的计算机（single interface），可是主机或路由器，发送到unicast的包将选择一条最短的路径到达目的站
 - ▣ 任播—anycast（集群...）：目的地址是共享一个IP地址的计算机接口集合（a set of interfaces），典型的情况是在不同物理网络上的不同节点，发送到anycast地址的包将选择一条最近路径到达该集群（路由度量距离最近的节点）中一个
 - ▣ 组播-multicast：目的地址是一组计算机（a set of interfaces），典型情况是属于不同网络的不同节点，发送到一个multicast地址的报文将投递给组中的每个成员。IPv6中没有广播地址，其功能由组播取代
- 所有IPv6地址都是分配给interface而不是node的，所有接口都必须有至少一个link-local unicast，一个单接口可分配任何一种类型的多重地址（uni/any/multicast）或地址范围

地址类型的表示

- 一个IPv6地址的具体类型是由其前面的bits决定的
- 包含这些比特的变长字段称为格式前缀FP (Format Prefix)
- 其最初的分配表如下

分配	前缀		所占比例
	2进制	16进制	
保留	0000 0000	0::/8	1/256
未分配	0000 0001	100::/8	1/256
为NSAP分配保留	0000 001	200::/7	1/128
为IPX分配保留	0000 010	400::/7	1/128
未分配	0000 011	600::/7	1/128
未分配	0000 1	800::/5	1/32
未分配	0001	1000::/4	1/16
可聚类全局 unicast地址	001	2000::/3	1/8
未分配	010	4000::/3	1/8
未分配	011	6000::/3	1/8
未分配	100	8000::/3	1/8
未分配	101	A000::/3	1/8
未分配	110	C000::/3	1/8
未分配	1110	E000::/4	1/16
未分配	1111 0	F000::/5	1/32
未分配	1111 10	F800::/6	1/64
未分配	1111 110	FC00::/7	1/128
未分配	1111 1110 0	FE00::/19	1/512
Link-local Unicast 地址	1111 1110 10	FE80::/10	1/1024
Site-local Unicast地址	1111 1110 11	FEC0::/10	1/1024
Multicast地址	1111 1111	FF00::/8	1/256

地址类型：保留地址

28

- 不要把保留地址和未分配地址混淆,保留地址不等于未分配地址
- 保留地址有3种, 共占 $2^{128-8}/2^{128} = \text{占} 1/256$, 由前缀0000 0000表示,
 - 全零地址: 没有规定的地址
 - Loopback地址: 回送地址
 - 嵌入到了IPV4地址的IPV6地址
- 其它保留地址
 - 0000 001:为NSAP (Network Service Access Point) 保留,占空间1/128
 - 0000 010:为Novell的IPX保留, 占空间1/128

地址类型：保留地址

29

- 未规定的地址：全零地址， **0:0:0:0:0:0:0:0**；不分配给任何节点，表示一个缺失地址。应用例之一是，还未分配IP地址的主机初始化中，要发送IPV6包时，用全零地址作为自己的**暂时源地址**。它不能作为信宿地址
- Loopback—回环地址， **0:0:0:0:0:0:0:1**；可被任何节点用于向自身发送IPV6包，它不能分配给任何物理接口，不能作为任何包的信源地址，以此为信宿的包永远不能发出该节点，永远不能被路由器转发
- 包含IPv4地址的IPv6地址：RFC 1993定义了IPv6 over IPv4 tunnel 机制。应用这种技术的IPv6节点被分配给一种特殊的IPv6 unicast 地址，其最低32位是IPv4地址

地址类型：保留地址

30

- 这种支持IPV4的地址又分2种
 - ▣ 既支持V4，也支持V6：如图1，是与V4兼容的V6地址，
 - ▣ 仅支持V4，不支持V6：如图2，对不支持V6的R/H，应屏蔽最低32位以上的部分，V6还规定33--48位全部为F，高位全0时其为V4地址

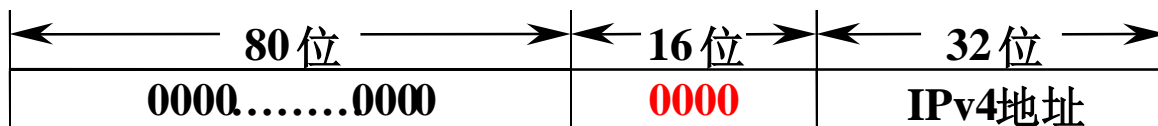


图. IPV4 兼容的IPV6地址

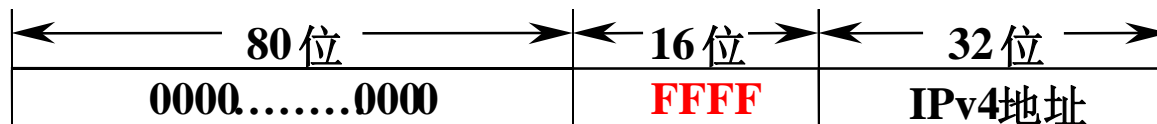


图2. IPV4 映射的IPV6地址

地址类型：Unicast和Multicast地址

31

- 前缀为001—111的地址，除Multicast地址（1111 1111）外，都必须包含64 Bits的EUI-64G格式的接口标识符
- 表上的分配仅仅使用了地址空间的15%，剩余的85%留待今后使用
- 这种分配方式支持聚类地址、本地使用地址和Multicast地址的直接分配，剩余空间既可支持现有使用扩展（如附加可聚类地址），也可用于新的领域
- Unicast 和 multicast 的地址靠最高8位来区分，FF表示是multicast地址，其它是Unicast地址
- Anycast 地址从Unicast中分出，格式上同Unicast没有区分

Unicast 地址(RFC2374)

32

- Unicast地址严格聚类，具有连续前缀
- 目前的Unicast地址分配的几种形式是
 - ▣ 全局可聚类地址
 - ▣ NASP地址
 - ▣ IPX层次地址
 - ▣ Link-Local地址
 - ▣ Site-Local地址
 - ▣ IPv4兼容主机地址
 - ▣ 将来可定义的其它地址类型

IPv6 Unicast 地址的结构

33

- IPv6 Unicast 地址结构的解释依据节点所扮演的角色（主机或者是路由器），最简单情况下可认为是没有任何内在的结构，如图1
- 稍复杂情况，包括与其相连子网的 n 位前缀，如图2

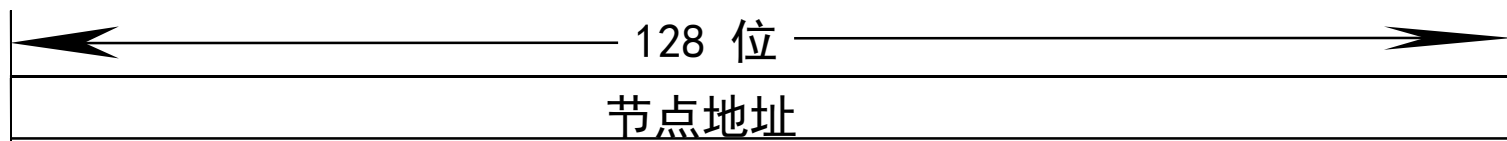


图1. 没有内在结构的Unicast地址

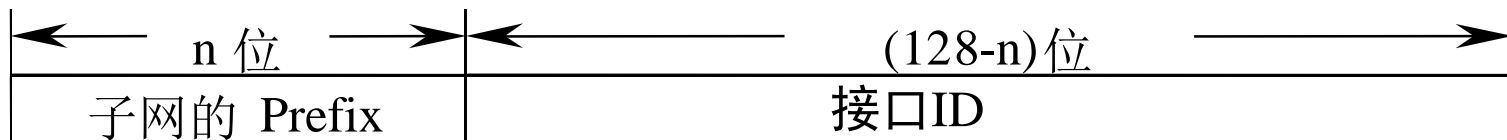
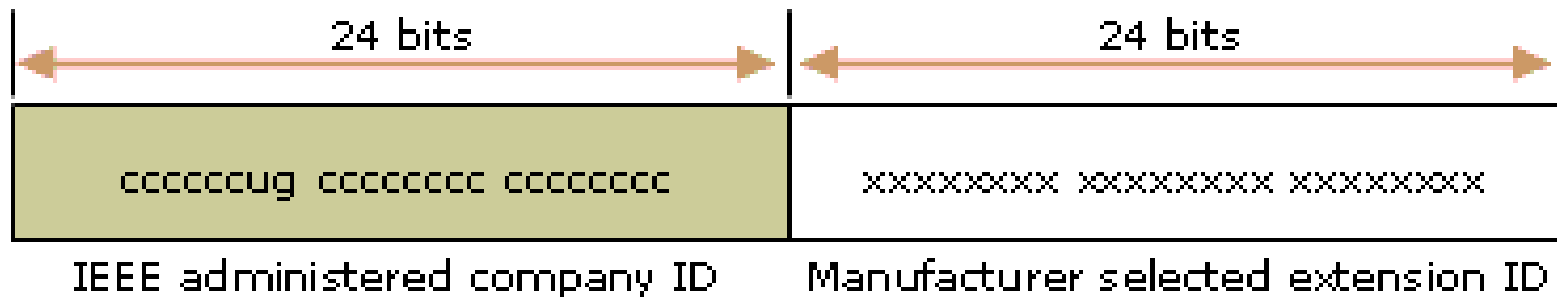


图2. 有子网前缀的Unicast地址

接口标识符

34

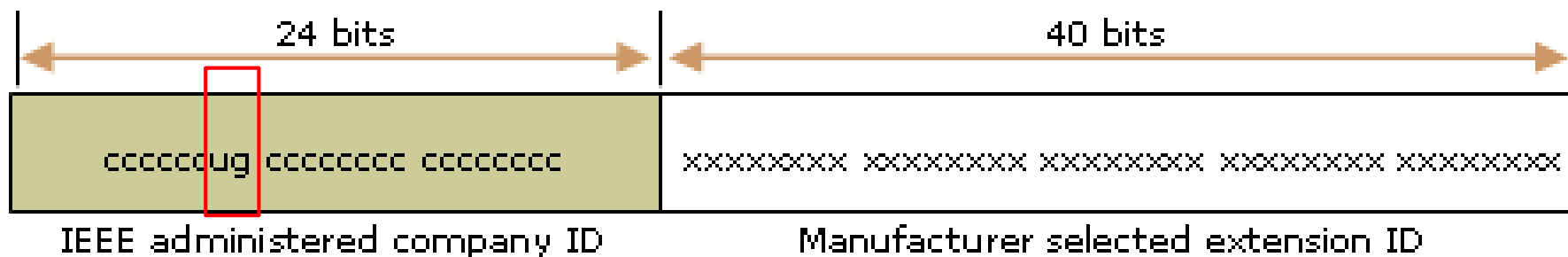
- IEEE EUI-Extended Unique Identifier: 扩展唯一标识符
 - ▣ RFC 2373 规定所有unicast地址必须有64比特的EUI-64的接口ID
 - ▣ EUI来自IEEE 802 Address, 包括24位制造商地址, 24位扩展地址, 称为硬件、物理或MAC地址



IEEE 802 Address

□ EUI-64的接口ID

- The IEEE EUI-64 address 表示网络物理接口寻址的新标准
- 公司ID仍然 24-bits， 扩展ID为40bits， 给网卡商更大的地址空间
- EUI-64地址中的 U/L 和 I/G bits与IEEE 802 address的表示意义相同
- Universal/Local (U/L) 是第7位， 0：表示全局ID；1：表示局部ID；
- Individual/Group (I/G) 是第8位， 0：表示单播地址；1：组播地址
- C是制造商的标识符

