Internet Fundamentals

Contact: training@apnic.net

APNIC

Overview

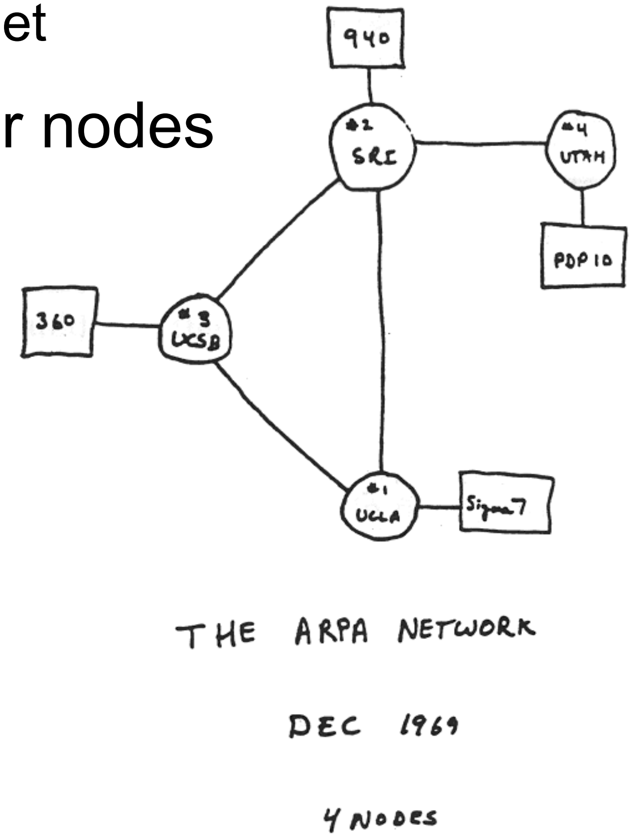
- History of the Internet
- Internet Operations Fundamentals
- Introduction to APNIC
- Managing Internet Resources
- Policy Development Process
- IP Addressing Basics
- IP Routing Basics
- DNS and Reverse DNS

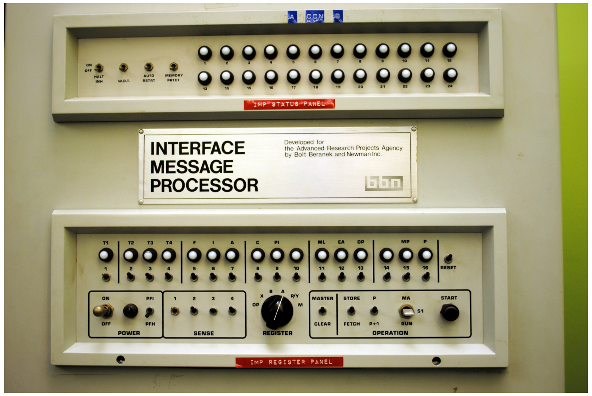
History of the Internet

APNIC

In the beginning…

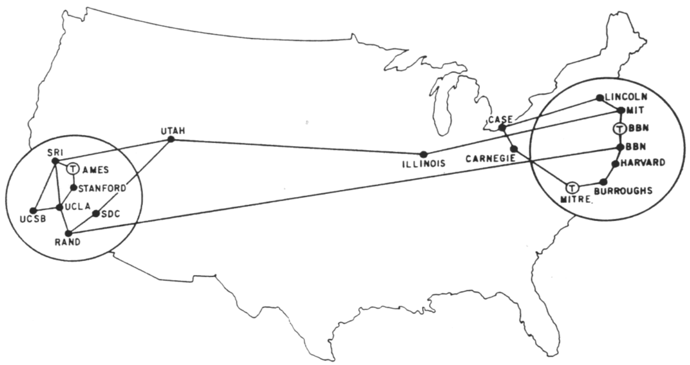
- 1968 - DARPA
  - (Defense Advanced Research Projects Agency) contracts with BBN to create ARPAnet
- 1969 – First four nodes





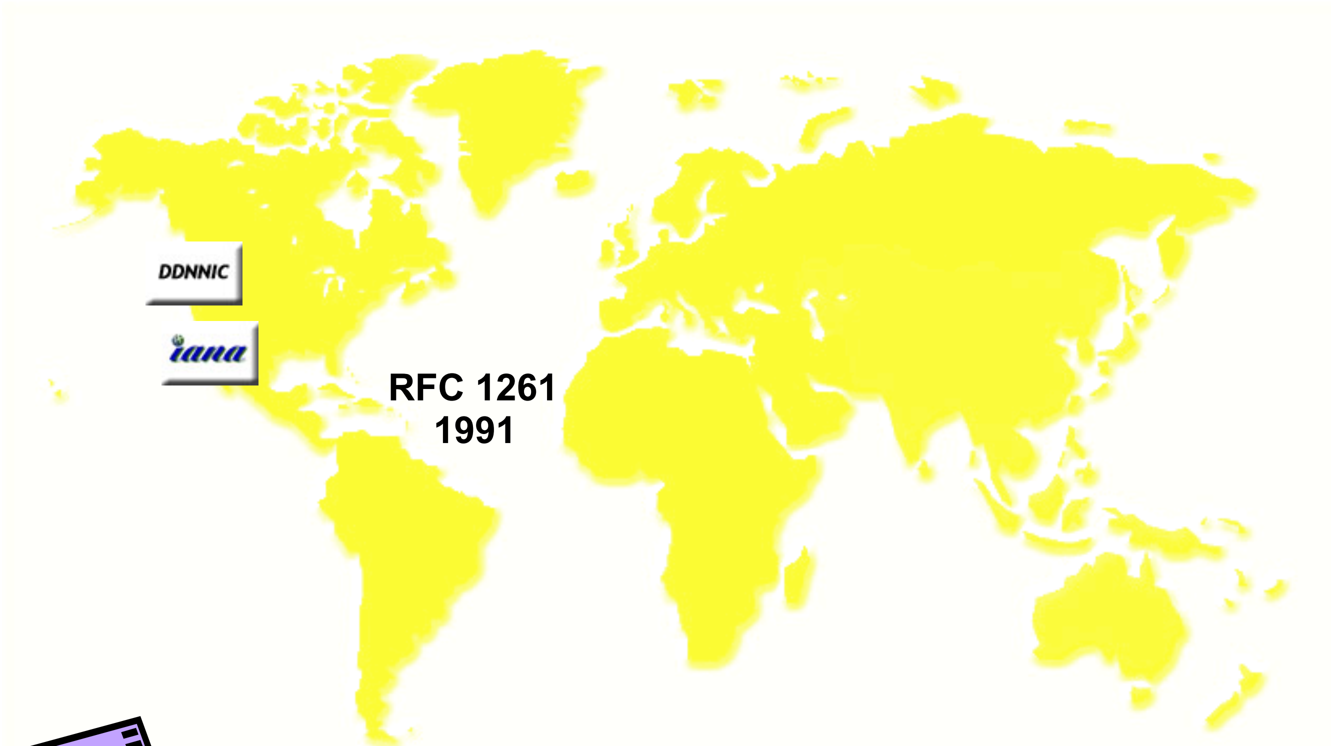
The Internet is born…

- 1970 - Five nodes:
  - UCLA – Stanford - UC Santa Barbara - U of Utah – BBN
- 1971 – 15 nodes, 23 hosts connected

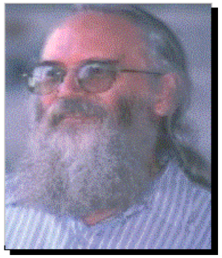


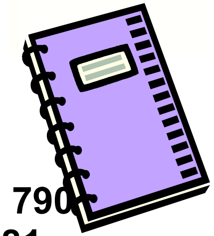
- 1974 – TCP specification by Vint Cerf & Bob Kahn
- 1983 – TCP/IP
  - On January 1, the Internet with its 1000 hosts converts en masse to using TCP/IP for its messaging

Pre 1992



RFC 1020
1987





RFC 790
1981

“The assignment of numbers is also handled by Jon. If you are developing a protocol or application that will require the use of a link, socket, port, protocol, or network number **please contact Jon to receive a number assignment**.”

Address Architecture - History

- Initially, only 256 networks in the Internet!
- Then, network “classes” introduced:
  - Class A (128 networks x 16M hosts)
  - Class B (16,384 x 65K hosts)
  - Class C (2M x 254 hosts)

Address Architecture - Classful

Class A: 128 networks x 16M hosts (50% of all address space)

	A (7 bits)	Host address (24 bits)

0

0-127

Class B: 16K networks x 64K hosts (25%)

	B (14 bits)	Host (16 bits)

10

128-191

Class C: 2M networks x 254 hosts (12.5%)

	C (21 bits)	Host (8 bits)

110

192-223

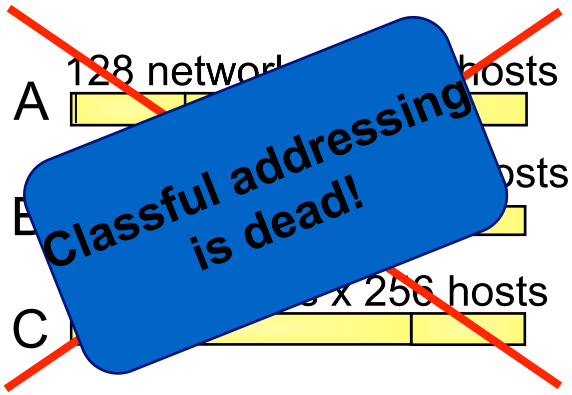
Internet Challenges 1992

- Address space depletion
  - IPv4 address space is finite
  - Historically, many wasteful allocations
- Routing chaos
  - Legacy routing structure, router overload
  - CIDR & aggregation are now vital
- Inequitable management
  - Unstructured and wasteful address space distribution

Classless & Classful addressing

Best Current Practice

Classful



Obsolete
- *inefficient*
- *depletion of B space*
- *too many routes from C space*

Classless

Addresses	Prefix	Classful	Net Mask
...	...	...	...
8	/29		255.255.255.248
16	/28		255.255.255.240
32	/27		255.255.255.224
64	/26		255.255.255.192
128	/25		255.255.255.128
256	/24	1 C	255.255.255.0
...	...	...	...
4096	/20	16 C's	255.255.240.0
8192	/19	32 C's	255.255.224
16384	/18	64 C's	255.255.192
32768	/17	128 C's	255.255.128
65536	/16	1 B	255.255.0.0
...	...	...	... *

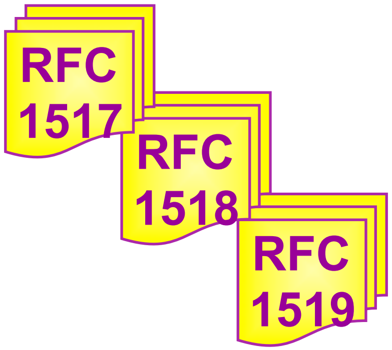
- Network boundaries may occur at *any* bit

Evolution of Internet Eco System

APNIC

Evolution of Internet Resource Management

- 1993: Development of “CIDR”
  - addressed both technical problems



Address depletion

→ Through more accurate assignment

- variable-length network address

Routing table overload

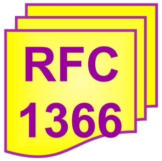
→ Through address space aggregation

- “ supernetting”

Evolution of Internet Resource Management

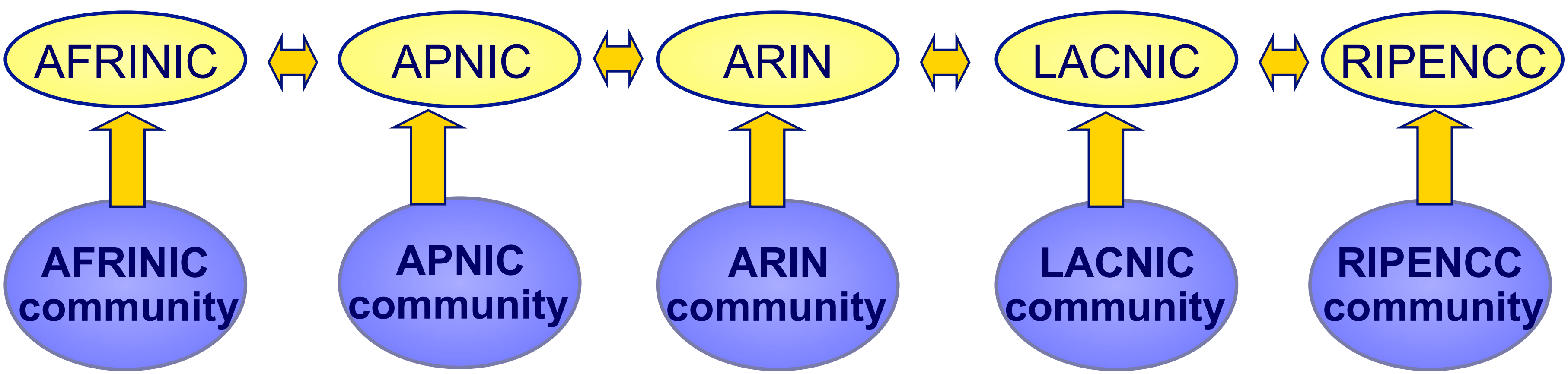
- Administrative problems remained
  - Increasing complexity of CIDR-based allocations
  - Increasing awareness of conservation and aggregation
  - Need for fairness and consistency

- RFC 1366 (1992)
  - Described the "growth of the Internet and its increasing globalization"
  - Additional complexity of address management
  - Set out the basis for a regionally distributed Internet registry system