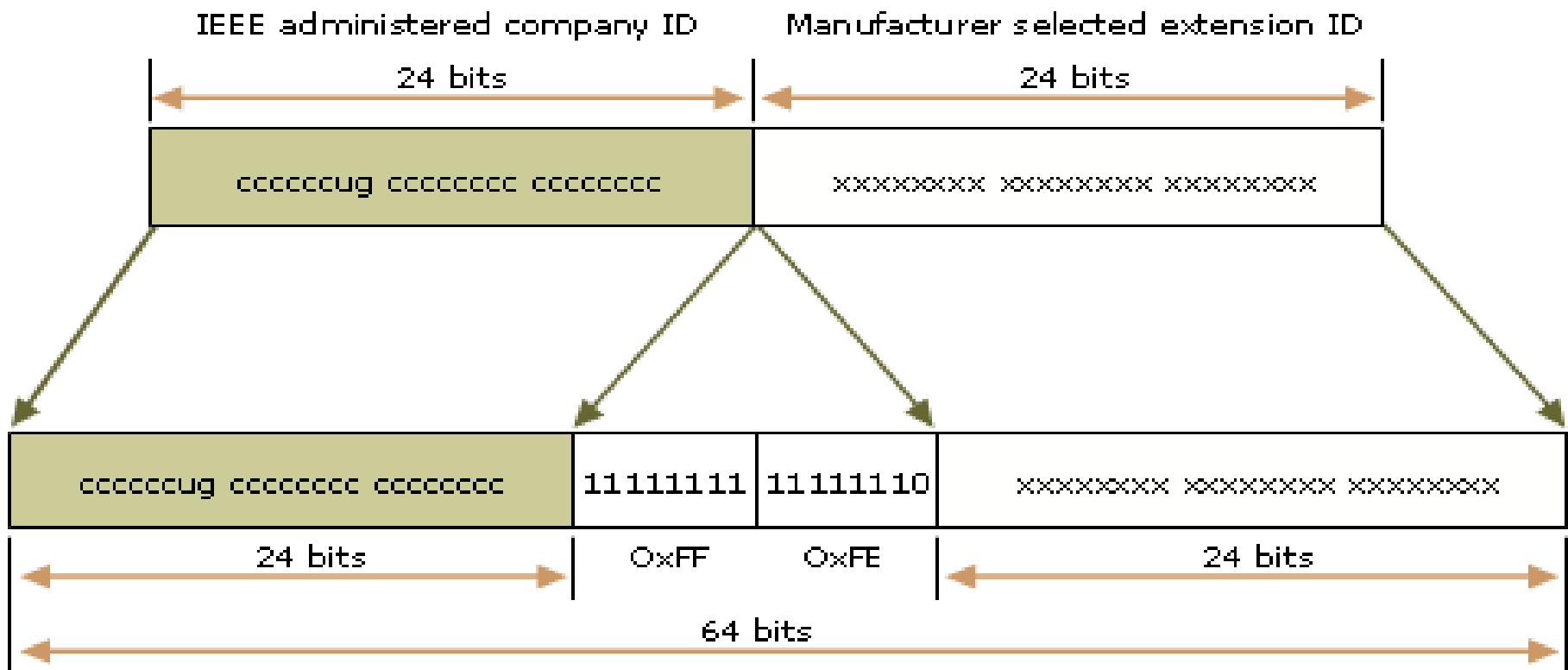


IEEE EUI-64 Address

Mapping IEEE 802 to EUI-64 addresses

36

- To create an EUI-64 address from an IEEE 802 address, the 16 bits of **11111111 11111110** (0x**FFFE**) are inserted into the IEEE 802 address between the company ID and the extension ID.



Mapping EUI-64 addresses to IPv6 ID

37

- To obtain the 64-bit interface identifier for IPv6 unicast addresses, the U/L bit in the EUI-64 address is complemented (取补) if it is a 1, it is set to 0; and if it is a 0, it is set to 1).

EUI-64 address

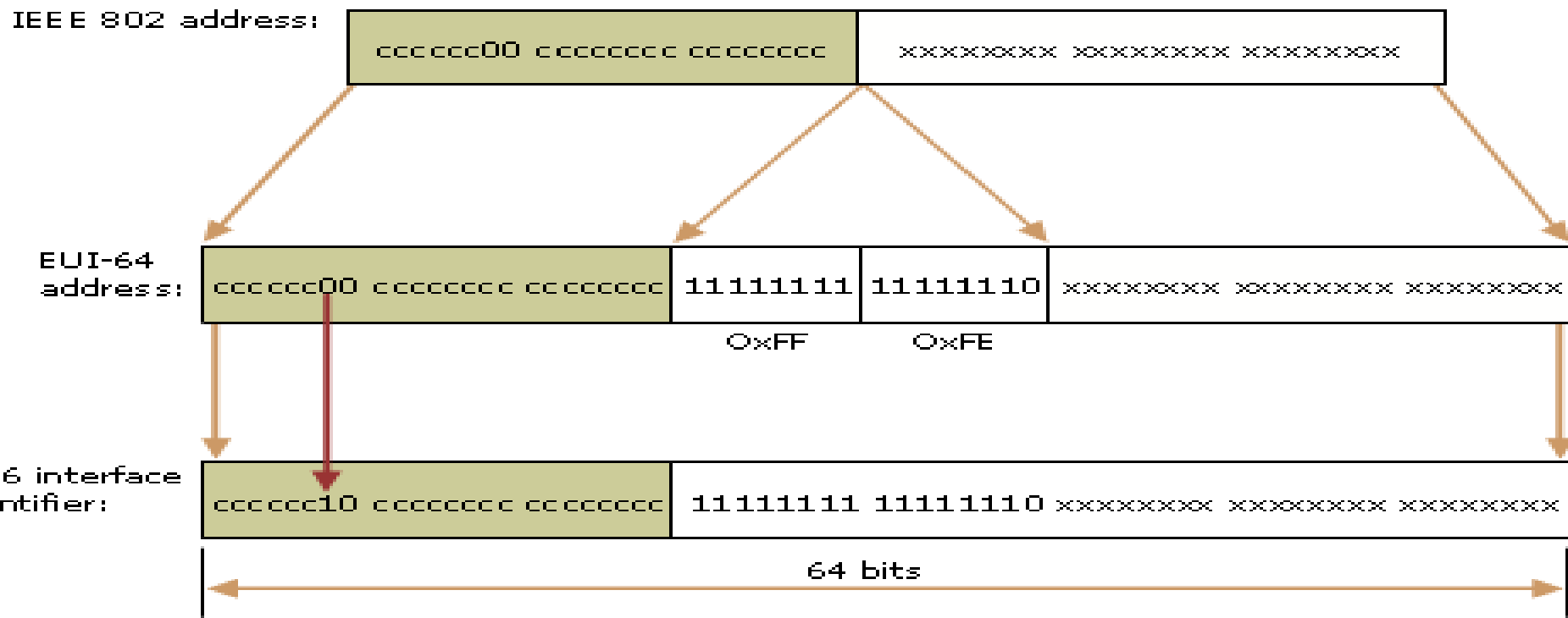


IPv6 interface identifier

Mapping IEEE 802 to IPv6 ID

38

- To obtain an IPv6 interface identifier from an IEEE 802 address, you must first map the IEEE 802 address to an EUI-64 address, and then complement the U/L bit.
- The following illustration shows the conversion process for a universally administered, unicast IEEE 802 address.



Mapping IEEE 802 to IPv6 ID

39

但是Windows 10产生IPv6地址的时候并没有完全遵循这个规范，如果用PowerShell命令查看IPv6网络配置：

- Get-NetIPv6Protocol如果发现RandomizeIdentifiers选项enable，那么就不会使用EUI-64作为接口标识符，而是在MAC address基础上随机产生：

```
MulticastForwarding : Disabled
GroupForwardedFragments : Disabled
RandomizeIdentifiers : Enabled
AddressMaskReply : Disabled
UseTemporaryAddresses : Enabled
```

- Set-NetIPv6Protocol -RandomizeIdentifiers Disabled，才会遵循原来的规范

Name	InterfaceDescription	ifIndex	Status	MacAddress
以太网	Realtek PCIe GBE Family Controller	18	Up	70-4D-7B-61-E2-37

ifIndex	IPAddress	PrefixLength	PrefixOrigin	SuffixOrigin
3	fe80::250:56ff:fec0:8%3	64	WellKnown	Link
9	fe80::250:56ff:fec0:1%9	64	WellKnown	Link
1	::1	128	WellKnown	WellKnown
18	fe80::724d:7bff:fe61:e237%18	64	WellKnown	Link

Mapping IEEE 802 to IPv6 ID

40

- 现在的华中科技大学网络中IPv6地址从路由器提供前缀方式变为了DHCPv6方式，所以不能再看到以太网上两个全局IPv6地址了

```
PS C:\WINDOWS\system32> Get-NetIPAddress | format-table
```

ifIndex	IPAddress	PrefixLength	PrefixOrigin	SuffixOrigin
3	fe80::250:56ff:fec0:8%3	64	WellKnown	Link
9	fe80::250:56ff:fec0:1%9	64	WellKnown	Link
1	::1	128	WellKnown	WellKnown
18	fe80::724d:7bff:fe61:e237%18	64	WellKnown	Link
18	2001:250:4000:4160:e757:bbed:7923:759c	128	Dhcp	Dhcp
3	192.168.157.1	24	Dhcp	Dhcp
9	192.168.88.1	24	Dhcp	Dhcp
1	127.0.0.1	8	WellKnown	WellKnown
18	202.114.23.188	24	Dhcp	Dhcp

可聚类全局Unicast地址

41

- FP为001（2000到3FFF开头）都是全局可聚类Unicast地址，由RFC 2374给出
- 基本假设：路由系统基于“最长前缀匹配”算法来选择转发路径
- 特点：
 - 与IPv4的CIDR不同，IPv6强制规定，地址中的前64个bit才能作为网络地址。
 - 在理想情况下，一个核心主干网路由器只须维护不超过8192个表项（TLA顶级聚合为13个bit）
 - IPv6改变了地址的分配方式，从用户拥有变成了ISP拥有。全球网络号由因特网地址分配机构（IANA）分配给ISP，用户的全球网络地址是ISP地址空间的子集。每当用户改变ISP时，全球网络地址必须更新为新ISP提供的地址。这样ISP能有效地控制路由信息，避免路由爆炸现象的出现。

可聚类全局Unicast地址

42

□ 共分3级，6个部分

- 001:全球可聚合Unicast地址
- TLA ID (Top Level Aggregator) : 顶级聚合标识符, 分配给大型ISP, 从IANA直接获得。
- RES:留做将来使用—Reserved for future use
- NLA ID (Next Level Aggregator) :次级聚合标识符, 中型ISP从TLA获取。
- SLA ID (Site Level Aggregator) : 站点级聚合标识符, 小型ISP从NLA获得
- 接口ID: 接口标识符Interface Identifier

3	13	8	24	16	64 bit
FP	TLA ID	RES	NLA ID	SLA ID	Interface ID
公共拓扑 48 bits				站点拓扑	接口 ID

华科地址: 2001:0250:4000::/48

顶级聚类标识符 - TLA ID

43

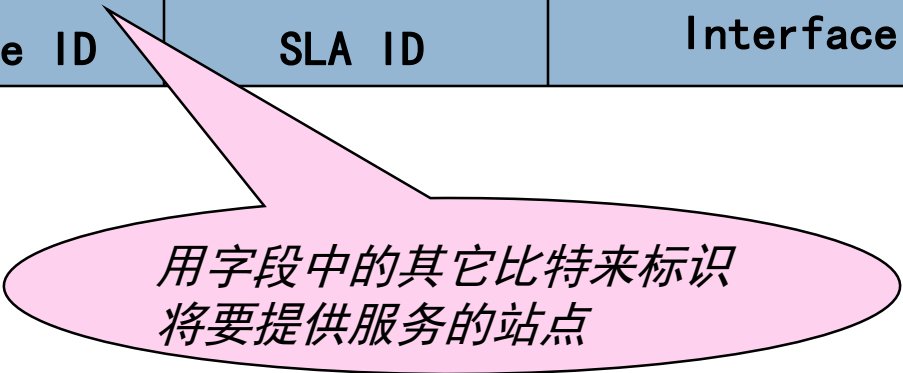
- 位于路由层次的顶层
- 共支持 $2^{13} = 8192$ 个 TLA ID，扩展方式：
 - ▣ 将 TLA 字段扩展到保留字段中去
- 保留字段 — Reserved 2^8
 - ▣ 现在必须设置成 0
 - ▣ 可为 TLA NLA 提供增长的可能

次级聚类标识符 - NLA ID

44

- 为已得到TLA ID的机构创建的地址层次
- 每1个TLA ID空间允许该机构创建大致相当于目前IPv4 internet所支持网络总数
- 下面是一个NLA ID的可能结构

<i>n bit</i>	<i>(24-n) bit</i>	<i>16 bit</i>	<i>64 bit</i>
NLA 1	Site ID	SLA ID	Interface ID

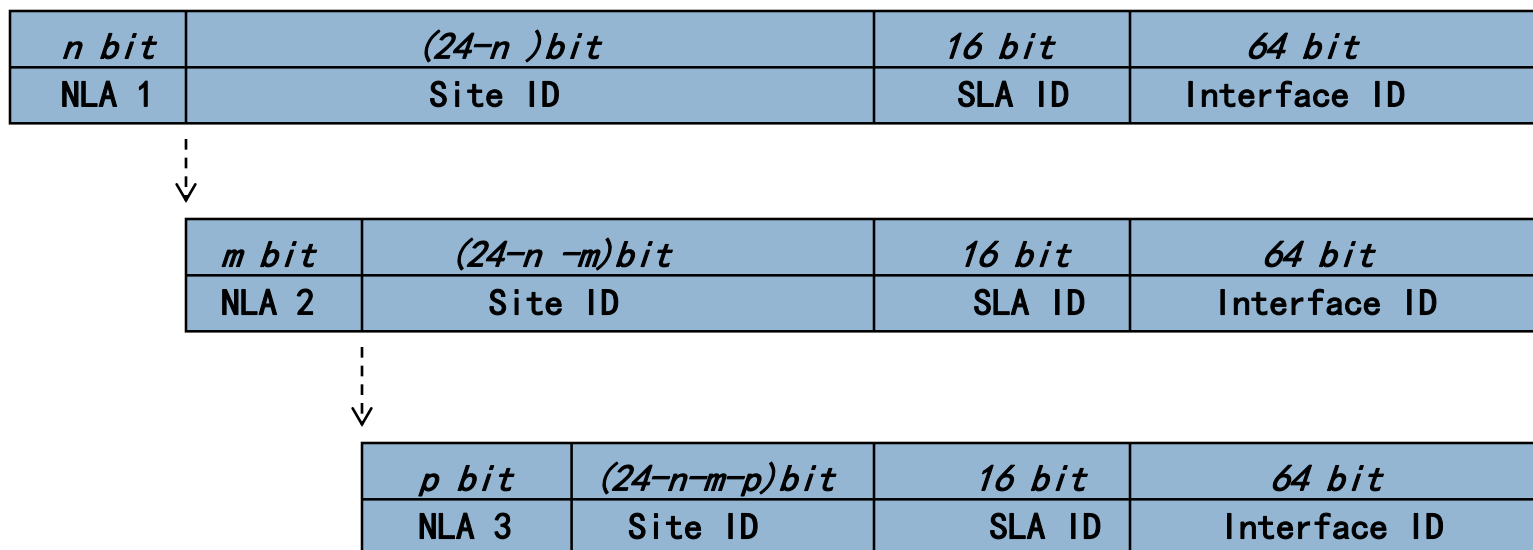


用字段中的其它比特来标识
将要提供服务的站点

多层NLA结构

45

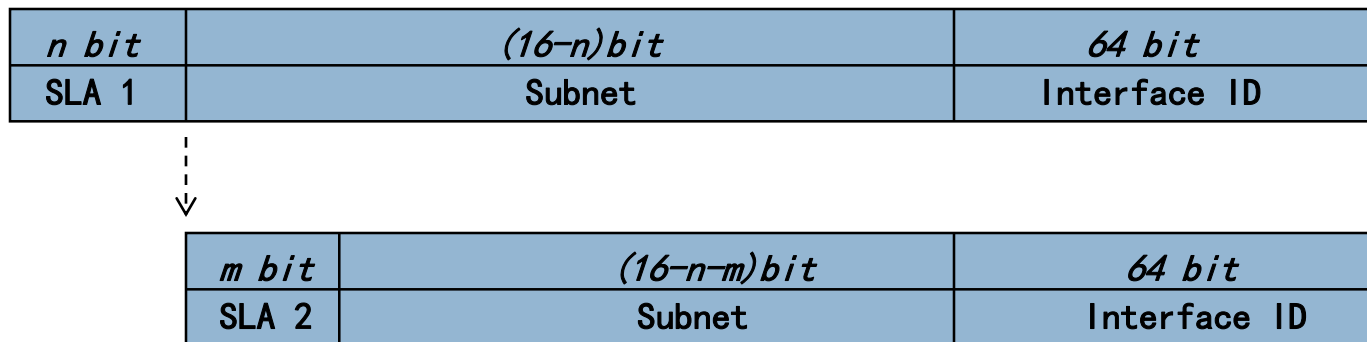
- 得到TLA ID的机构可以在自身的NLA ID的空间中支持多级NLA ID
- 上前一级NLA ID机构负责管理下一级的NLA ID空间的bit
- NLA ID的分级是路由聚类效率和灵活性的折中，层次越多，则聚类可能越大，同时减少了路由表的规模；而平坦（flat）的NLA ID分配简单且灵活，同时也增加了路由表的规模



地点聚类标识符 - SLA ID

46

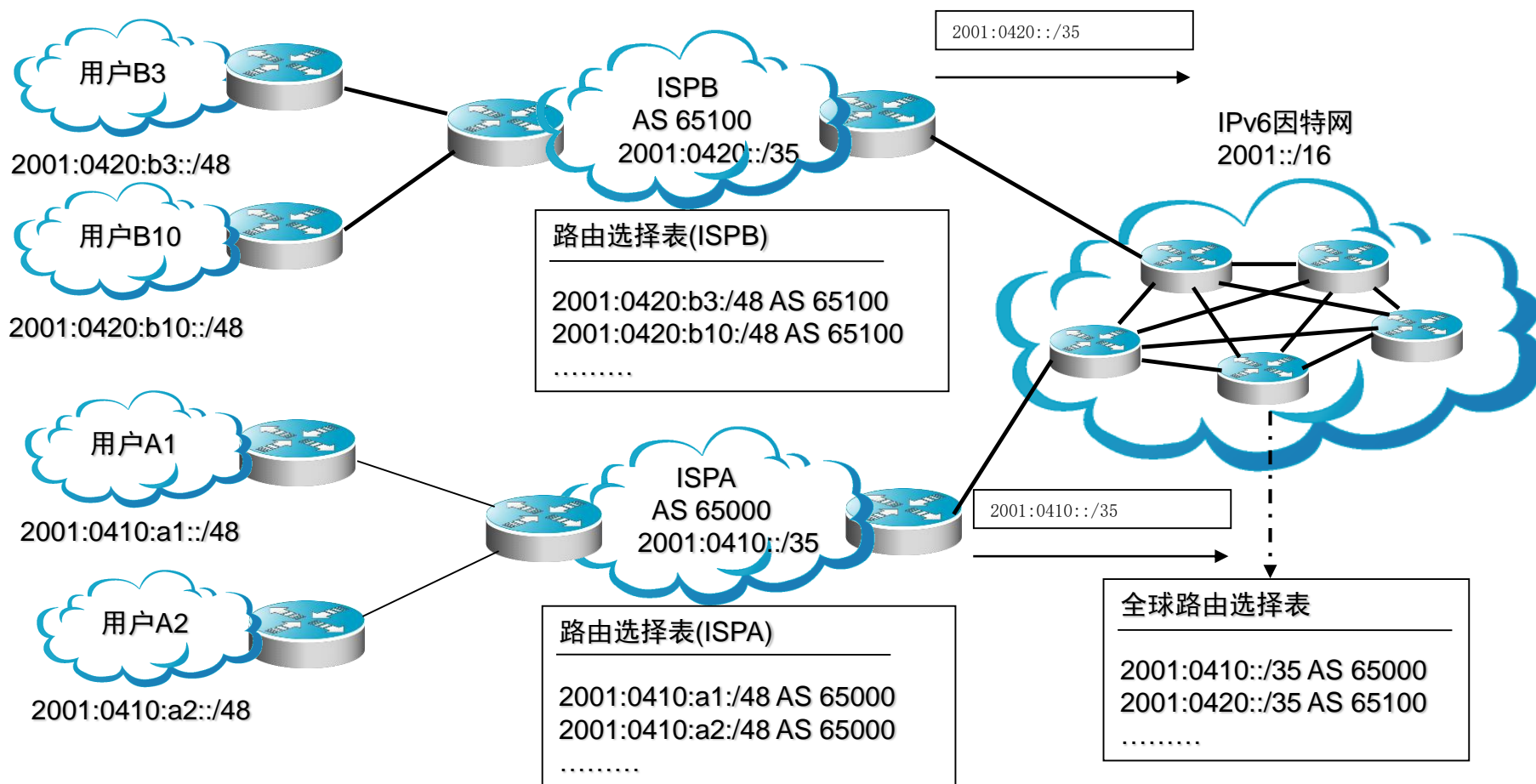
- SLA ID字段由单独机构用决定来创建其内部的地址层次并标识子网（这类似v4的子网划分），SLA ID有 $2^{16} = 65535$ 个子网
- 可选择将SLA ID设置成flat路由（SLA内无任何逻辑关系，但增加路由表规模）；也可在SLA ID字段中创建更多的层次和级别（将减少路由表规模）



可能的SLA结构

有效的、分级的寻址和路由结构

下面以一个实例来说明IPv6路由的聚合：

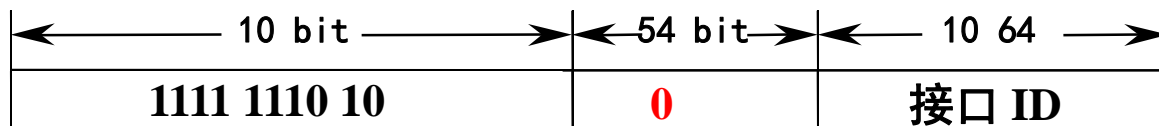


提供商聚合客户的前缀并公告他们的前缀到IPv6因特网

本地使用的2类Unicast地址

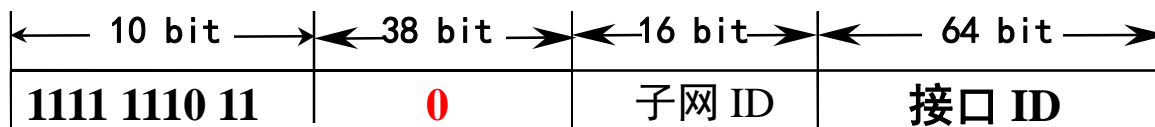
48

- Link-Local address: 前缀PF = 1111 1110 10 (**FE80::/10**)
- 用于本地链路上的地址分配, 例如
 - ▣ 自动地址配置—auto-address configuration
 - ▣ 邻站发现—neighbor discovery
 - ▣ 没有路由器时
- 路由器不能转发任何以Link-Local为源目的的包到其它链路



Link local的地址格式

- Site-Local address: 前缀PF = 1111 1110 11
- 用于一个单独的站点，站点内不需要全局前缀的地址分配
- 路由器不能转发任何以Site-Local为源目的的包到其它链路



Site local的地址格式

2.4 IPv6 邻居发现协议

- 基于ICMPv6报文实现其功能（RFC 4861）
- 路由器请求（Router Solicitation）
- 路由器通告（Router Advertisement）
- 邻居请求（Neighbor Solicitation）
- 邻居通告（Neighbor Advertisement）

邻居发现协议作用

- ❑ **Router discovery:** hosts can locate routers residing on attached links.
- ❑ **Prefix discovery:** hosts can discover address prefixes that are on-link for attached links.
- ❑ **Parameter discovery:** hosts can find link parameters (e.g., MTU).
- ❑ **Address auto-configuration:** optional stateless configuration of addresses of network interfaces.
- ❑ **Address resolution:** mapping between IP addresses and link-layer addresses.
- ❑ **Next-hop determination:** hosts can find next-hop routers for a destination.
- ❑ **Neighbor unreachability detection (NUD):** determine that a neighbor is no longer reachable on the link.
- ❑ **Duplicate address detection (DAD):** nodes can check whether an address is already in use.