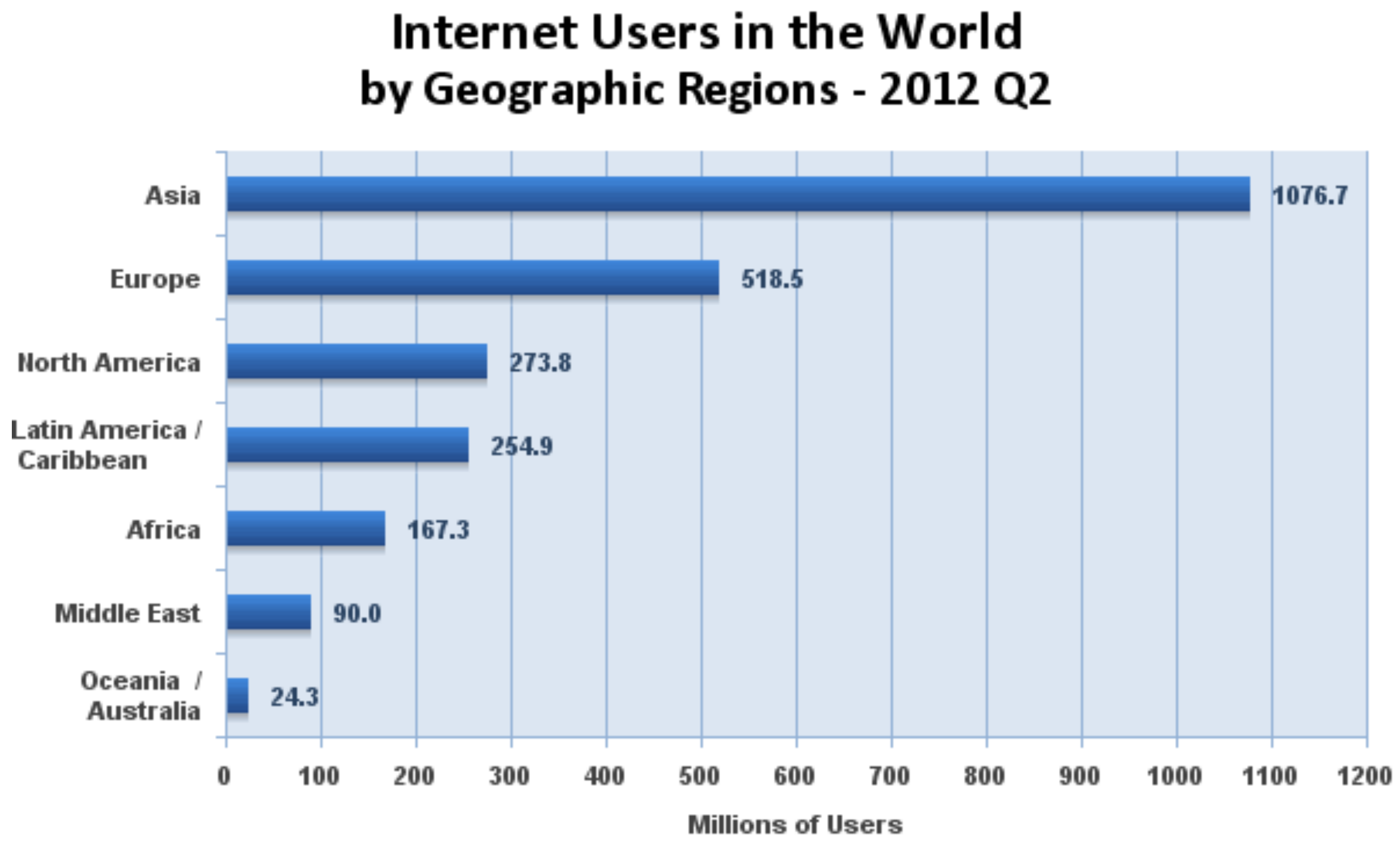


Evolution of Address Policy

- Establishment of RIRs
  - Regional open processes
  - Cooperative policy development
  - Industry self-regulatory model
    - bottom up

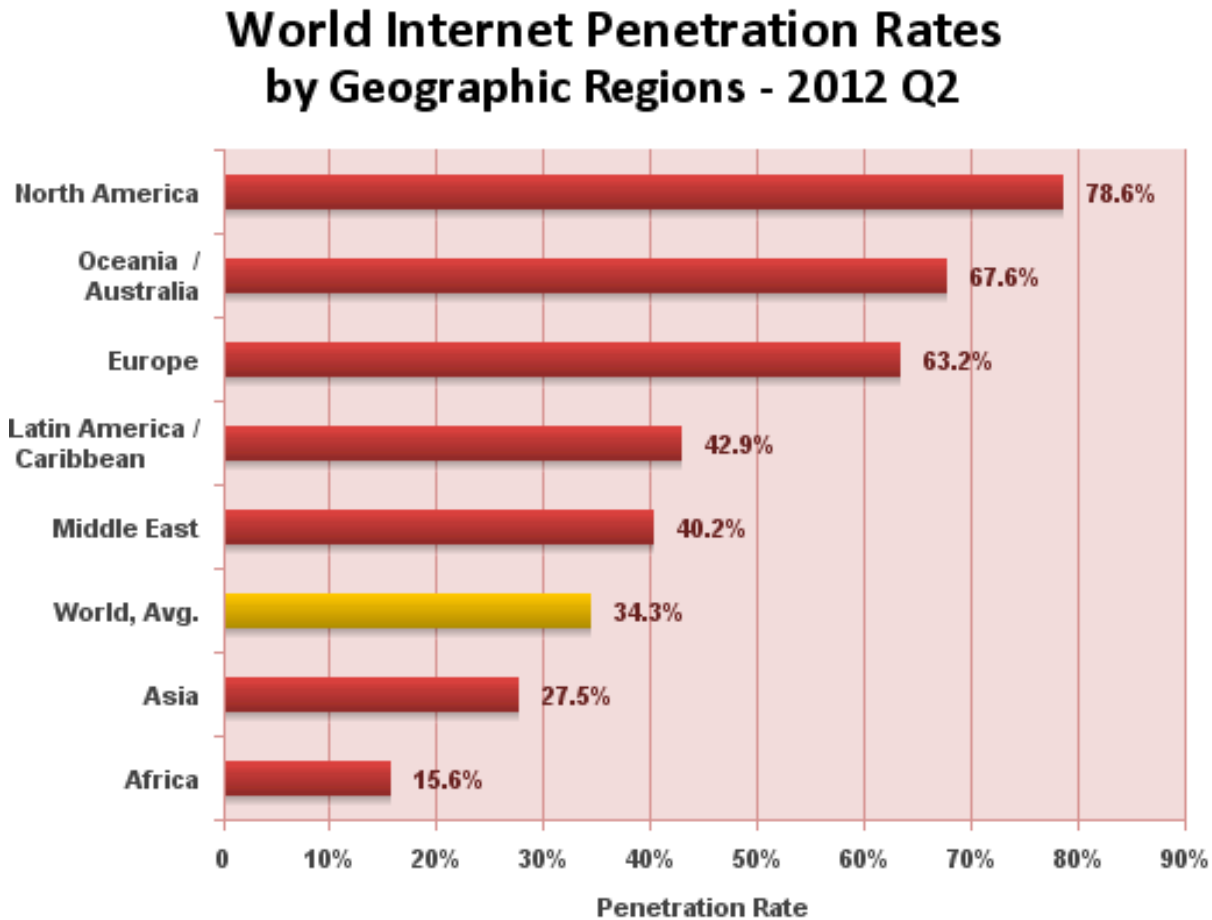


World Internet Users Today



Source: Internet World Stats - www.internetworldstats.com/stats.htm
2,405,518,376 Internet users estimated for June 30, 2012

World Internet Penetration Today



Source: Internet World Stats - www.internetworldststs.com/stats.htm
Penetration Rates are based on a world population of 7,017,846,922 and 2,405,518,376 estimated Internet users on June 30, 2012.

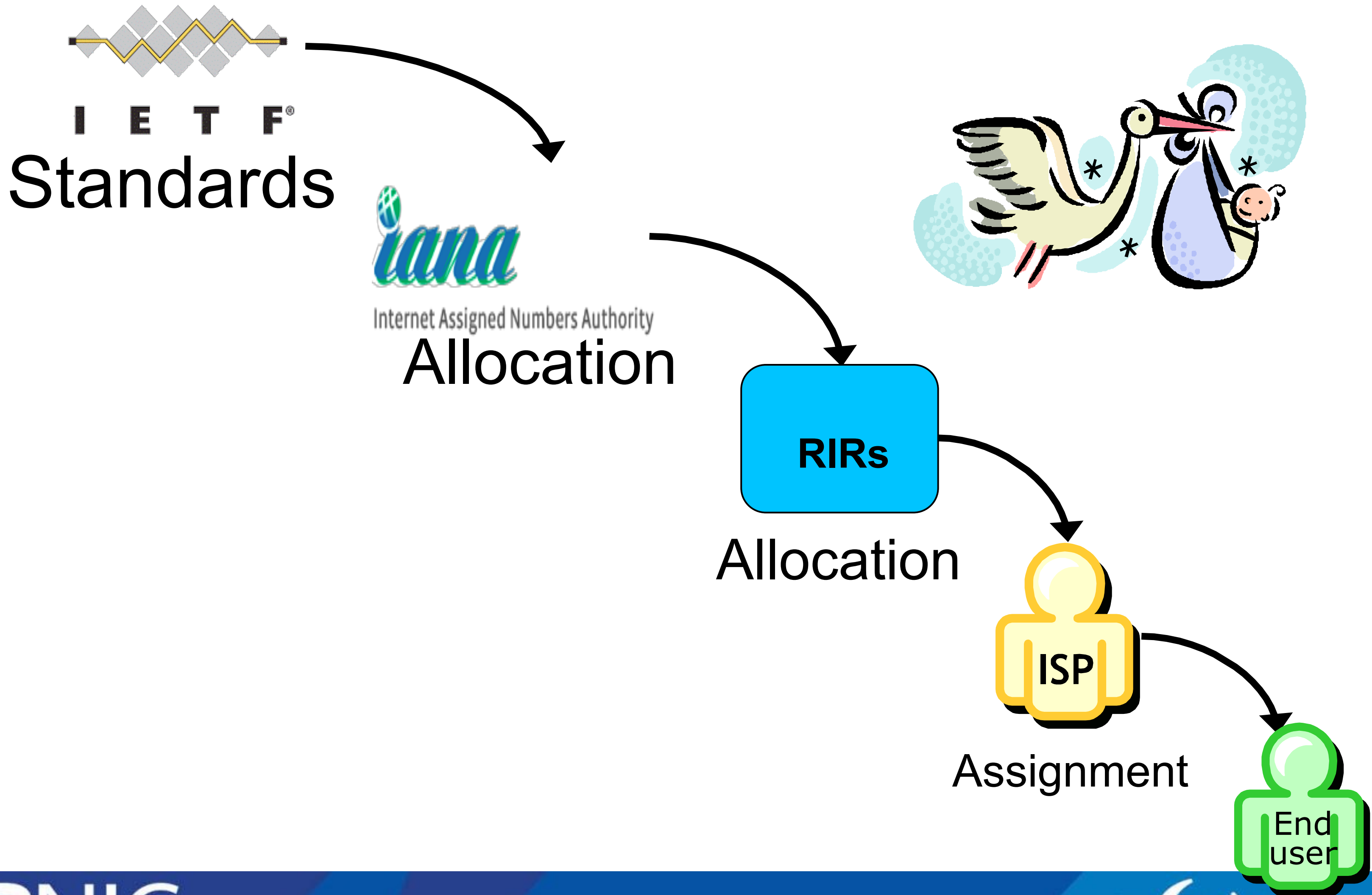
Internet Operational Fundamentals

APNIC

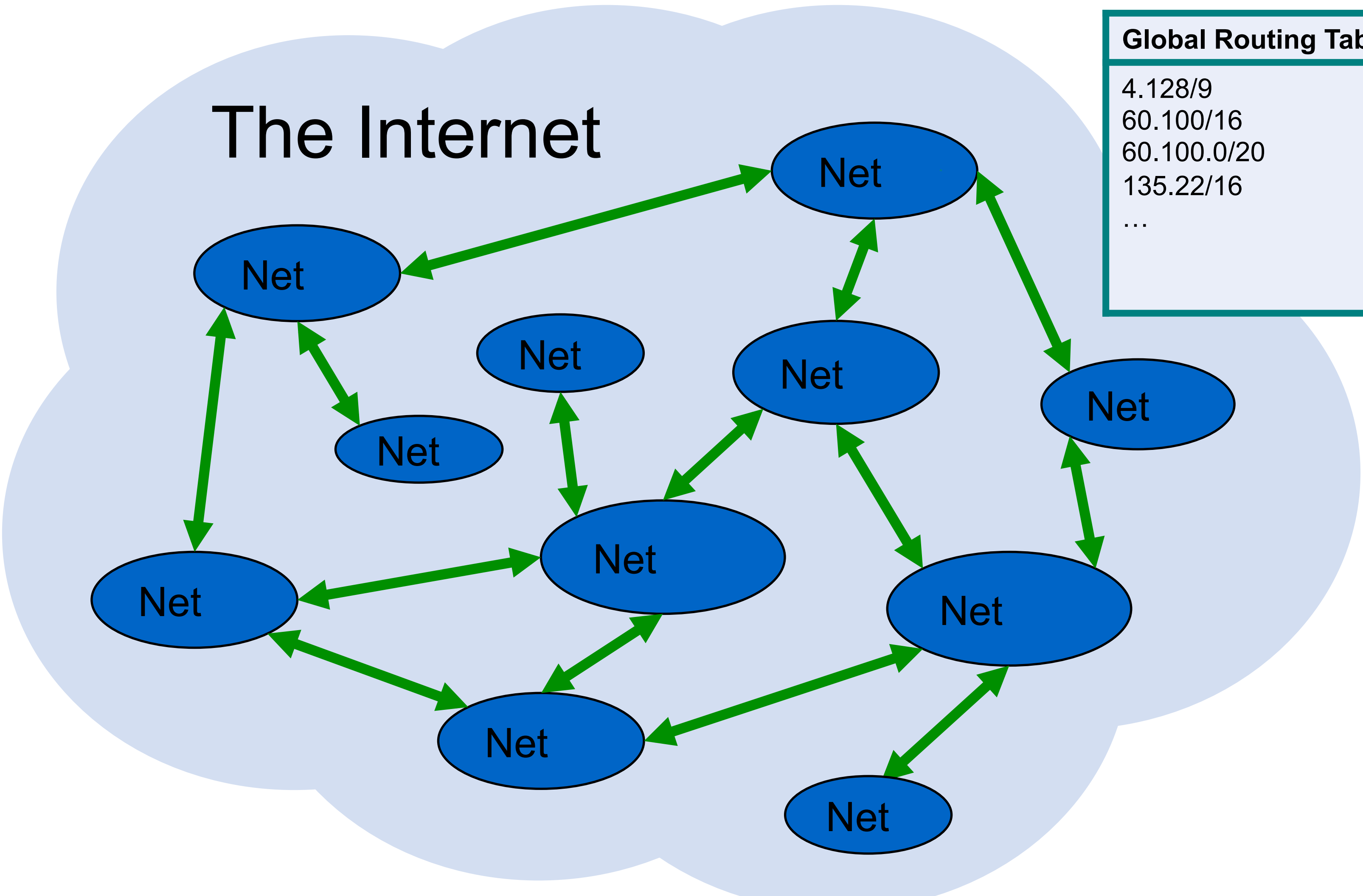
How does the Internet work

- Physical connectivity and reachability
  - Packet switching
- Protocols – common communication and rules
  - TCP/IP
- Addressing – global accessibility
  - IPv4, AS numbers, IPv6
    - IANA - RIRs

Where do IP addresses come from?