Address Autoconfiguration (1)

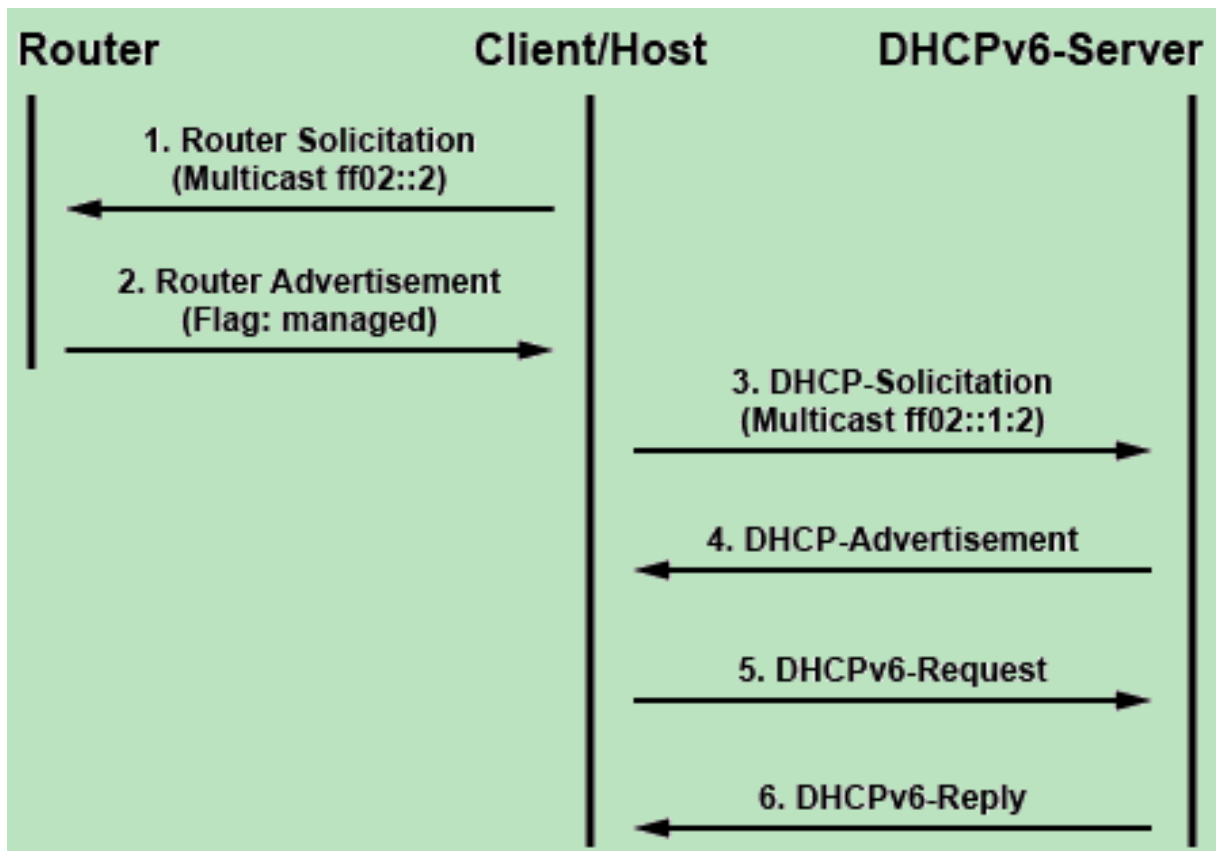
- 允许即插即用
- BOOTP and DHCP are used in IPv4
- DHCPng will be used with IPv6
- 两种方法: Stateless and Stateful
- Stateless通过邻居发现来完成

Address Autoconfiguration (2)

- Stateful有状态的地址自动化配置:
 - ▣ Routers ask the new host to go DHCP server (by setting managed configuration bit)
 - ▣ The new host multicasts to "All DHCP servers"
 - ▣ There's also a reserved, link-scoped multicast address (**FF02::1:2**) and New UDP port numbers clients listen for DHCP messages on UDP port 546. Servers listen for DHCP messages on UDP port 547.
 - ▣ DHCP server assigns the new host an address

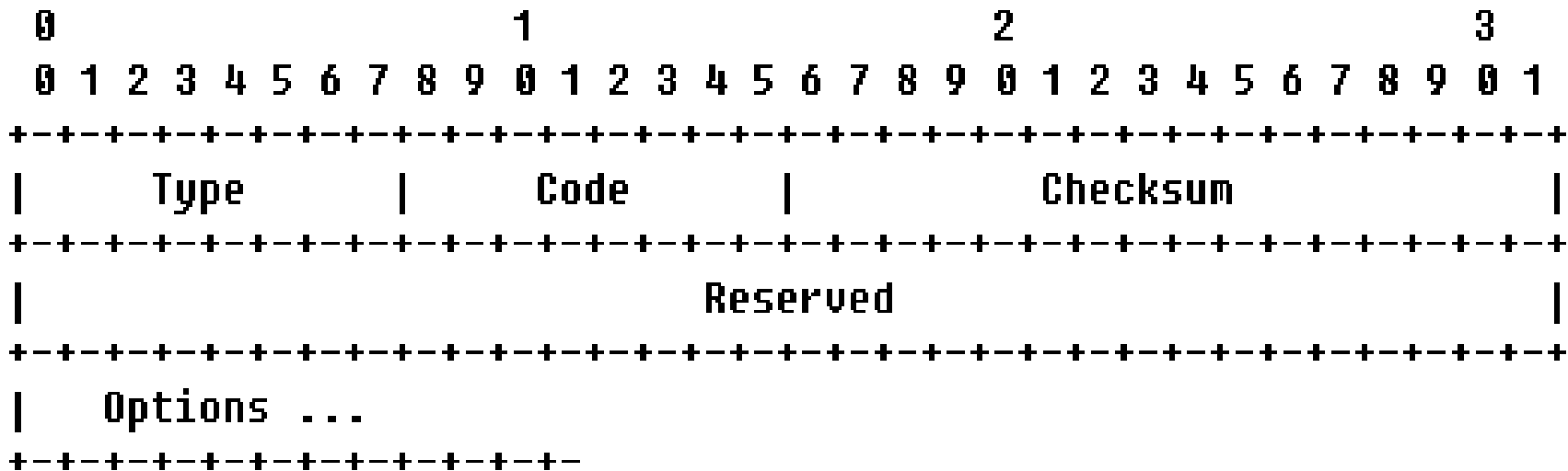
Address Autoconfiguration (2)

- Stateful有状态的地址自动化配置:



Router Solicitation 报文

- RS是主机发送的报文，触发路由器迅速产生路由器通告。
- 回应报文为RA报文
- 报文结构（**ICMP**）如下：



Router Solicitation报文结构

□ IP 部分

- ▣ 源地址：接口（link-local）的地址或者unspecified（全0）。
- ▣ 目的地址：全部路由器组播地址FF02::02
- ▣ 跳数：255（If a node receives an RS with Hop-Limit less than 255, the packet is deemed invalid. Hence, the Hop-Limit of 255 ensures the packet has not traversed through a router）

□ ICMP部分

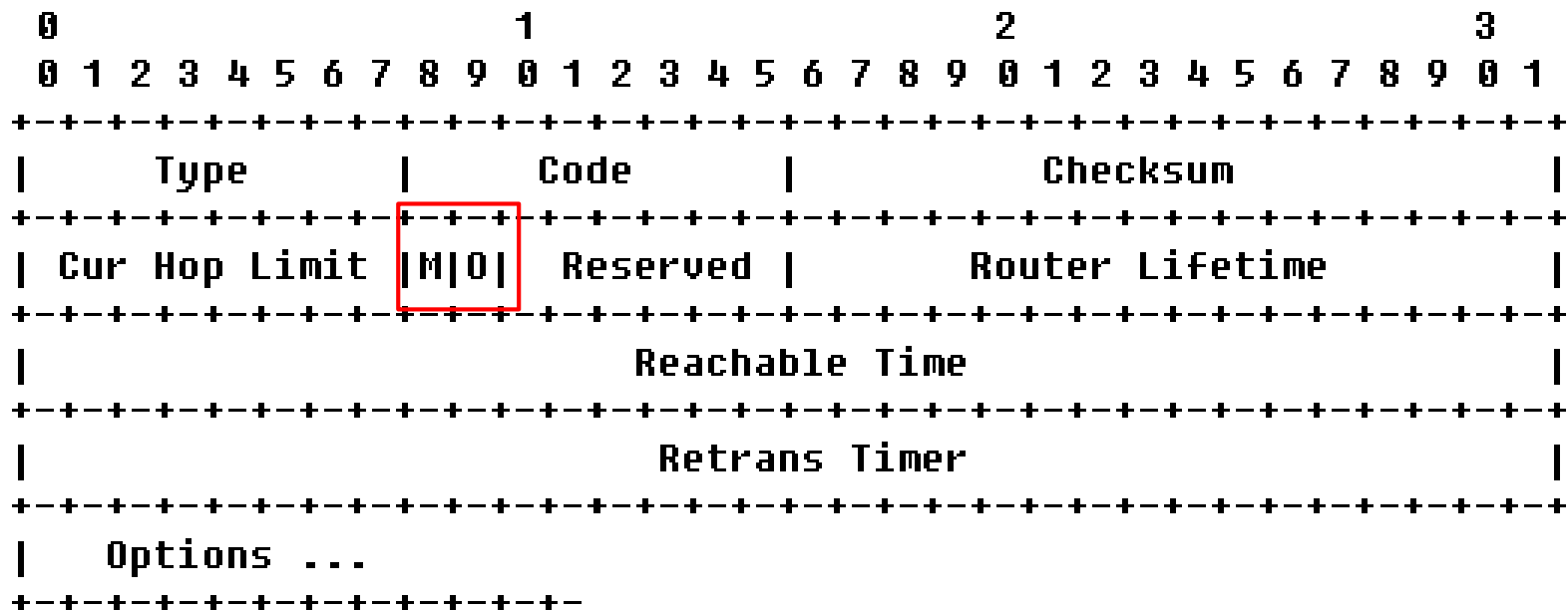
- ▣ Type=133
- ▣ Code=0
- ▣ 选项部分包含了发送者的link-layer地址

Router Solicitation报文结构

131	49.785307	fe80::200:86ff:fe05:80da	ff02::2	ICMPv6
> Frame 131: 62 bytes on wire (496 bits), 62 bytes captured (496 bits)				
> Ethernet II, Src: GatewayC_05:80:da (00:00:86:05:80:da), Dst: IPv6mcast_02 (33:33:00:00:00:02)				
> Internet Protocol Version 6, Src: fe80::200:86ff:fe05:80da, Dst: ff02::2				
✓ Internet Control Message Protocol v6				
Type: Router Solicitation (133)				
Code: 0				
Checksum: 0x7557 [correct]				
[Checksum Status: Good]				
Reserved: 00000000				
0000	33 33 00 00 00 02 00 00	86 05 80 da 86 dd 60 00	33.....`.	
0010	00 00 00 08 3a ff fe 80	00 00 00 00 00 00 02 00:... ..	
0020	86 ff fe 05 80 da ff 02	00 00 00 00 00 00 00 00	
0030	00 00 00 00 00 02 85 00	75 57 00 00 00 00 00 00 uW....	

Router Advertisement报文

- 由路由器发出
- 路由器周期性地发送路由器通告消息，或者对路由器请求作出响应
- 报文（**ICMP**）结构如下：



Router Advertisement报文结构

- IP部分
 - ▣ 源地址：发送者Link-local地址
 - ▣ 目的地址：全部节点组播地址FF02::1或发送RS的主机单播地址
 - ▣ 跳数：255
- ICMP部分
 - ▣ Type=134
 - ▣ Code=0
 - ▣ Cur hop limit = 当一台主机发送IPv6报文，其报文头部的hop limit应该设置的默认值是多少
 - ▣ 选项部分包含了发送者的link-layer地址
 - ▣ 选项部分包含了MTU、地址前缀

Router Advertisement报文结构 (续)

□ ICMP部分

- ▣ M=0, 表示使用stateless 地址自动配置
- ▣ M=1, 表示使用stateful 地址自动配置(DHCPv6)
- ▣ O bit, other information, such as DNS
- ▣ Router Lifetime, 表示存在于主机default router缓存中的时间
- ▣ Reachable Time, 表示存在于主机邻居缓存中的时间
- ▣ Retrans Timer, 表示进行邻居检测时的重新发送间隔

Router Advertisement报文结构 (续)

132	50.080559	fe80::260:97ff:fe07:69ea	ff02::1	ICMPv6
<p>> Frame 132: 118 bytes on wire (944 bits), 118 bytes captured (944 bits)</p> <p>> Ethernet II, Src: 3com_07:69:ea (00:60:97:07:69:ea), Dst: IPv6mcast_01 (33:33:00:00:00:01)</p> <p>> Internet Protocol Version 6, Src: fe80::260:97ff:fe07:69ea, Dst: ff02::1</p> <p>▼ Internet Control Message Protocol v6</p> <p>Type: Router Advertisement (134)</p> <p>Code: 0</p> <p>Checksum: 0x4625 [correct]</p> <p>[Checksum Status: Good]</p> <p>Cur hop limit: 64</p> <p>> Flags: 0x00, Prf (Default Router Preference): Medium</p> <p>Router lifetime (s): 1800</p> <p>Reachable time (ms): 30000</p> <p>Retrans timer (ms): 1000</p> <p>> ICMPv6 Option (Source link-layer address : 00:60:97:07:69:ea)</p> <p>> ICMPv6 Option (MTU : 1500)</p> <p>> ICMPv6 Option (Prefix information : 3ffe:507:0:1::/64)</p>				
0000	33 33 00 00 00 01 00 60	97 07 69 ea 86 dd 60 00	33.....`	..i...`.
0010	00 00 00 40 3a ff fe 80	00 00 00 00 00 00 02 60	...@:...`
0020	97 ff fe 07 69 ea ff 02	00 00 00 00 00 00 00 00i...
0030	00 00 00 00 00 01 86 00	46 25 40 00 07 08 00 00F%@....	
0040	75 30 00 00 03 e8 01 01	00 60 97 07 69 ea 05 01	u0.....`	..i...
0050	00 00 00 00 05 dc 03 04	40 c0 00 36 ee 80 00 36@..6...6	
0060	ee 80 00 00 00 00 3f fe	05 07 00 00 00 01 00 00?.	
0070	00 00 00 00 00 00		

Router Advertisement报文结构 (续)

□ IPv6 Neighbor Discovery Option

- 大概有40多种
- Source link-layer address表示路由器的链路层接口地址，可以用来做负载均衡
- MTU是网络内部最大传输单元

```
✓ ICMPv6 Option (Source link-layer address : 00:60:97:07:69:ea)
  Type: Source link-layer address (1)
  Length: 1 (8 bytes)
  Link-layer address: 3com_07:69:ea (00:60:97:07:69:ea)
✓ ICMPv6 Option (MTU : 1500)
  Type: MTU (5)
  Length: 1 (8 bytes)
  Reserved
  MTU: 1500
```

Router Advertisement报文结构 (续)

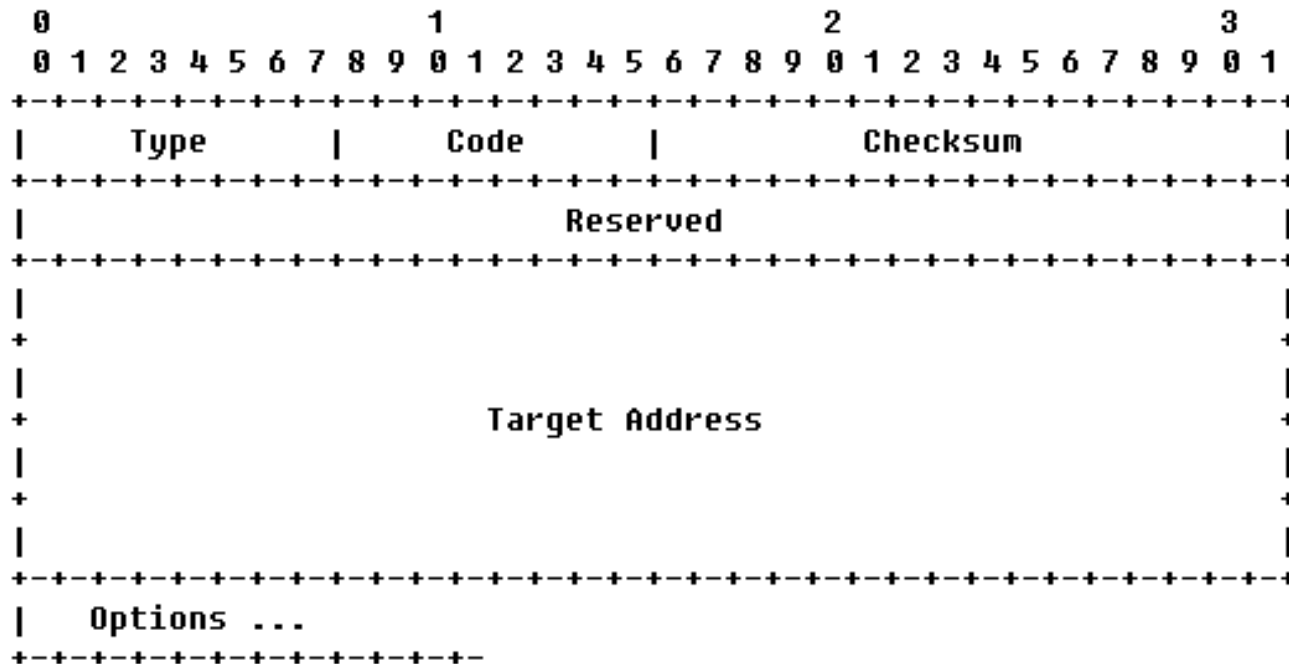
□ IPv6 Neighbor Discovery Option

- 路由器告诉主机IPv6网络的prefix
- Prefix长度为64
- Prefix具体值为3FFE:507:0:1::

```
▼ ICMPv6 Option (Prefix information : 3ffe:507:0:1::/64)
  Type: Prefix information (3)
  Length: 4 (32 bytes)
  Prefix Length: 64
  > Flag: 0xc0, On-link flag(L), Autonomous address-configuration flag(A)
  Valid Lifetime: 3600000
  Preferred Lifetime: 3600000
  Reserved
  Prefix: 3ffe:507:0:1::
```

Neighbor Solicitation报文

- 用途有两个：
 - ▣ 链路层地址解析（等价原来的ARP协议）
 - ▣ 地址重复检测（DAD）
- 报文结构如下：



Neighbor Solicitation报文结构

□ IP部分

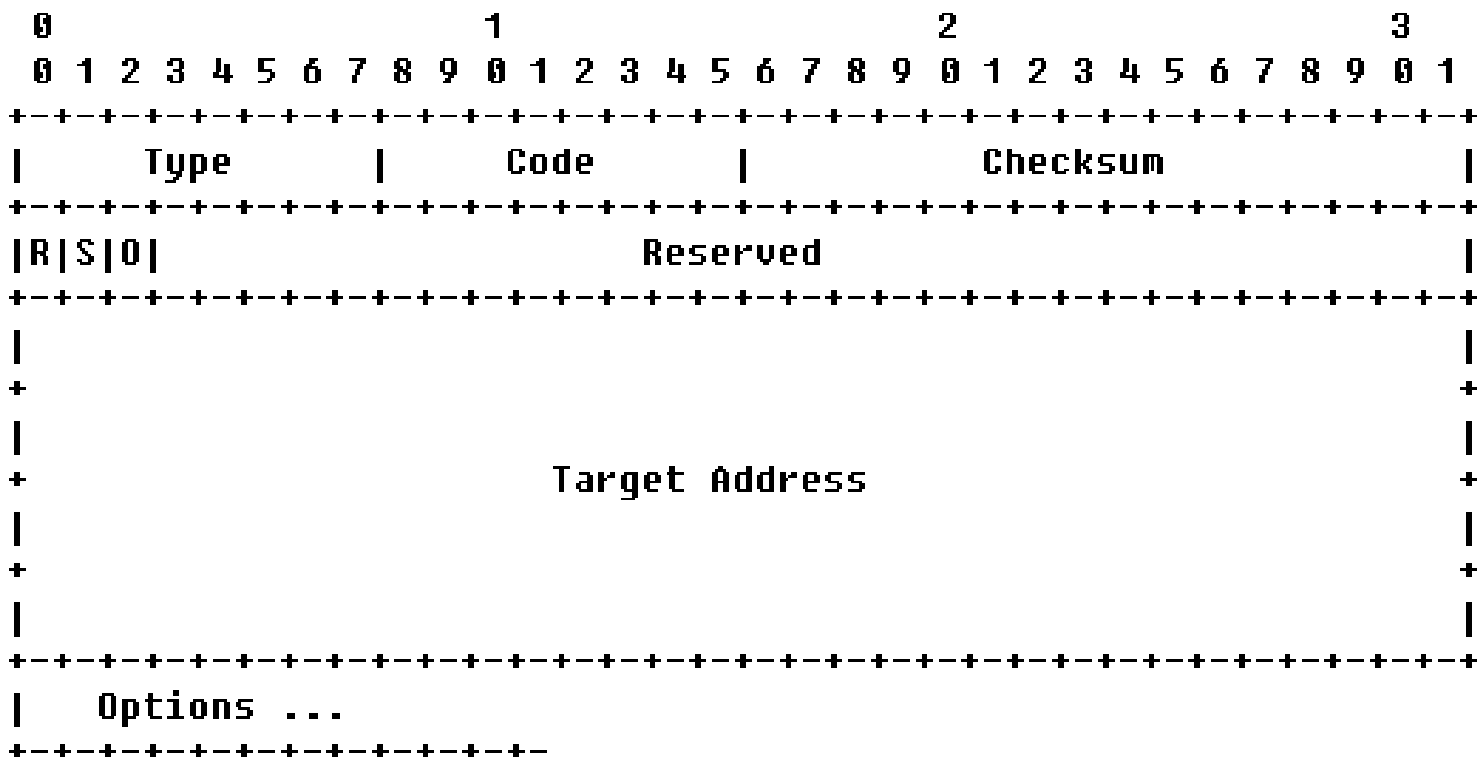
- ▣ 源地址：发送者IPv6地址（地址解析用）或unspecified地址（DAD重复地址检测）
- ▣ 目的地址：Neighbor Solicitations are multicast when the node needs to resolve an address and unicast when the node seeks to verify the reachability of a neighbor.
- ▣ 跳数：255

□ ICMP部分

- ▣ Type=135
- ▣ Code=0
- ▣ Target address=需要解析的IPv6地址或者需要DAD的IPv6地址
- ▣ 选项部分包含发送者链路层地址