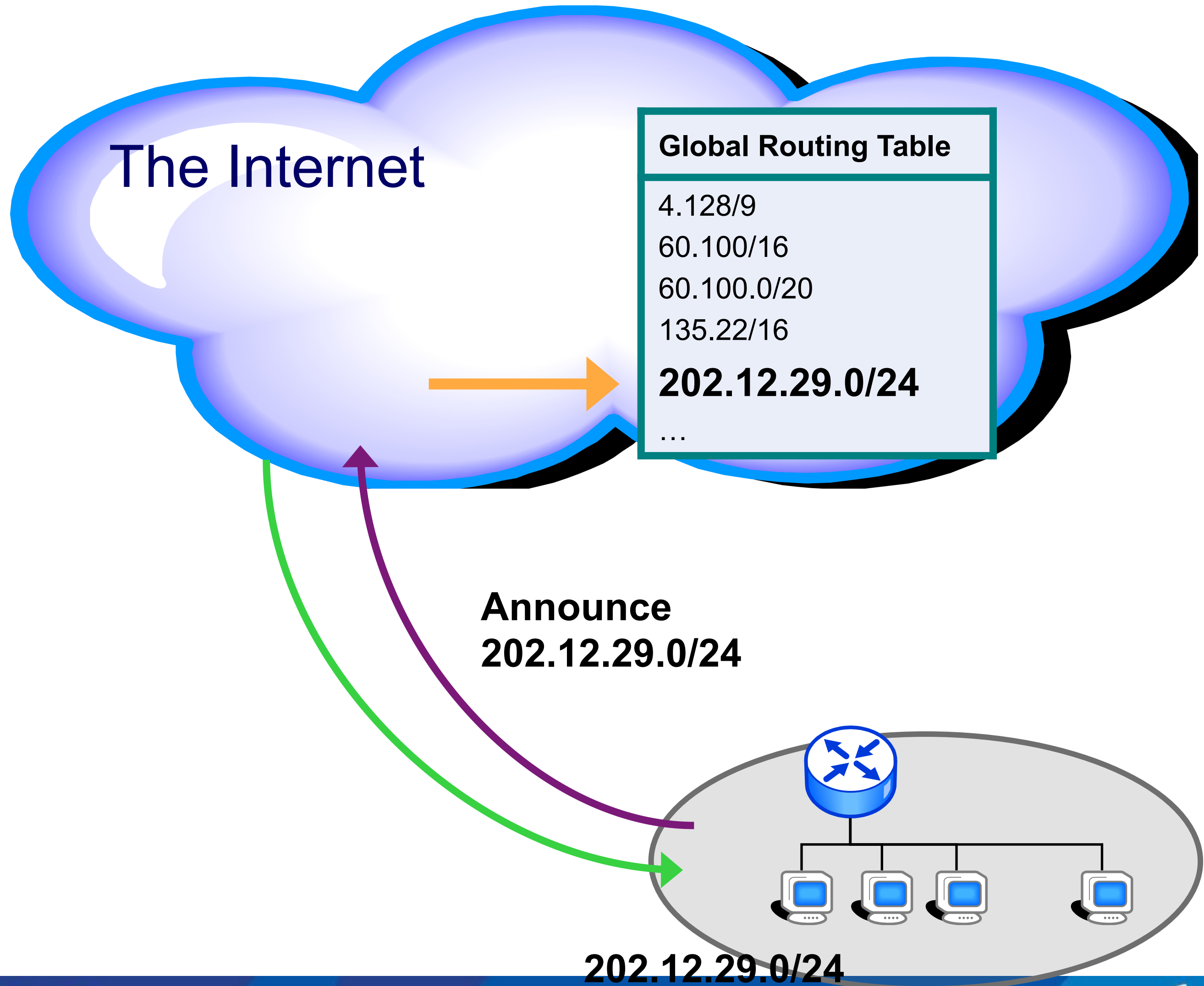
Internet Routing



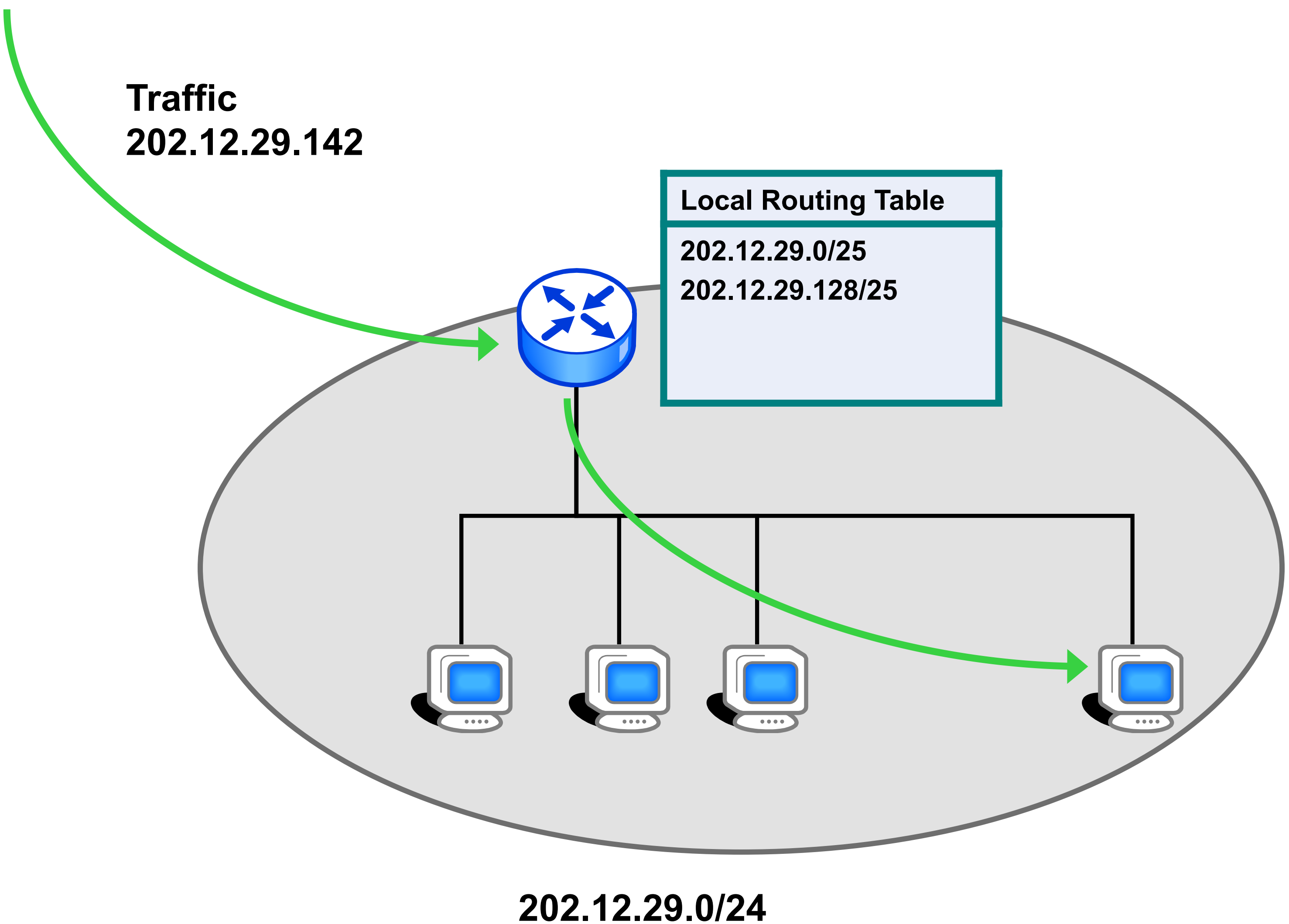
Global Routing Table

4.128/9
60.100/16
60.100.0/20
135.22/16
…

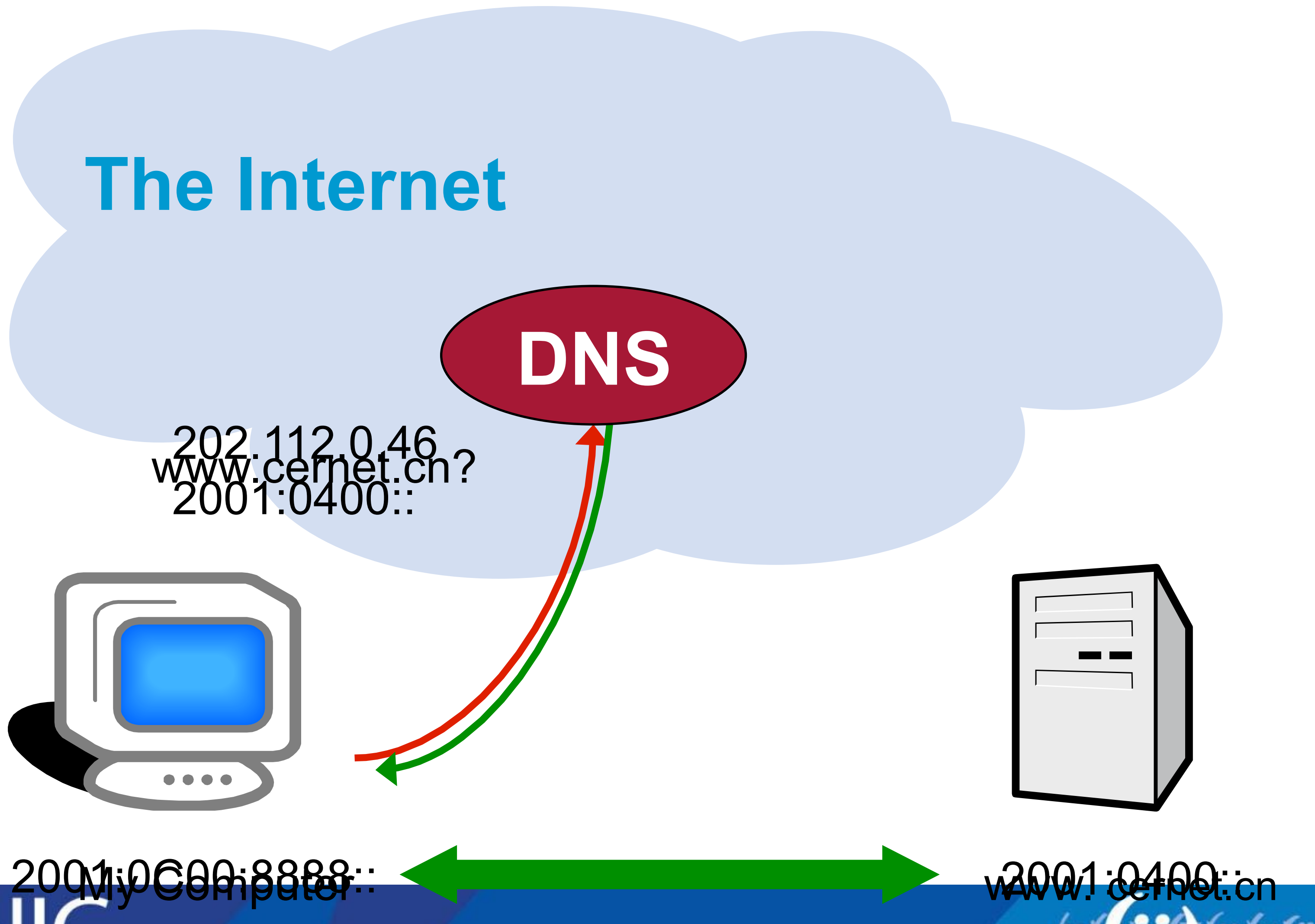
Internet Routing



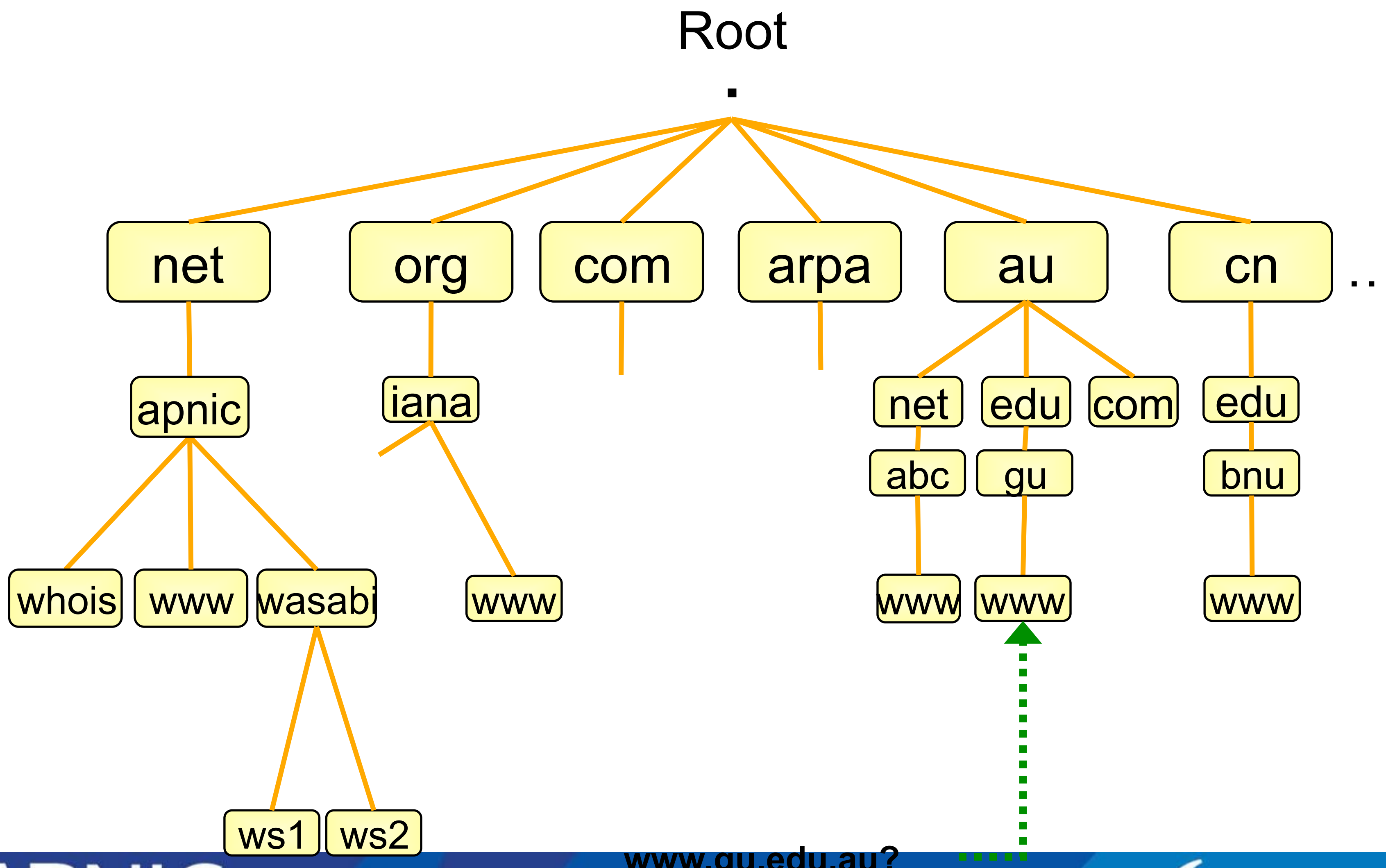
Internet Routing



IP Addresses vs Domain Names

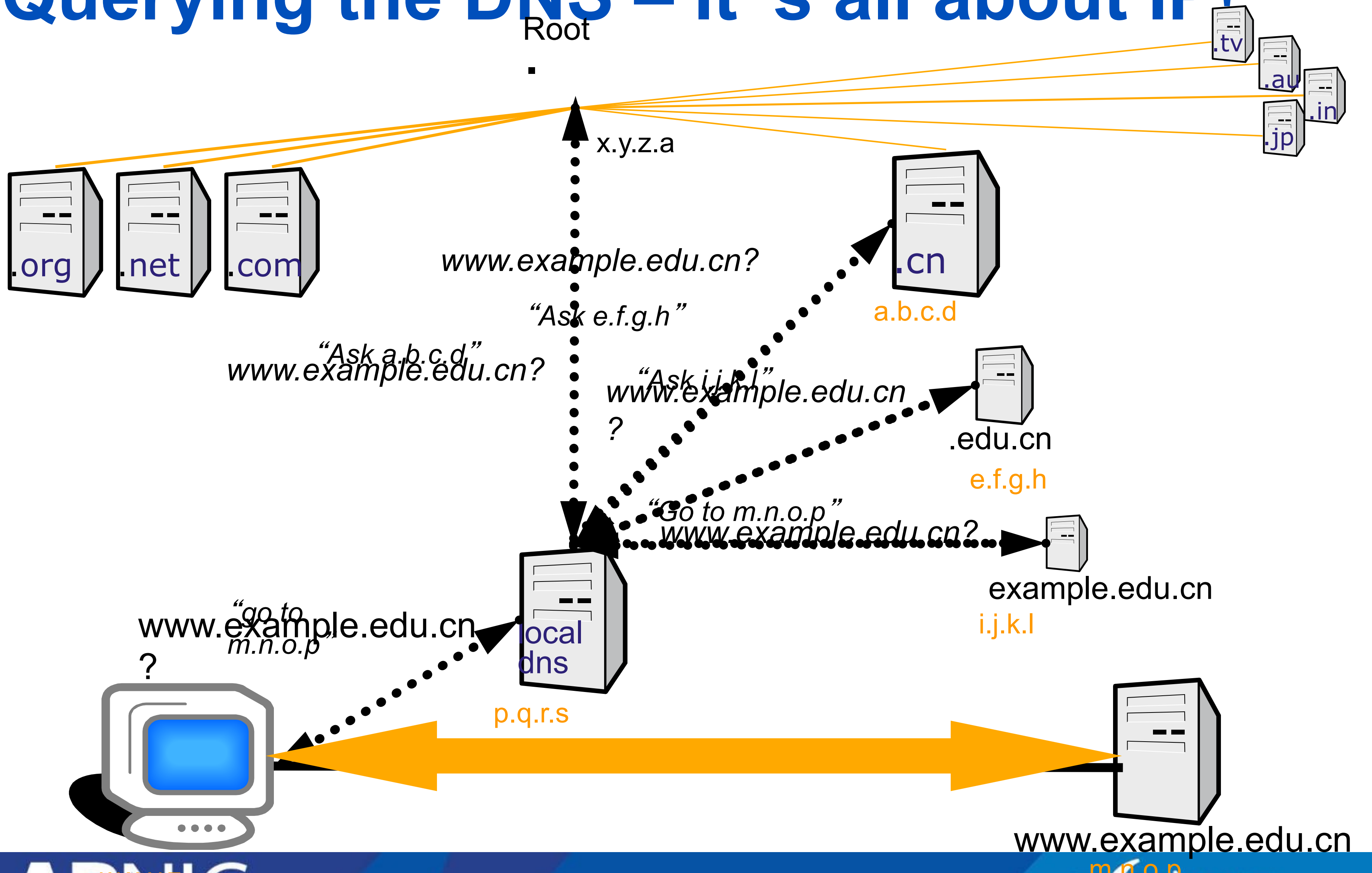


The DNS tree



Root
.
net
org
com
arpa
au
cn
…
apnic
iana
net
edu
com
edu
abc
gu
bnu
whois
www
wasabi
www
www
www
www
ws1
ws2
www.gu.edu.au?

Querying the DNS – It's all about IP!



Who Runs the Internet?

- No one
- (Not ICANN, not the RIRs, not the governments…)
- It is decentralized

How does it Keep on Working