Neighbor Advertisement 报文

- NA回复NS报文
- 报文结构如下：



Neighbor Advertisement报文结构

□ IP部分

- ▣ 源地址：发送者IPv6地址
- ▣ 目的地址：全部节点组播地址FF02::1（DAD用）或发送NS的主机单播地址（地址解析用）
- ▣ 跳数：255

□ ICMP部分

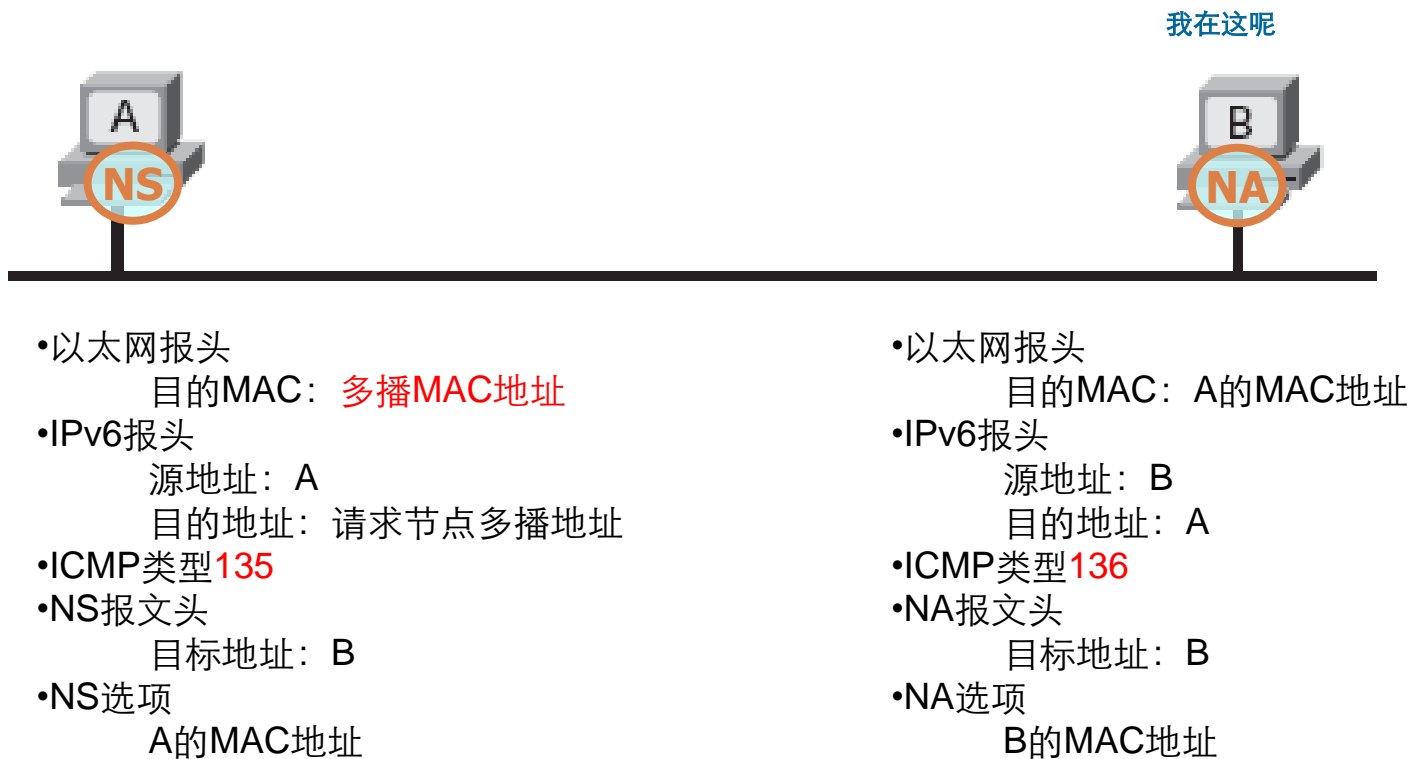
- ▣ Type=136
- ▣ Code=0
- ▣ 选项部分包含了发送者链路层地址
- ▣ R : Router Flag. 发送这个报文的是路由器
- ▣ S : Solicited Flag. 发送这个报文是对Solicitation的回应
- ▣ O : Override flag. 覆盖cache的内容

邻居发现协议—地址解析

- 地址解析在三层完成，不同的二层介质可以采用相同的地址解析协议
- 可以使用三层的安全机制（例如IPSec）避免地址解析攻击
- 使用组播方式发送请求报文，减少了二层网络的性能压力

邻居发现协议—地址解析

- 使用两种ICMPv6报文完成交互过程
 - ▣ 邻居请求NS
 - ▣ 邻居通告NA



邻居发现协议—地址解析实例

R1#ping 2001:db2::1F5C:7A92

Type escape sequence to abort.

Sending 5, 100-byte ICMP Echos to 2001:DB2::1F5C:7A92, timeout is 2 seconds:

! ! ! ! !

Success rate is 100 percent (5/5), round-trip min/avg/max = 16/23/44 ms

```
R1(config-if)#do show ipv6 nei
```

IPv6 Address

Age	Link-layer	Addr	State	Interface
-----	------------	------	-------	-----------

2001:DB2::1F5C:7A92

```
0 c003.2168.0000 REACH Fa0/0
```

NS报文的目的IPv6地址为请求节点组播地址：FF02::1:FF5C:7A92

No.	Time	Source	Destination	Protocol	Length	Info
5	14.8378490	2001:db2::7729:c0ad	ff02::1:ff5c:7a92	ICMPv6	86	Neighbor Solicitation for 2001:db2::1f5c:7a92 from c0:02:21:68:00:00
<div> <div>+</div> <div>Frame 5: 86 bytes on wire (688 bits), 86 bytes captured (688 bits) on interface 0</div> </div>						
<div> <div>+</div> <div>Ethernet II, Src: c0:02:21:68:00:00 (c0:02:21:68:00:00), Dst: IPv6mcast_ff5c:7a:92 (33:33:ff:5c:7a:92)</div> </div>						
<div> <div>-</div> <div>Internet Protocol Version 6, Src: 2001:db2::7729:c0ad (2001:db2::7729:c0ad), Dst: ff02::1:ff5c:7a92 (ff02::1:ff5c:7a92)</div> </div>						
<div> <div>+</div> <div>0110 = Version: 6</div> </div>						
<div> <div>+</div> <div>.... 1110 0000 = Traffic class: 0x000000e0</div> </div>						
<div> <div>.... 0000 0000 0000 0000 0000 = Flowlabel: 0x00000000</div> </div>						
<div> <div>Payload length: 32</div> </div>						
<div> <div>Next header: ICMPv6 (58)</div> </div>						
<div> <div>Hop limit: 255</div> </div>						
<div> <div>Source: 2001:db2::7729:c0ad (2001:db2::7729:c0ad)</div> </div>						
<div> <div>Destination: ff02::1:ff5c:7a92 (ff02::1:ff5c:7a92)</div> </div>						
<div> <div>[Source GeoIP: Unknown]</div> </div>						
<div> <div>[Destination GeoIP: Unknown]</div> </div>						
<div> <div>-</div> <div>Internet Control Message Protocol v6</div> </div>						
<div> <div>Type: Neighbor Solicitation (135)</div> </div>						
<div> <div>Code: 0</div> </div>						
<div> <div>Checksum: 0xf019 [correct]</div> </div>						
<div> <div>Reserved: 00000000</div> </div>						
<div> <div>Target Address: 2001:db2::1f5c:7a92 (2001:db2::1f5c:7a92)</div> </div>						
<div> <div>+</div> <div>ICMPv6 Option (Source link-layer address : c0:02:21:68:00:00)</div> </div>						

邻居发现协议—地址解析实例

NS报文的目的IPv6地址为请求节点组播地址：FF02::1:FF5C:7A92

- 首先IPv6规定，所有的请求节点组播地址前缀为：
FF02::1:FF00:0/104
- FF02表示是本地链接组播
- 前缀为104个bit，剩下的24个bit，来自于邻居发现中的请求地址，
所以两者拼起来得到： FF02::1:FF5C:7A92

为什么这么做呢？

- 因为每个接口配置IPv6地址后都会自动加入到一些多播组中，例如
2001:db2::1F5C:7A92，会自动加入到FF02::1:FF5C:7A92
- 所以用请求节点组播地址，总是可以找到请求节点，同时限制报文接收者

邻居发现协议—地址解析实例

Windows操作系统下，ipconfig无法直接看到组播地址

以太网适配器 以太网:

```
连接特定的 DNS 后缀 . . . . . :  
描述. . . . . : Realtek PCIe GBE Family Controller  
物理地址. . . . . : 70-4D-7B-61-E2-37  
DHCP 已启用 . . . . . : 是  
自动配置已启用. . . . . : 是  
IPv6 地址 . . . . . : 2001:250:4000:4160:d474:a4e6:48cc:918b(首选)  
临时 IPv6 地址. . . . . : 2001:250:4000:4160:8039:5ab2:59ab:8191(首选)  
本地链接 IPv6 地址. . . . . : fe80::d474:a4e6:48cc:918b%6(首选)  
IPv4 地址 . . . . . : 202.114.23.2(首选)  
子网掩码 . . . . . : 255.255.255.0  
获得租约的时间 . . . . . : 2018年1月4日 12:29:53  
租约过期的时间 . . . . . : 2018年1月4日 15:29:54  
默认网关. . . . . : fe80::1614:4bff:fe7d:4cbd%6  
202.114.23.254  
DHCP 服务器 . . . . . : 202.114.23.254  
DHCPv6 IAID . . . . . : 57691515  
DHCPv6 客户端 DUID . . . . . : 00-01-00-01-20-4A-2A-EC-70-4D-7B-61-E2-37  
DNS 服务器 . . . . . : 202.114.0.242  
202.114.0.131  
TCP/IP 上的 NetBIOS . . . . . : 已启用
```


邻居发现协议—地址解析实例

Windows操作系统下：netsh interface ipv6 show joins 可以看出以太接口的两个IPv6地址都对应两个请求节点组播地址

接口 6: 以太网

作用域	参照	上一次	地址
0	0	是	ff01::1
0	0	是	ff02::1
0	1	是	ff02::c
0	2	是	ff02::fb
0	1	是	ff02::1:3
0	1	是	ff02::1:ffab:8191
0	2	是	ff02::1:ffcc:918b

邻居发现协议—重复地址检测 (DAD)

- 重复地址检测确保网络中无两个相同的单播地址
- 所有Unicast地址都需要做DAD
- 使用NS和NA完成DAD交互过程
- 每次系统默认一个主机在应用新的IP地址之前会发送3次DAD,如果三次以后均没有收到任何回应,那么该地址被认为是可以配置在接口上的
- 若发现有地址重复 (存在NA报文, 目的地址为FF02::1)
 - ▣ 全局单播地址: 不安排给接口
 - ▣ 链路本地地址: 将接口置于不可用状态

邻居发现协议—重复地址检测 (DAD)

IPv4自身是不具备重复地址检测的，依赖免费ARP（gratuitous ARP）又能称无偿ARP、无故ARP

- 指主机发送ARP request 报文查询自己的IP地址
- 确定网络中是否有其他的主机使用了该IP地址，如果有应答则产生错误消息
- 一般在ARP功能开启或者端口初始化配置完成时，主机向网络发送免费ARP来查询自己的IP地址确认地址唯一可用。

免费ARP的另外一个作用：

- 免费ARP可以更新ARP表项用，网络中其他主机收到该广播则在缓存中更新条目，收到主机强制更新，如果存在旧条目会将MAC更新为广播包中MAC

邻居发现协议—重复地址检测 (DAD)