- Inter-provider business relationships and the need for customer reachability ensures that the Internet by and large functions for the common good
- Driven by commerce – free market
- Engineers and the Internet community talk to each other

Regional Internet Registry System

APNIC

Regional Internet Registries

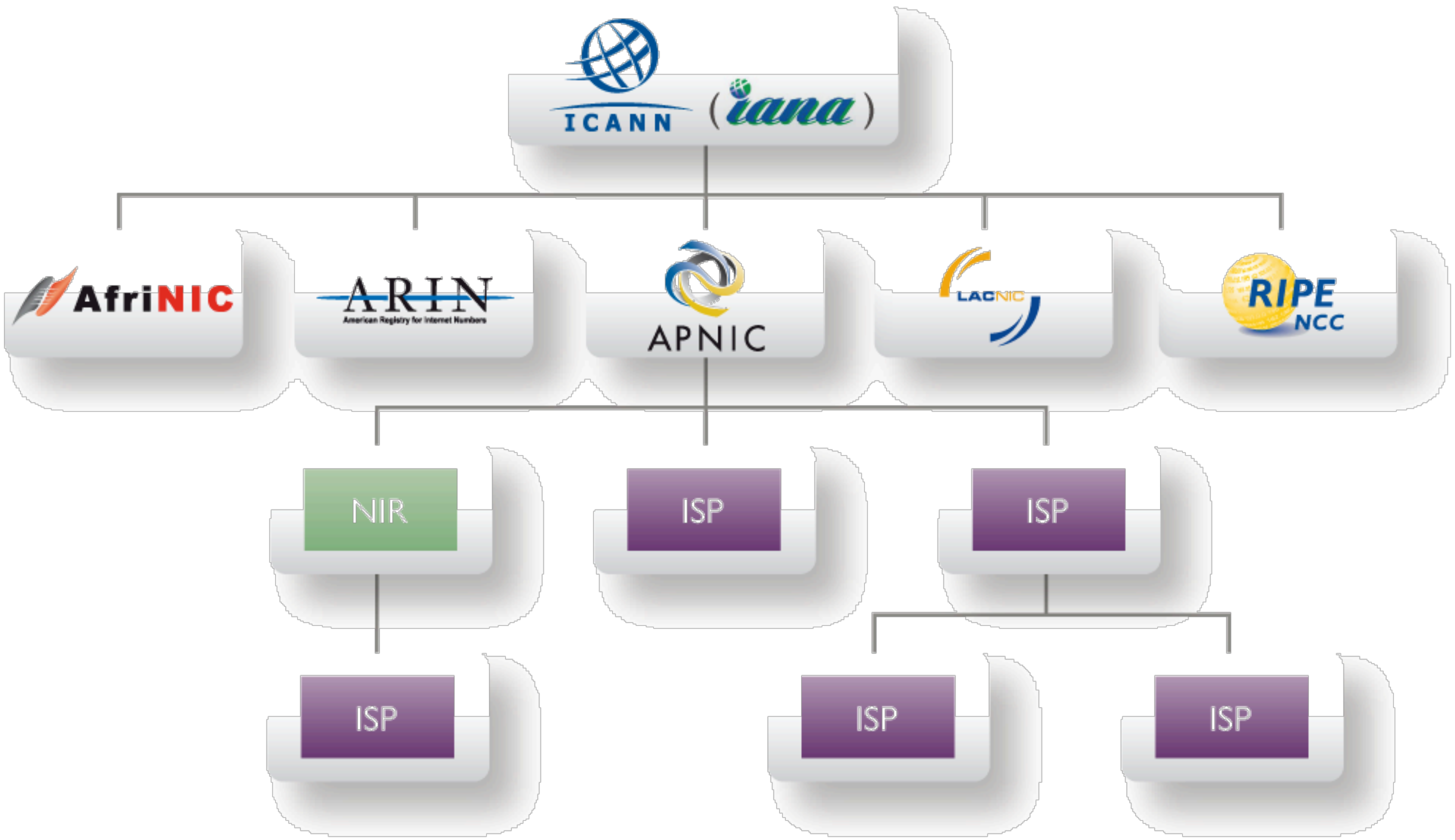
- RIRs manage, distribute, and register Internet number resources (IPv4 and IPv6 addresses and Autonomous System Numbers) within their respective regions.
  - Ensuring the fair distribution and responsible management
- Five RIRs:
  - AfriNIC, APNIC, ARIN, LACNIC, RIPE NCC

What are the Goals of the RIRs?

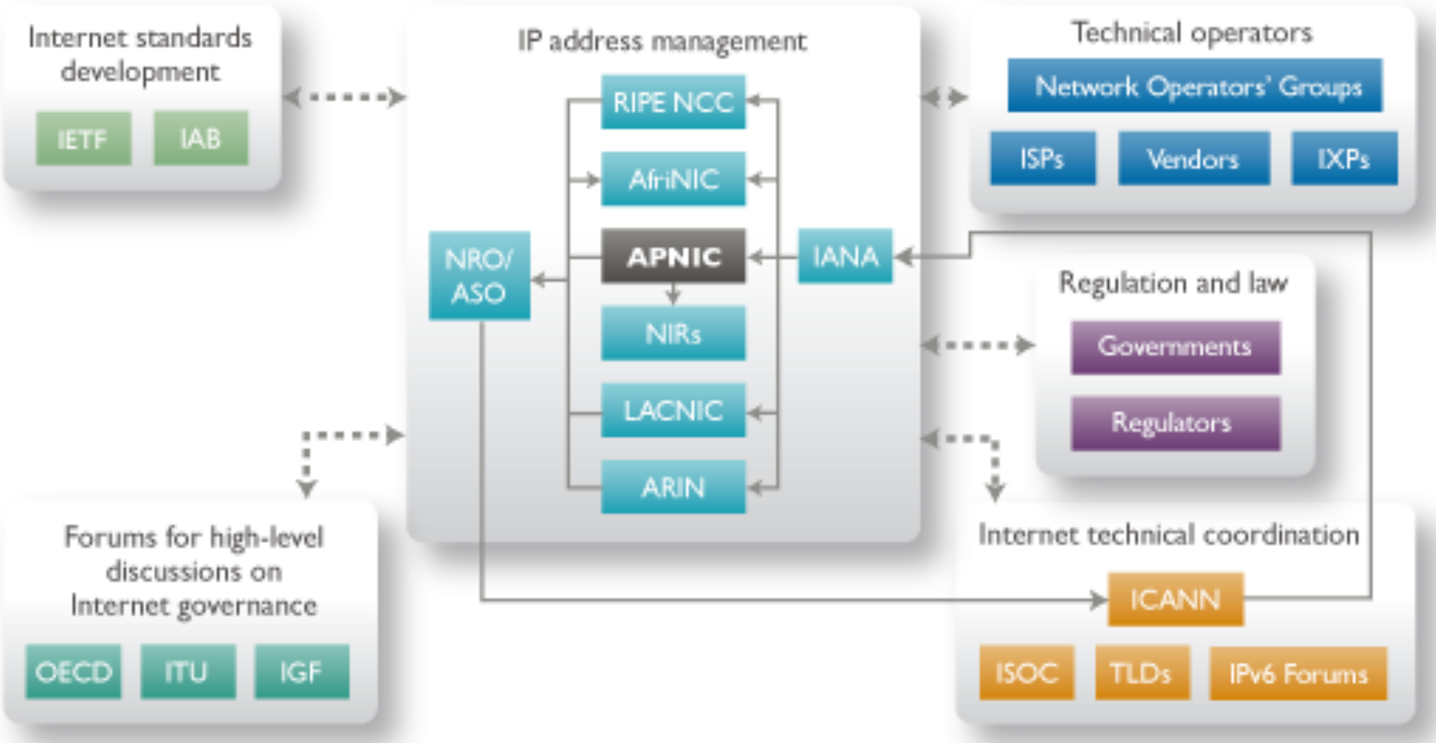
- The Regional Internet Registries have been charged with the following goals for the number resources they are responsible for:
  - Conservation
  - Aggregation
  - Registration

Where Are The RIR Regions?

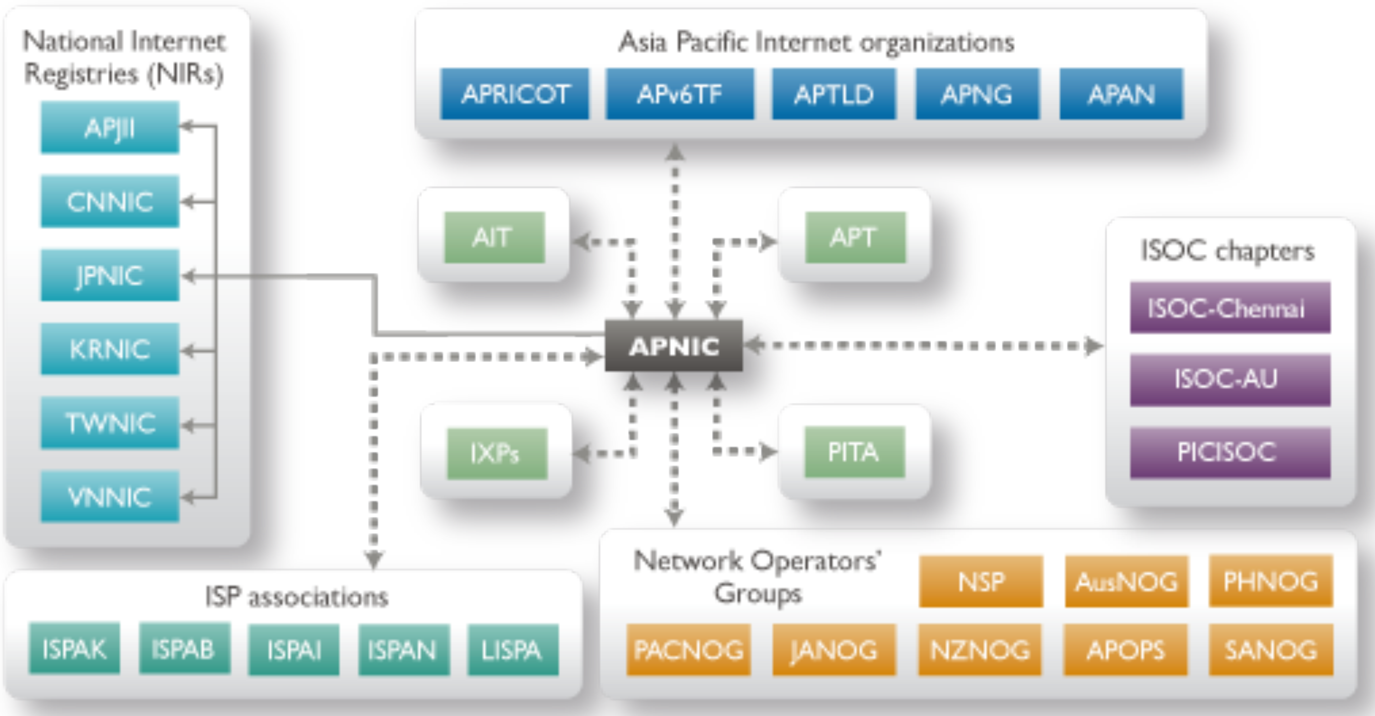
Internet Registry Structure



APNIC from a Global Perspective



APNIC in the Asia Pacific



Global Policy Coordination

The main aims of the NRO:

- To protect the unallocated number resource pool
- To promote and protect the bottom-up policy development process
- To facilitate the joint coordination of activities e.g., engineering projects
- To act as a focal point for Internet community input into the RIR system

Global Policy Coordination

The main function of ASO:

- ASO receives global policies and policy process details from the NRO
- ASO forwards global policies and policy process details to ICANN board

Introduction to APNIC

APNIC

What is APNIC?

- Regional Internet Registry (RIR) for the Asia Pacific region
  - One of five RIRs currently operating around the world
  - Non-profit, membership organisation
- Industry self-regulatory body
  - Open
  - Consensus-based
  - Transparent
- Meetings and mailing lists
  - http://meetings.apnic.net
  - http://www.apnic.net/mailing-lists

History of APNIC

- 1993
  - APNIC was established as a project of the Asia Pacific Networking Group (APNG)
- 1994
  - IANA authorized APNIC to commence allocating resources in its region
- 1995
  - Inaugural APNIC meeting in Bangkok
- 1998
  - APNIC relocated from Tokyo to Brisbane
- 2000
  - First independently-held three day Open Policy Meeting
- 2002
  - Introduced the Member Services Helpdesk with extended operating hours

What does APNIC do?

Resource service

- IPv4, IPv6, ASNs
- Reverse DNS delegation
- Resource registration
  - Authoritative registration server
  - Whois
  - IRR

Policy development

- Facilitating the policy development process
- Implementing policy changes

Information dissemination

- APNIC meetings
- Web and ftp site
- Publications, mailing lists
- Outreach seminars

Training

- Face to Face Training and Workshops
- eLearning

- Subsidised for members

Where is the APNIC region?

http://w

APNIC is NOT

- A network operator
  - Does not provide networking services
    - Works closely with APRICOT forum
- A standards body
  - Does not develop technical standards
    - Works within IETF in relevant areas (IPv6 etc)
- A domain name registry or registrar
    - Will refer queries to relevant parties