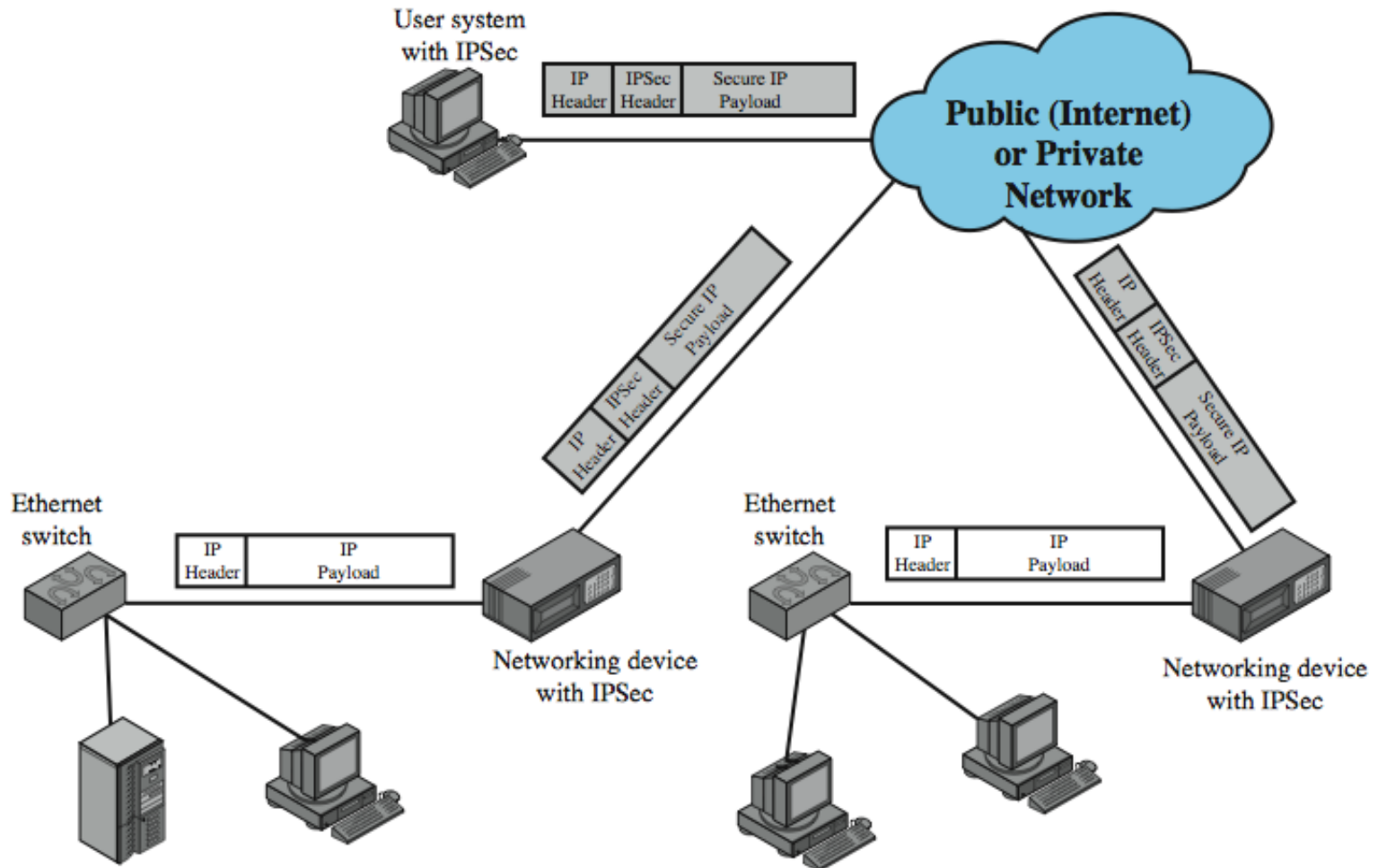
- 源地址为 ::
- 目标地址为请求节点组播地址
- Target地址为需要检测的IPv6地址

```
> Frame 1: 78 bytes on wire (624 bits), 78 bytes captured (624 bits) on interface 0
> Ethernet II, Src: Vmware_5f:11:9f (00:0c:29:5f:11:9f), Dst: IPv6mcast_ff:7e:4a:1e (33:33:ff:7e:4a:1e)
▼ Internet Protocol Version 6, Src: ::, Dst: ff02::1:ff7e:4a1e
    0110 .... = Version: 6
    > .... 0000 0000 .... .... .... .... = Traffic Class: 0x00 (DSCP: CS0, ECN: Not-ECT)
    .... .... 0000 0000 0000 0000 0000 = Flow Label: 0x000000
    Payload Length: 24
    Next Header: ICMPv6 (58)
    Hop Limit: 255
    Source: ::
    Destination: ff02::1:ff7e:4a1e
    [Source GeoIP: Unknown]
    [Destination GeoIP: Unknown]
▼ Internet Control Message Protocol v6
    Type: Neighbor Solicitation (135)
    Code: 0
    Checksum: 0xbdb1 [correct]
    [Checksum Status: Good]
    Reserved: 00000000
    Target Address: fe80::1816:c126:507e:4a1e
```

2.5 IPSec

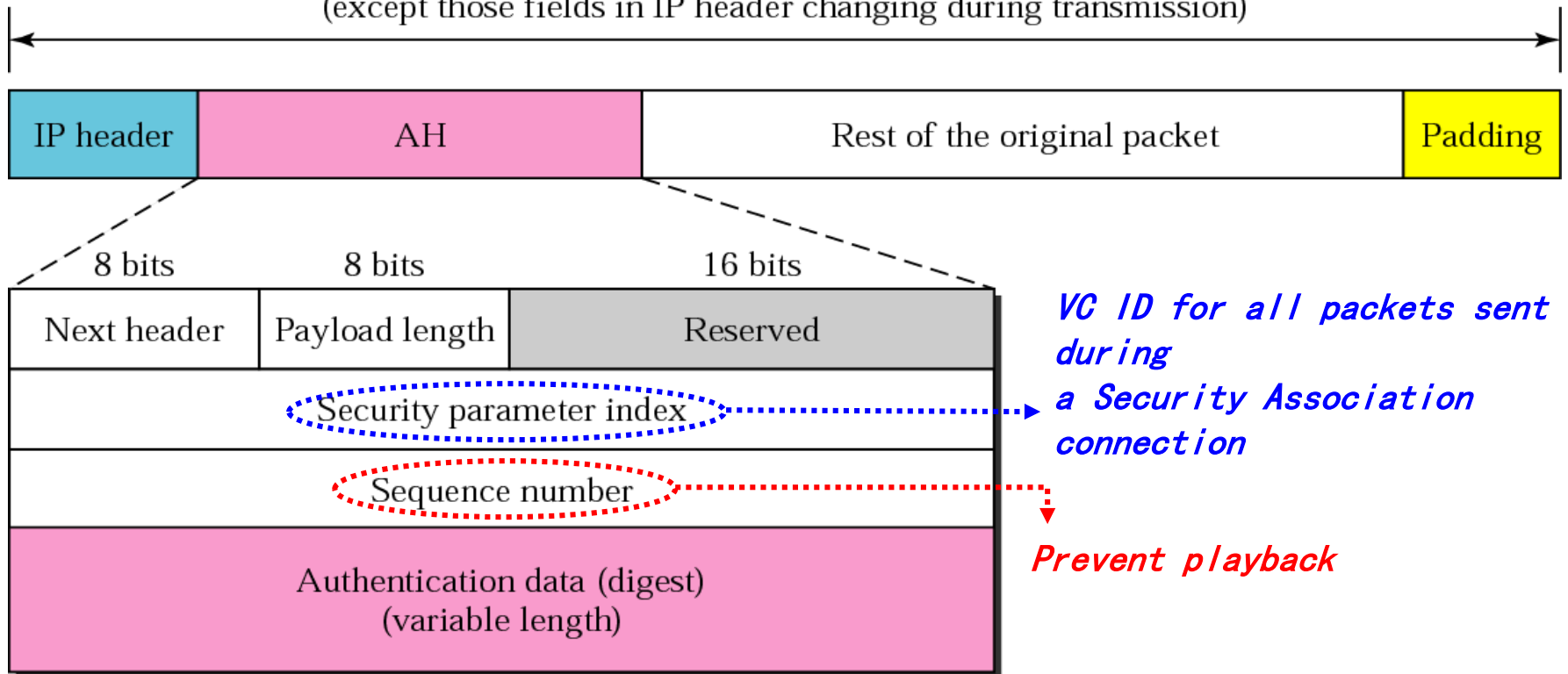
- ❑ IPSec是IETF（Internet Engineering Task Force, Internet工程任务组）的IPSec小组建立的一组IP安全协议集。IPSec定义了在网络层使用的安全服务，其功能包括数据加密、对网络单元的访问控制、数据源地址验证、数据完整性检查和防止重放攻击。
- ❑ IPv6强制安全标准
- ❑ IPv4可以选择实现