Managing Internet Resources

APNIC

Internet Resource Management Objectives

Conservation
- Efficient use of resources
- Based on demonstrated need

Aggregation
- Limit routing table growth
- Support provider-based routing

Registration
- Ensure uniqueness
- Facilitate trouble shooting

Uniqueness, fairness and consistency

IPv6 Allocation and Assignment

APNIC
Allocates
to APNIC Member

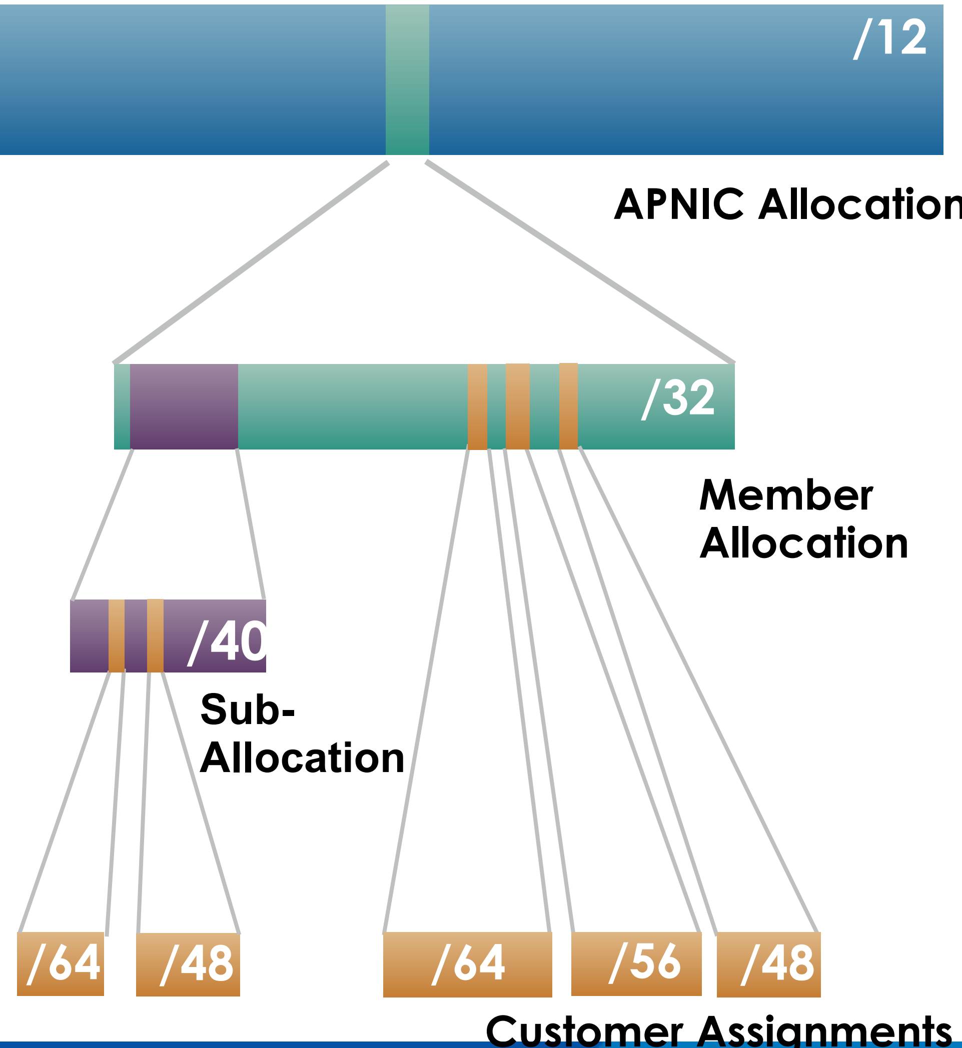
APNIC Member

Allocates
to downstream

Downstream
Assigns
to end-user

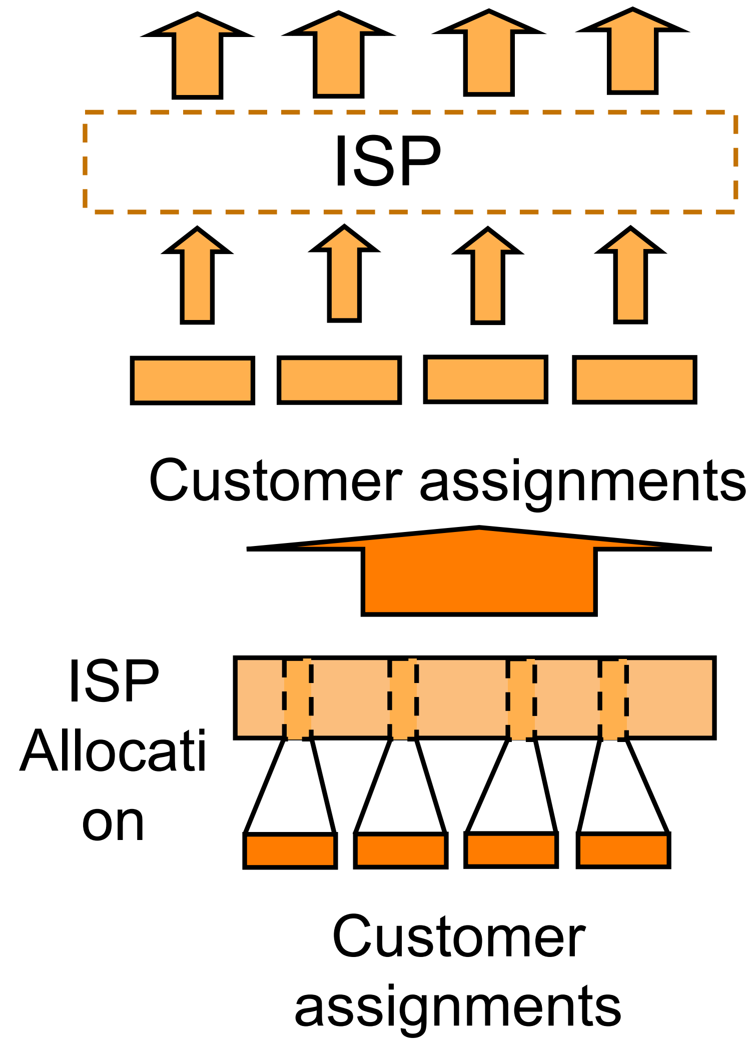
Assigns
to end-user

Customer / End User

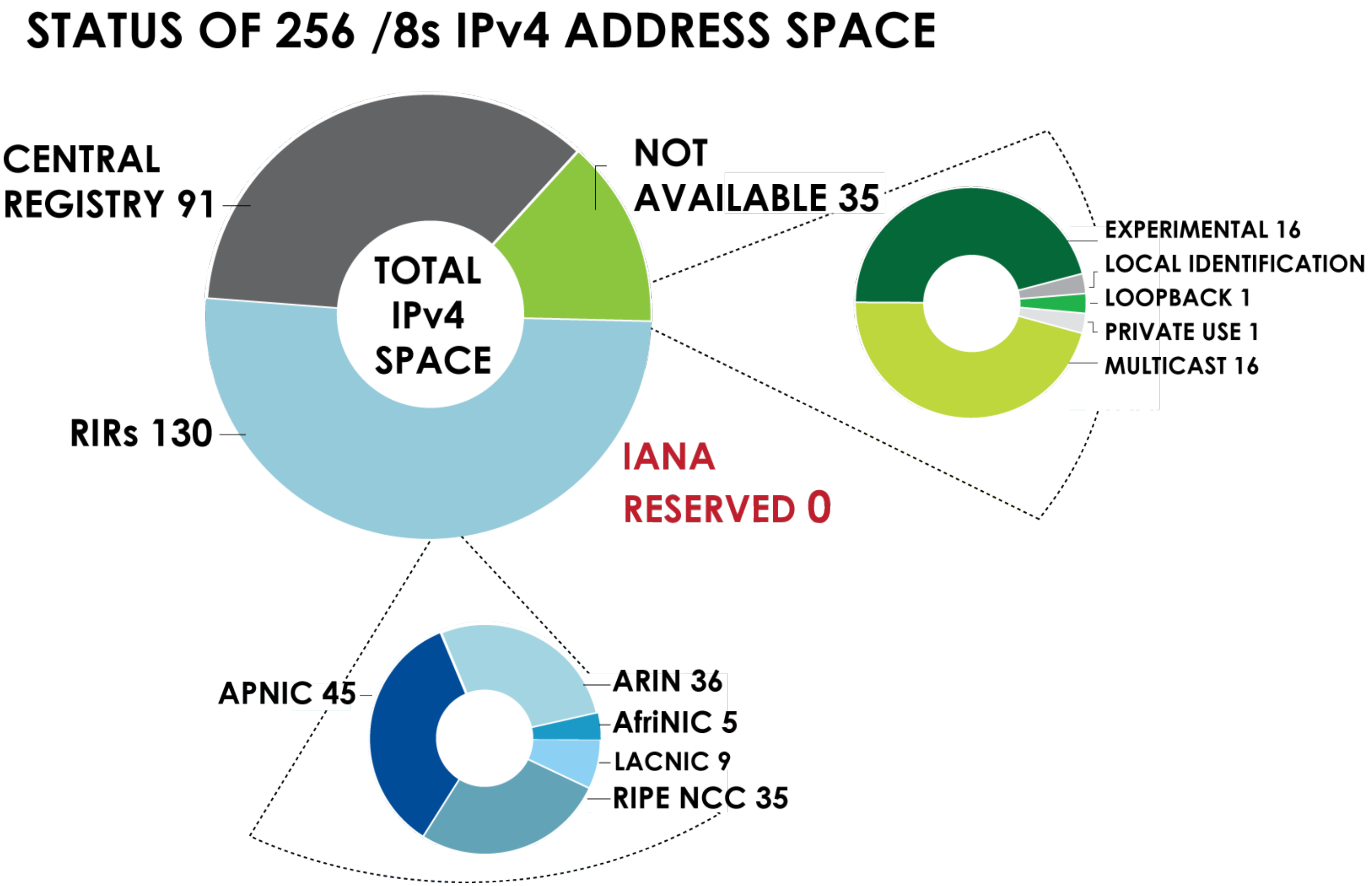


Portable and Non-Portable

- Portable Assignments
  - Customer addresses independent from ISP
  - Keeps addresses when changing ISP
  - Bad for size of routing tables
  - Bad for QoS: routes may be filtered, flap-dampened
- Non-portable Assignments
  - Customer uses ISP’s address space
  - Must renumber if changing ISP
  - Only way to effectively scale the Internet
- Portable allocations
  - Allocations made by APNIC/NIRs

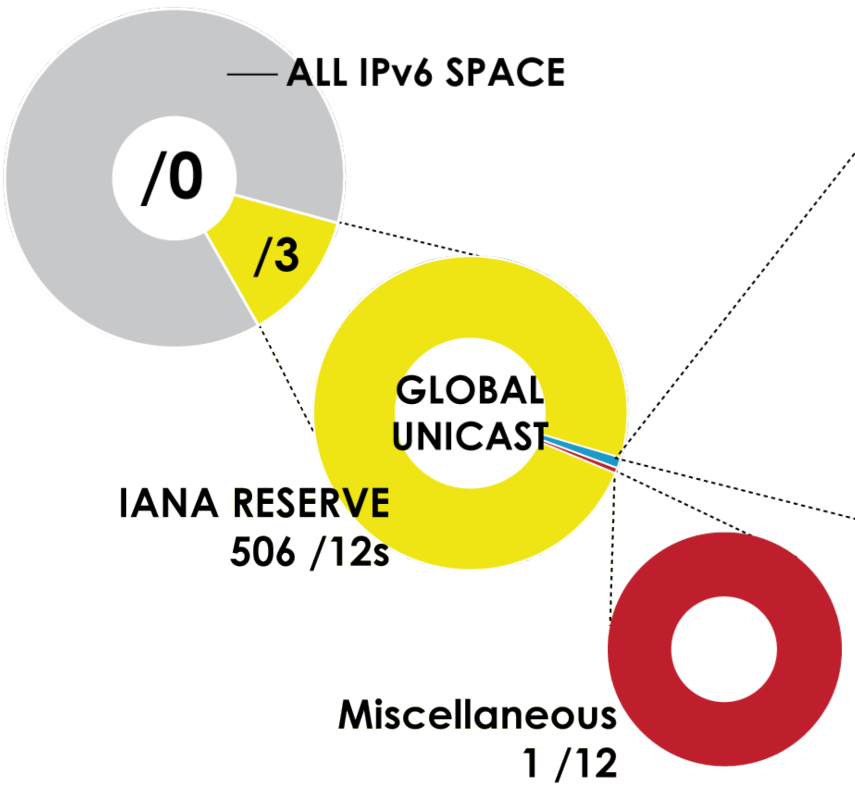


IPv4 Address Space



March 2011 - NRO

IPv6 Address Space



RIRs 5 /12s (October 2006)

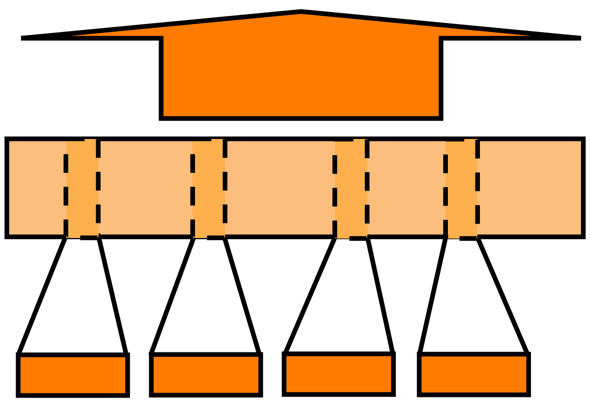
RIR	IPv6 ADDRESS
AfriNIC	2C00:0000::/12
APNIC	2400:0000::/12
ARIN	2600:0000::/12
LACNIC	2800:0000::/12
RIPE NCC	2A00:0000::/12

Aggregation and Portability

Aggregation

BGP Announcement (1)

ISP Allocation

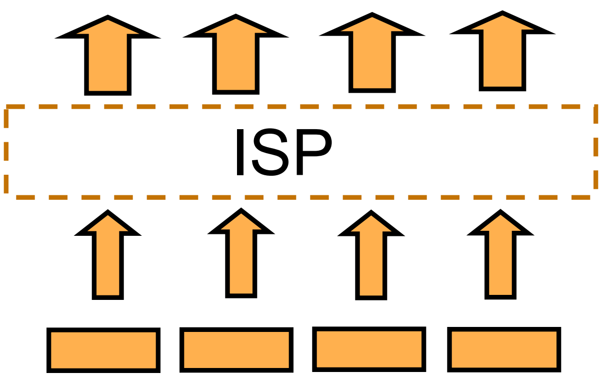


Customer assignments

(non-portable assignments)

No aggregation

BGP Announcements (4)

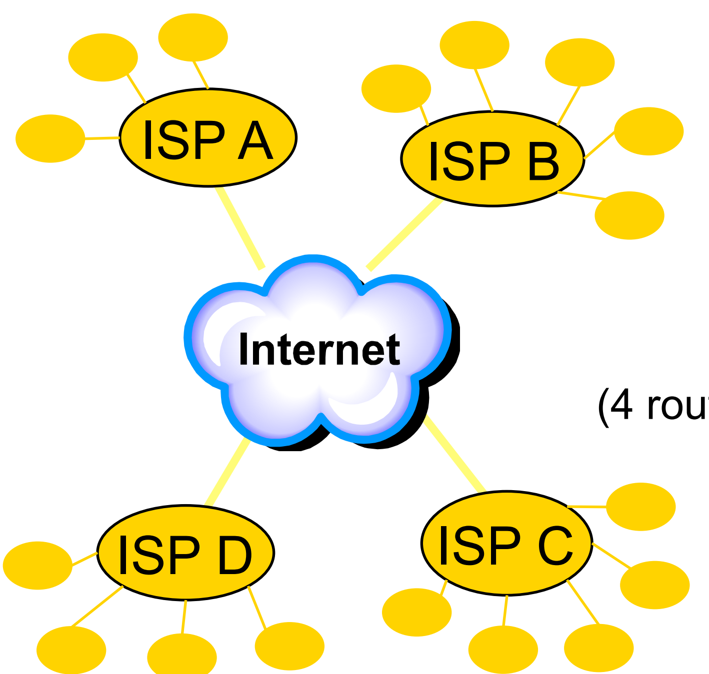


Customer assignments

(portable assignments)

Aggregation and Portability

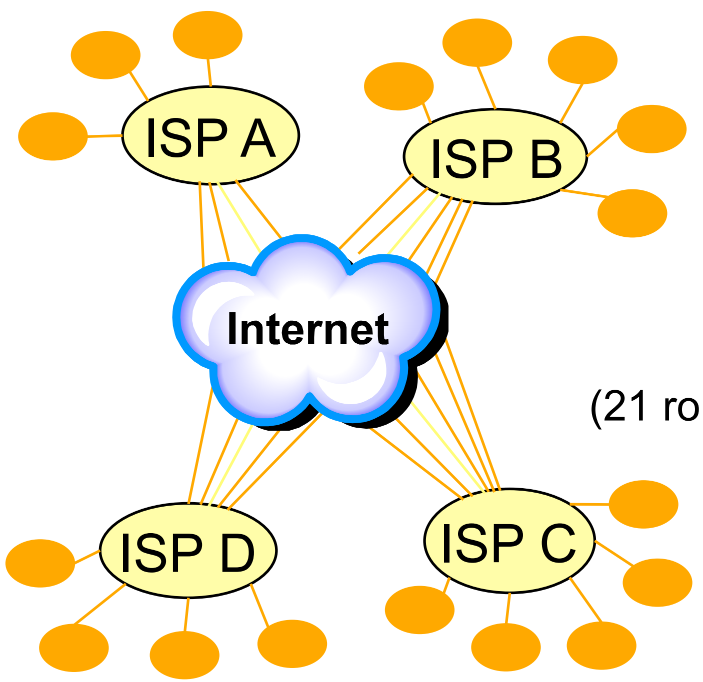
Aggregation



(4 routes)

(non-portable assignments)

No aggregation

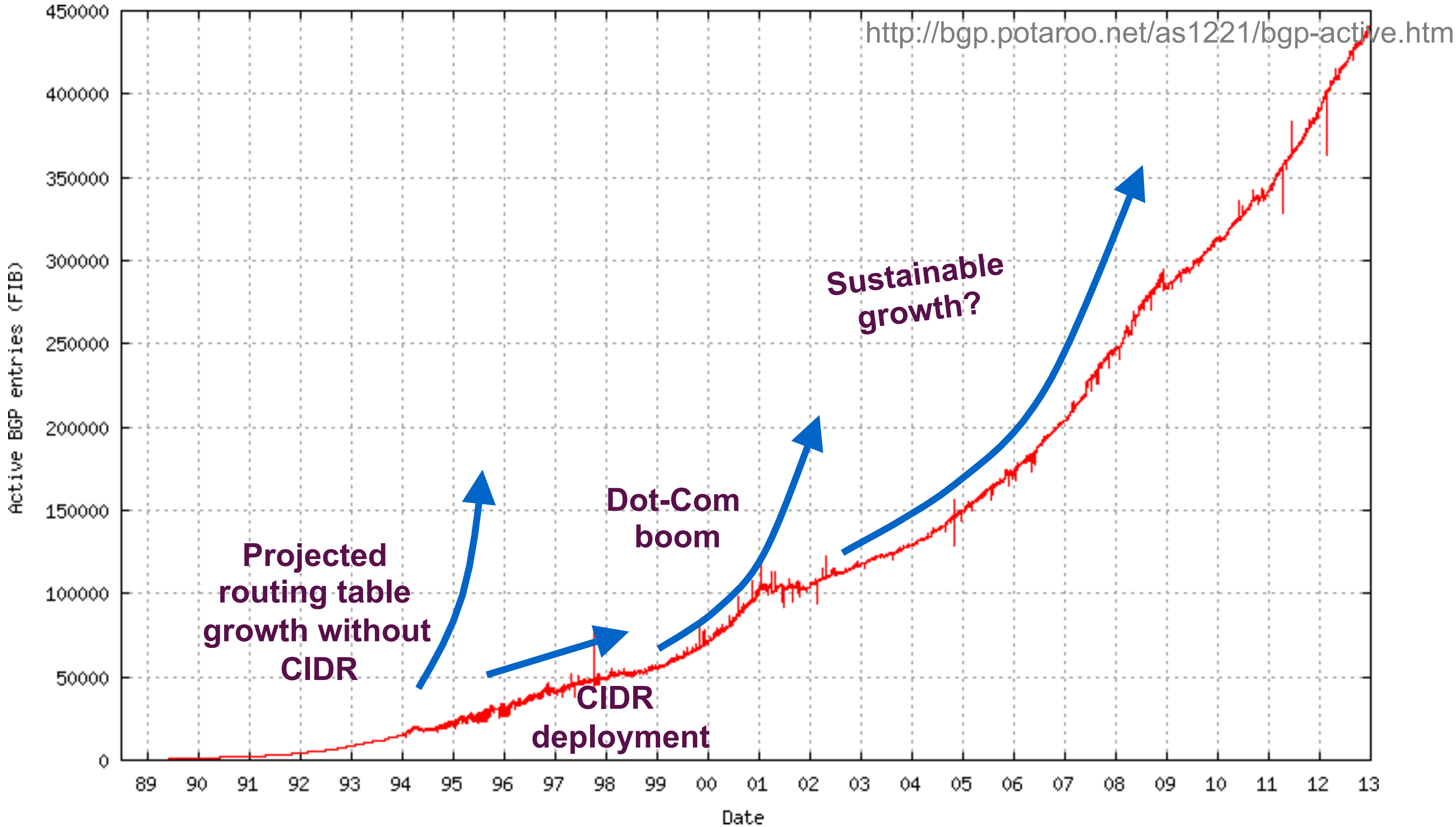


(21 routes)

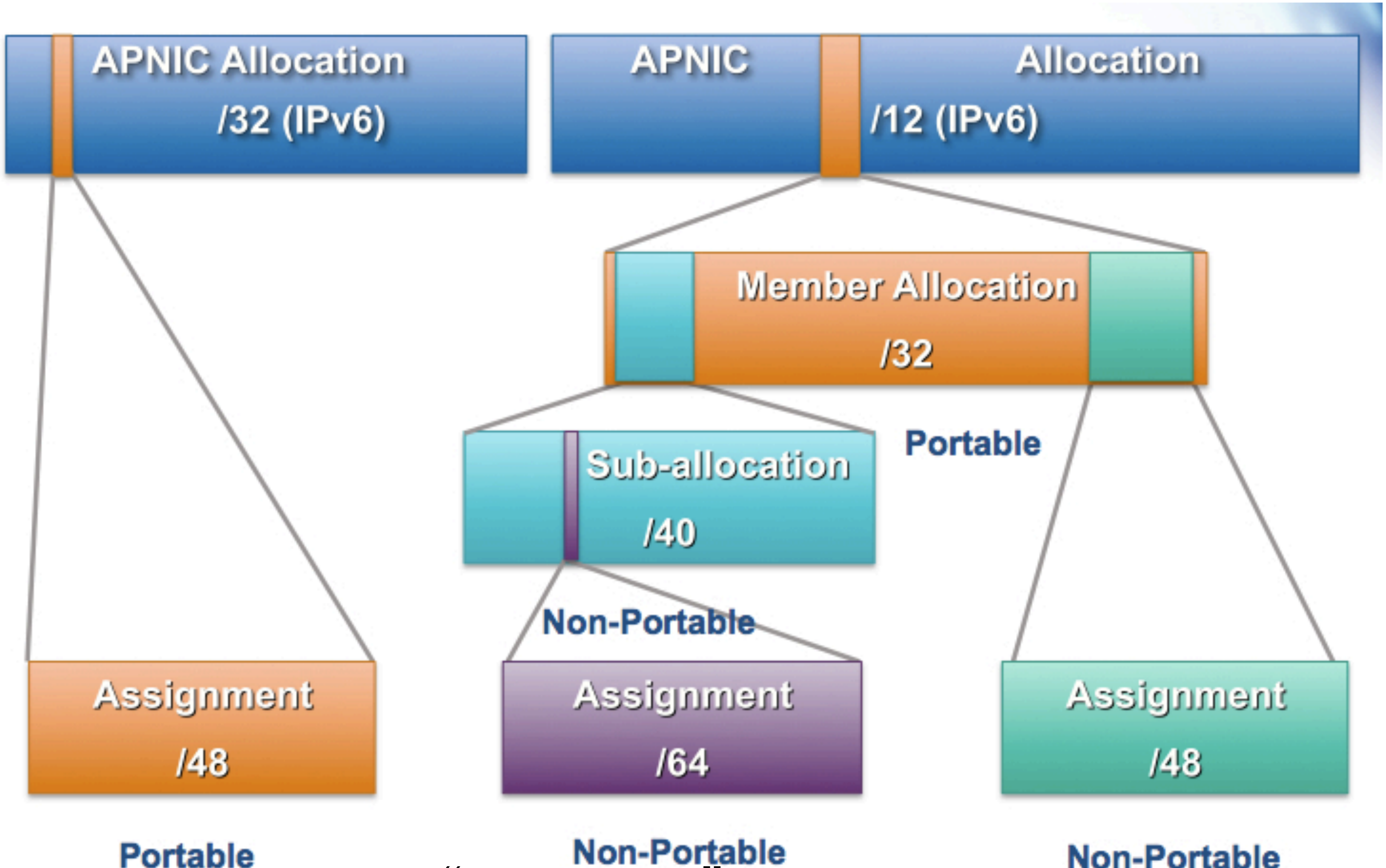
(portable assignments)

Growth of the Global Routing Table

441017 prefixes
As of 03 Jan 2013



Address Management Hierarchy



- Describes "portability" of the address space

Policy Development Process

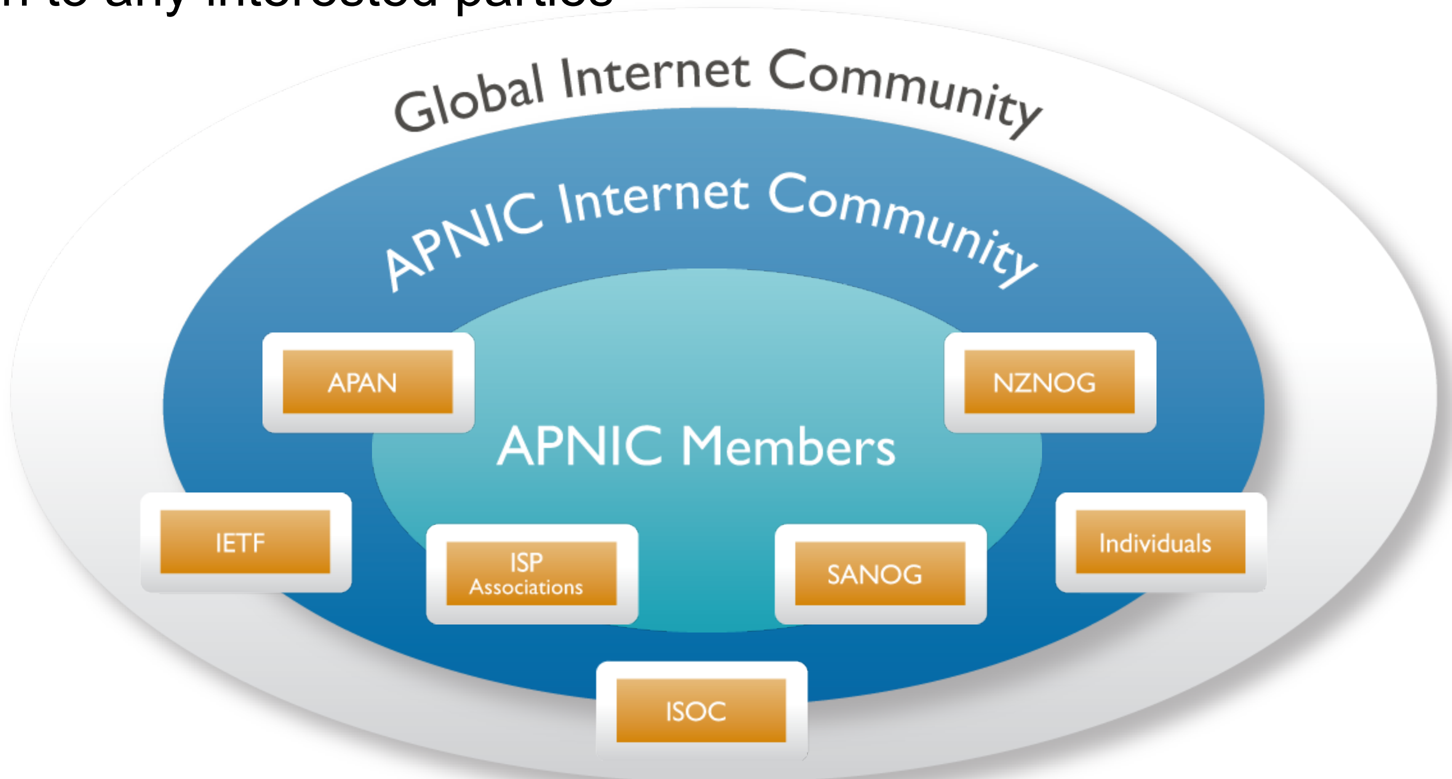
APNIC

Policies and their Development

- Policies are constantly changing to meet the technical needs of the Internet
- There is a system in place called the Policy Development Process
  - Anyone can participate
  - Anyone can propose a policy
  - All decisions & policies documented & freely available to anyone

You are Part of the APNIC Community!

- Open forum in the Asia Pacific
  - Open to any interested parties



A voice in regional Internet operations through participation in APNIC

Policy Development Process

OPEN

Anyone can participate

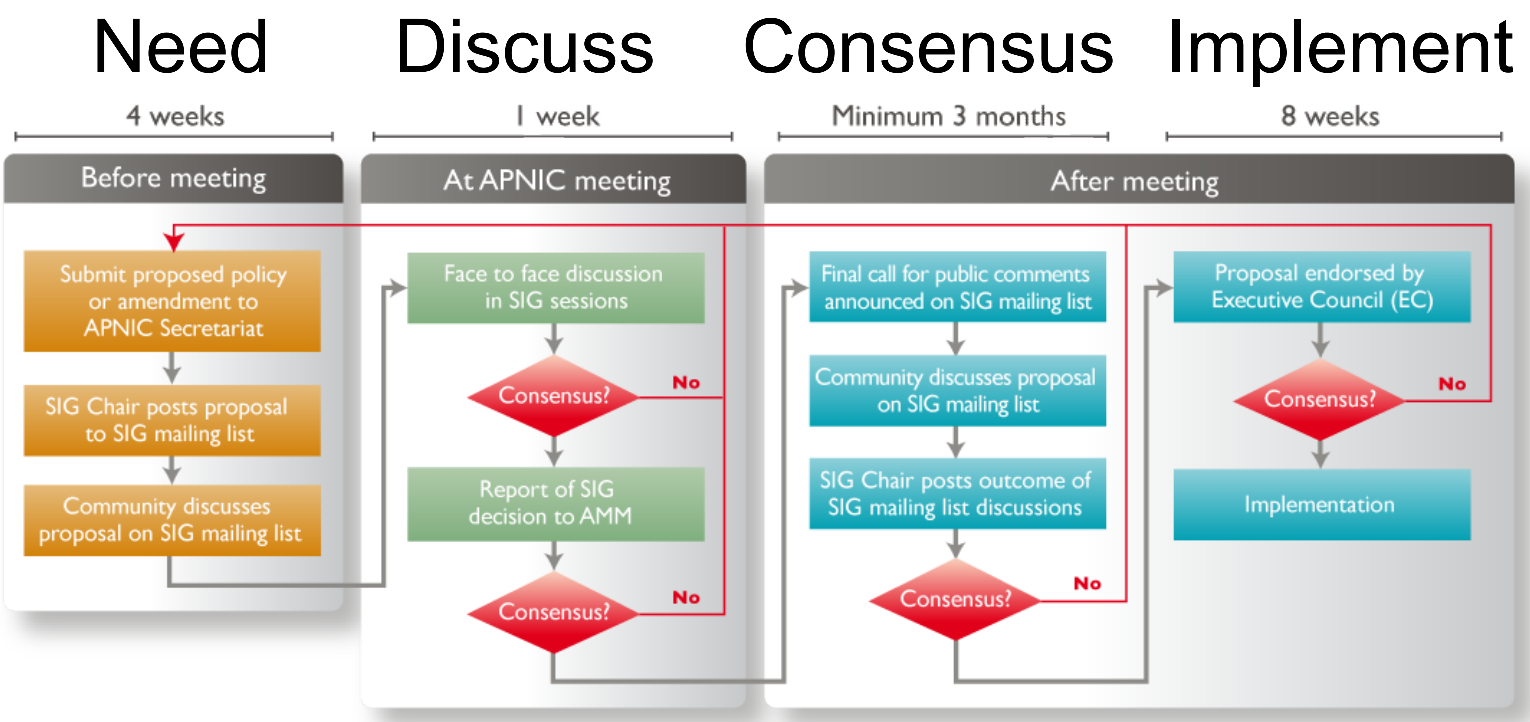
BOTTOM UP

Internet community proposes and approves policy

TRANSPARENT

All decisions & policies documented & freely available to anyone

Policy Development Process



You can participate!

More information about policy development can be found at:

http://www.apnic.net/policy

Why Participate?