IPsec Scenario

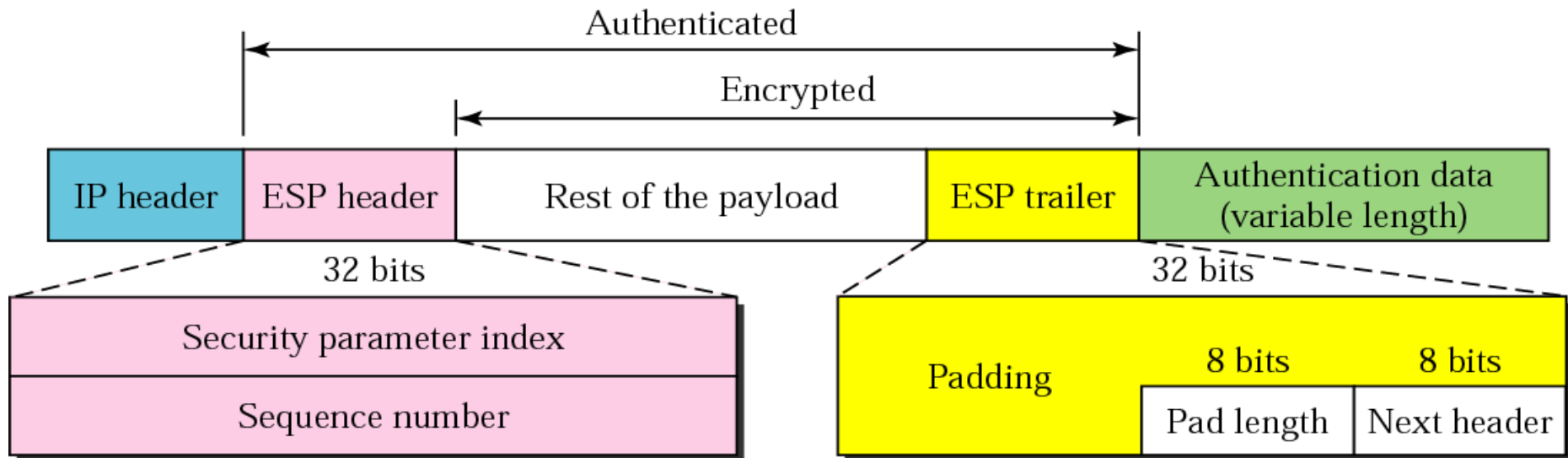


Authentication Header

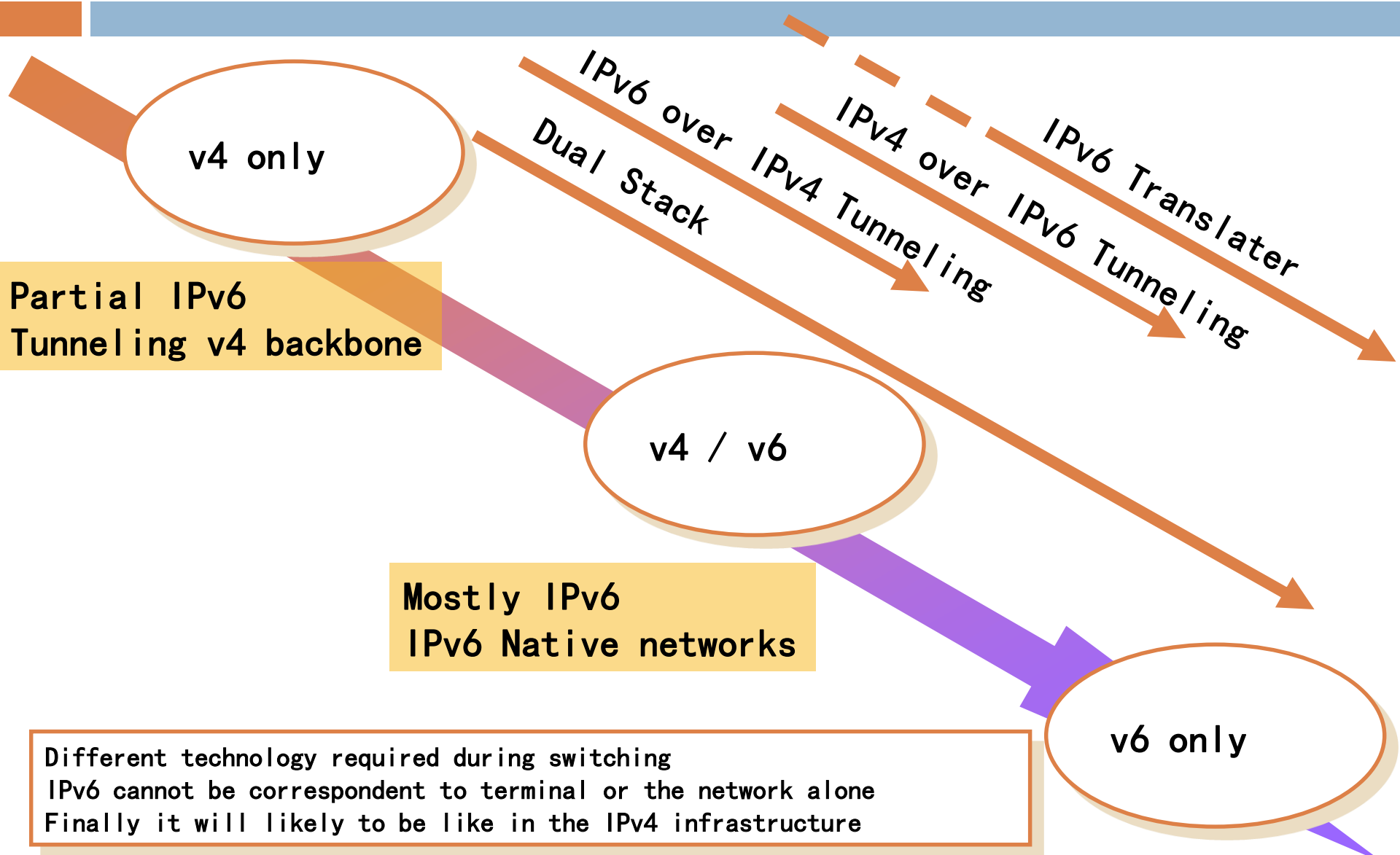
Data used in calculation of authentication data
(except those fields in IP header changing during transmission)



Encapsulating Security Payload



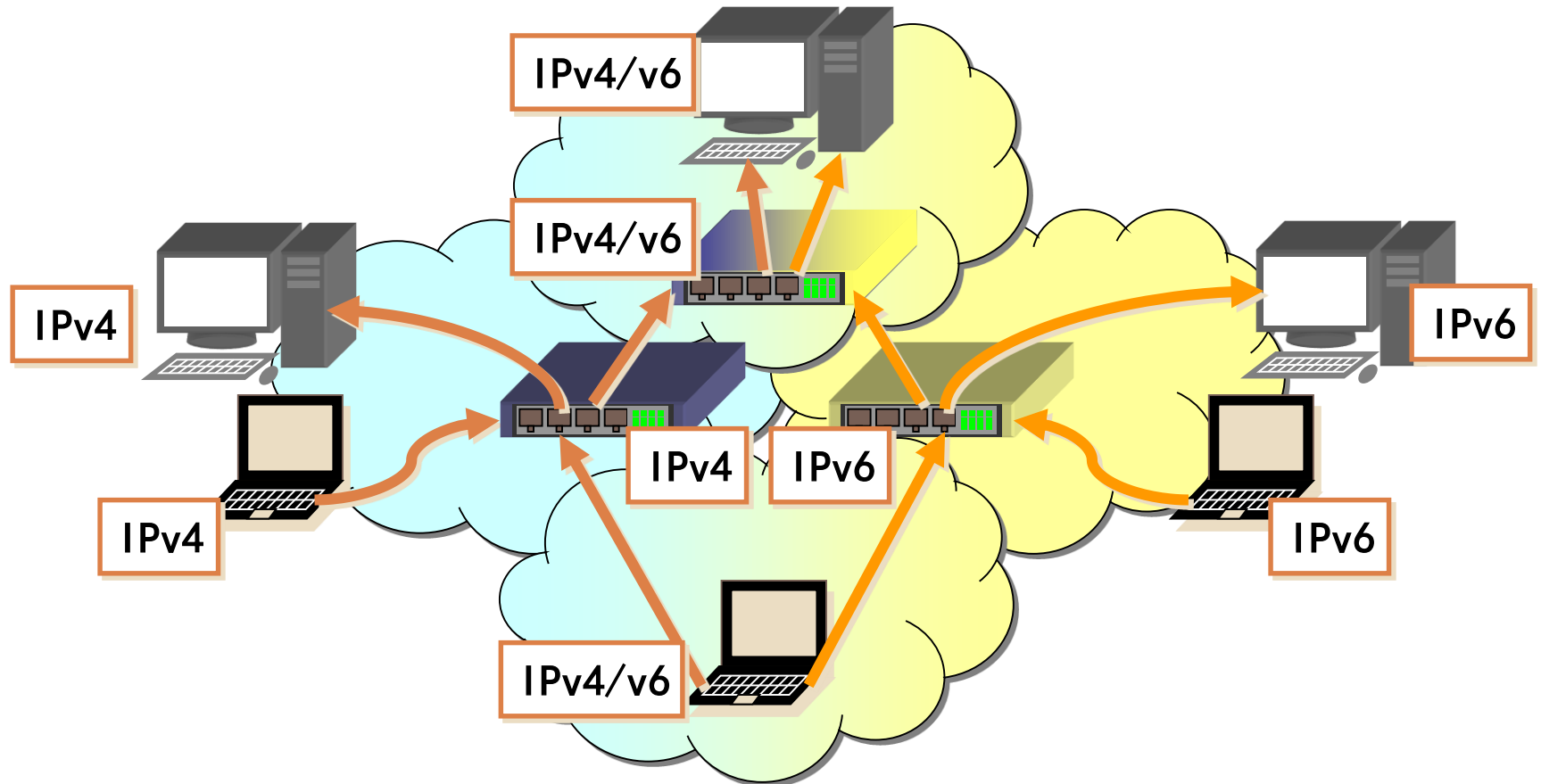
2.6 From IPv4 to IPv6



双栈 (Dual Stack)

82

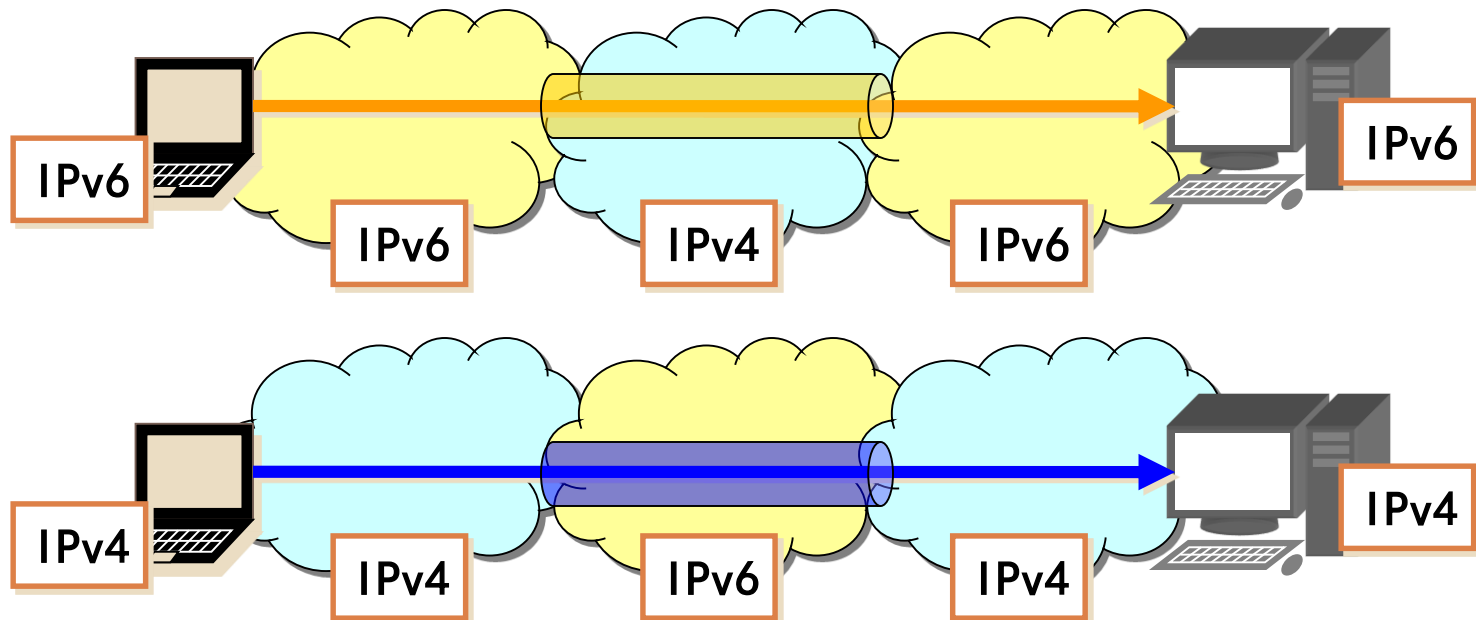
- IPv4/IPv6 can be used
- Server/Router/Client
- Until No IPv4 nodes are available



隧道 (Tunneling)

83

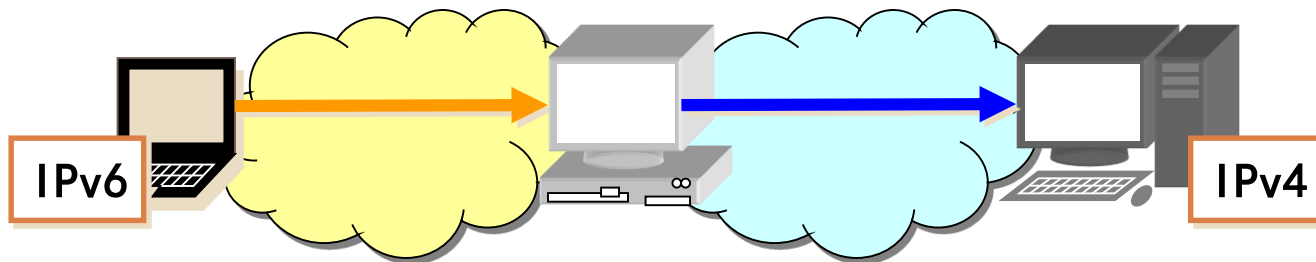
- IPv6 network tunnels through IPv4 network (IPv4 network tunnels through IPv6 network)
- Encapsulation mechanism



转换器 (Translator)

84

- To communicate IPv4 only supported host to IPv6 only supported host
- NAT, SOCKS, Layer realization



问题?

85

- 已知华中科技大学(总面积4517542平方米=约450万平方米=4.5平方公里)分配到的IPv6地址是
 - ▣ 2001: 0250: 4000::/48
- 请问这是一个何种类型的 Unicast / Multicast地址? 其相应FP/TLA ID/RES/NLA ID/SLA ID/接口ID分别是多少?
- HUST是一个TLA ID或 NLA ID或SLA ID机构?
- HUST所分得的地址空间相当于v4的多少个A类地址? 该地址占整个IPv6地址空间的比例是多大? 解答
 - ▣ 0010 0000 0000 0001: 0000 0010 0101 0000: 0100 0000 0000 0000: : /48
 - ▣ $2^{128-48} = 2^{80} \approx (2^{10})^8 = (10^3)^8 = 10^{24} / 450 \text{万} \text{m}^2 = 2.22 \times 10^{17} / \text{m}^2$
 - ▣ $2.22 \times 10^{17} / 6.02 \times 10^{23} \approx 0.368 \times 10^{-6} = \text{百万分之一的3分之一个摩尔数}$
 - ▣ $2^{80} / 2^{24} = 2^{56} \approx (10^3)^{5.6} = 10^{16.8} \text{个A类}$
 - ▣ $2^{80} / 2^{128} = 1 / 2^{48} \approx 1 / 10^{14.4} = \text{百万亿分之一}$