- You are part of the Community
  - APNIC policies are developed by the membership and broader Internet community
- Knowing and understanding the policies are important for your organization
  - This is your chance to comment on policies that may directly affect you
- Opportunity to learn and share experiences

How to Participate

- Joining APNIC conferences and meetings
- You can participate further

Ask questions and clarify points

Make your voice heard

Vote

- Attend remotely
  - Video, audio, text streaming, chat
- Trainings, seminars and outreach events
- Join the discussion in the mailing list

From Regional to Global Policies

While RIRs and their respective communities are responsible for policies specific to their regions, there are times when a policy needs to be global.

Global Policy Coordination

Supporting Internet Development

Projects - Root Server Deployment

- A number of mirrored root server sites have been placed into the Asia Pacific region
- Lowers the transit cost by using a nearby instance of a root server
- The sites are partially or fully funded by APNIC, but operate as "anycast" mirror copies of existing Root servers, by the applicable root server operator

Grants For Community Support

- The Information Society Innovation Fund is a small grants program funding innovative approaches to the extension of Internet infrastructure and services in the Asia Pacific region

isif asia

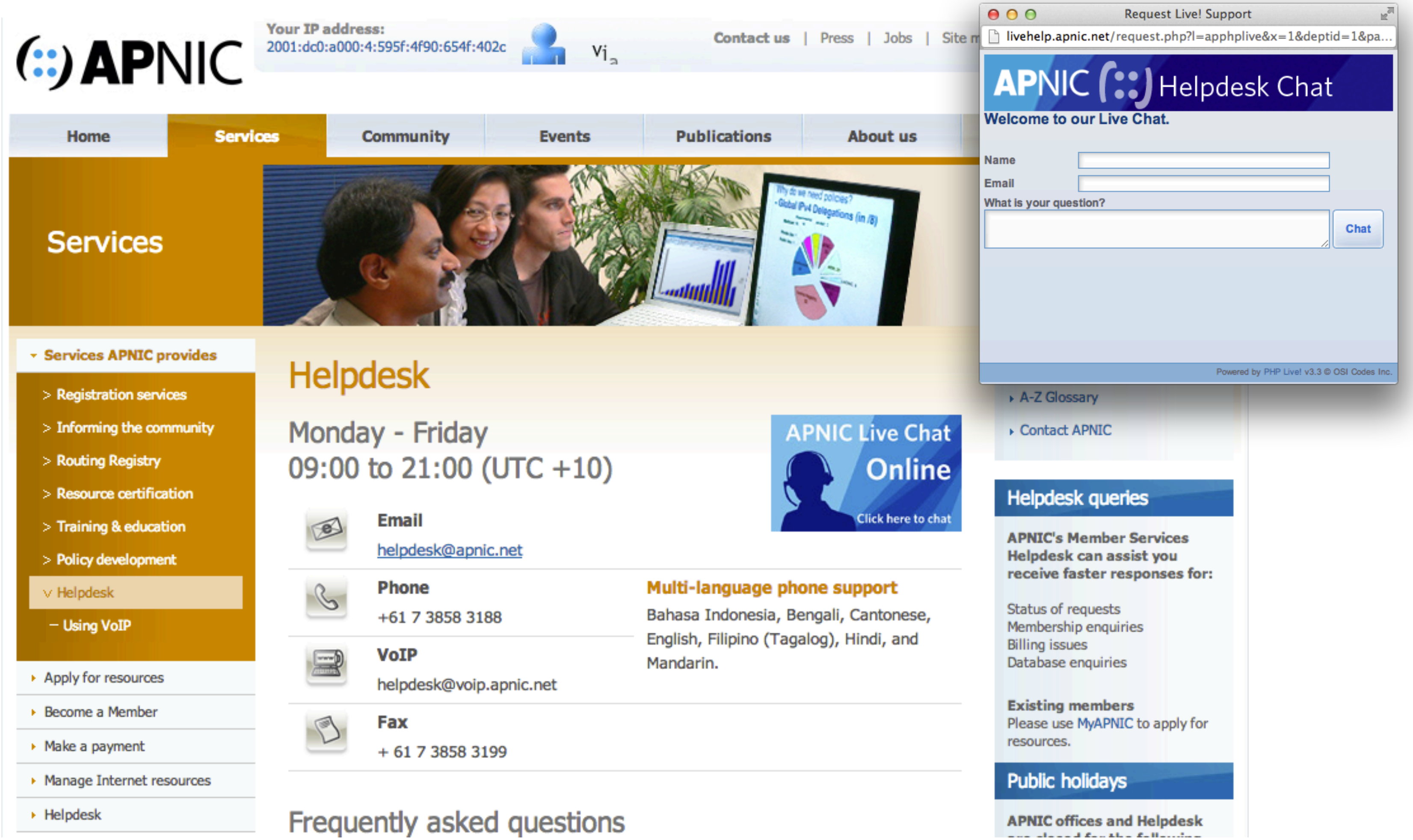
IPv6 Program

- Monitor: IPv6 technical development and BCP, deployment statistics, and challenges and solutions
- Outreach: Share timely, useful and customised information on IPv6 with the Internet stakeholders (network operators, content providers, content distribution networks, software developers, governments and inter-governmental organizations, civil society etc.)
- Facilitate: Encourage proactive communication and discussion among intra/inter Internet stakeholders on IPv6 deployment
- Assist: REAL and TANGIBLE IPv6 deployment

APNIC Labs

- IPv6 measurement
  - http://labs.apnic.net/ipv6-measurement/
- Resource Certification / RPKI

APNIC Helpdesk Chat



Introduction to Internet Protocols and Operations

APNIC

What is a Protocol?

- Set of rules that define the communications process
- defines the structure or pattern for the data transferred
  - functions or processes that need to be carried out in order to implement the data exchange
  - information required by processes in order for them to accomplish this
- All data is transmitted in the same way irrespective of what the data refers to, whether it is clear or encrypted.

The OSI Model

Application	Access to the network
Presentation	Manipulate data (Translate, encrypt)
Session	Manage sessions (connections)
Transport	Provide reliable delivery
Network	Internetwork - move packets from source to destination
Data Link	Configure data for direct delivery by physical layer
Physical	Physical delivery - electrical specs etc

OSI and TCP/IP Model

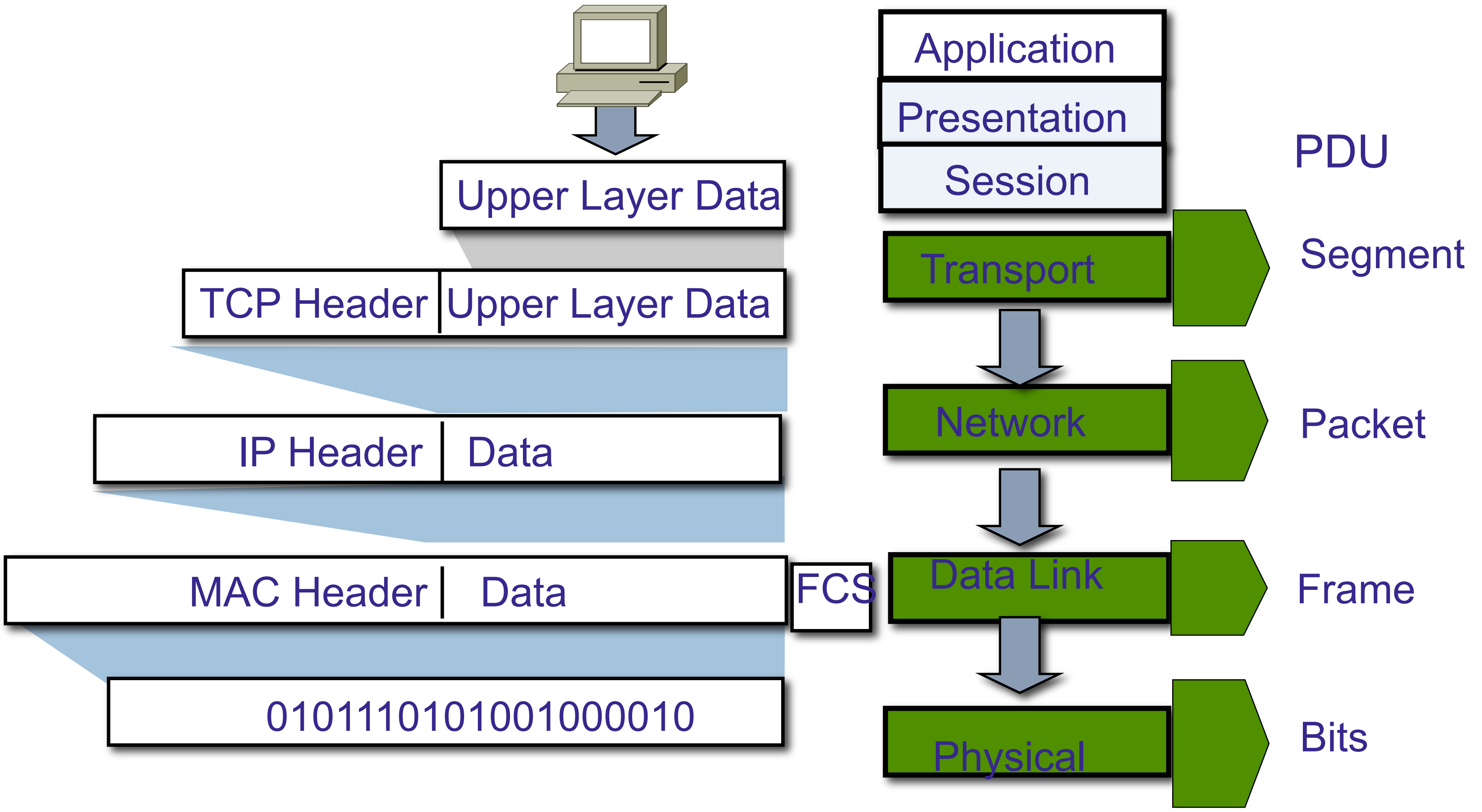
OSI Model

- Application
- Presentation
- Session
- Transport
- Network
- Data Link
- Physical

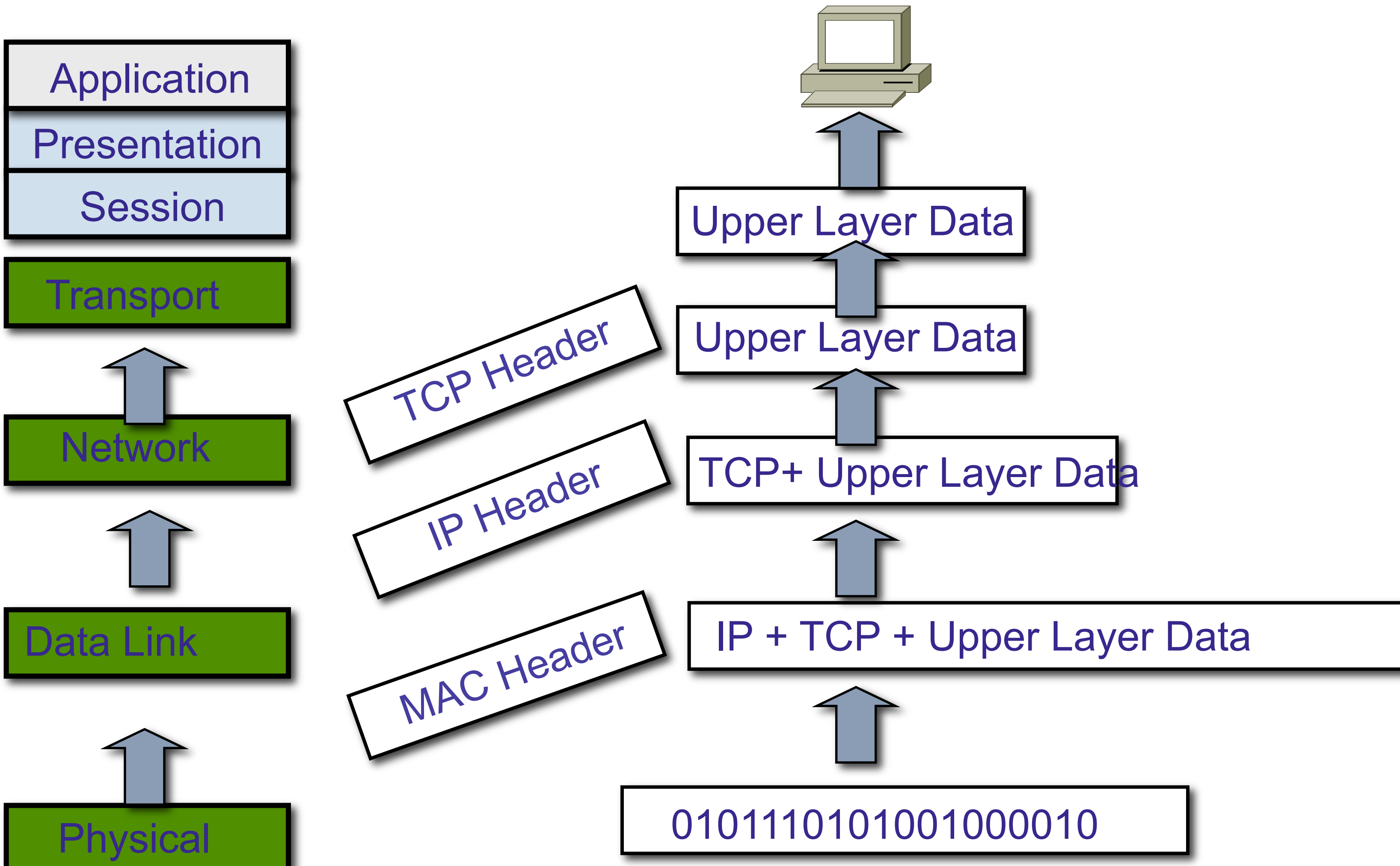
TCP/IP Model

- Application
- Transport
- Internet
- Network Access

Encapsulating Data



De-encapsulating Data

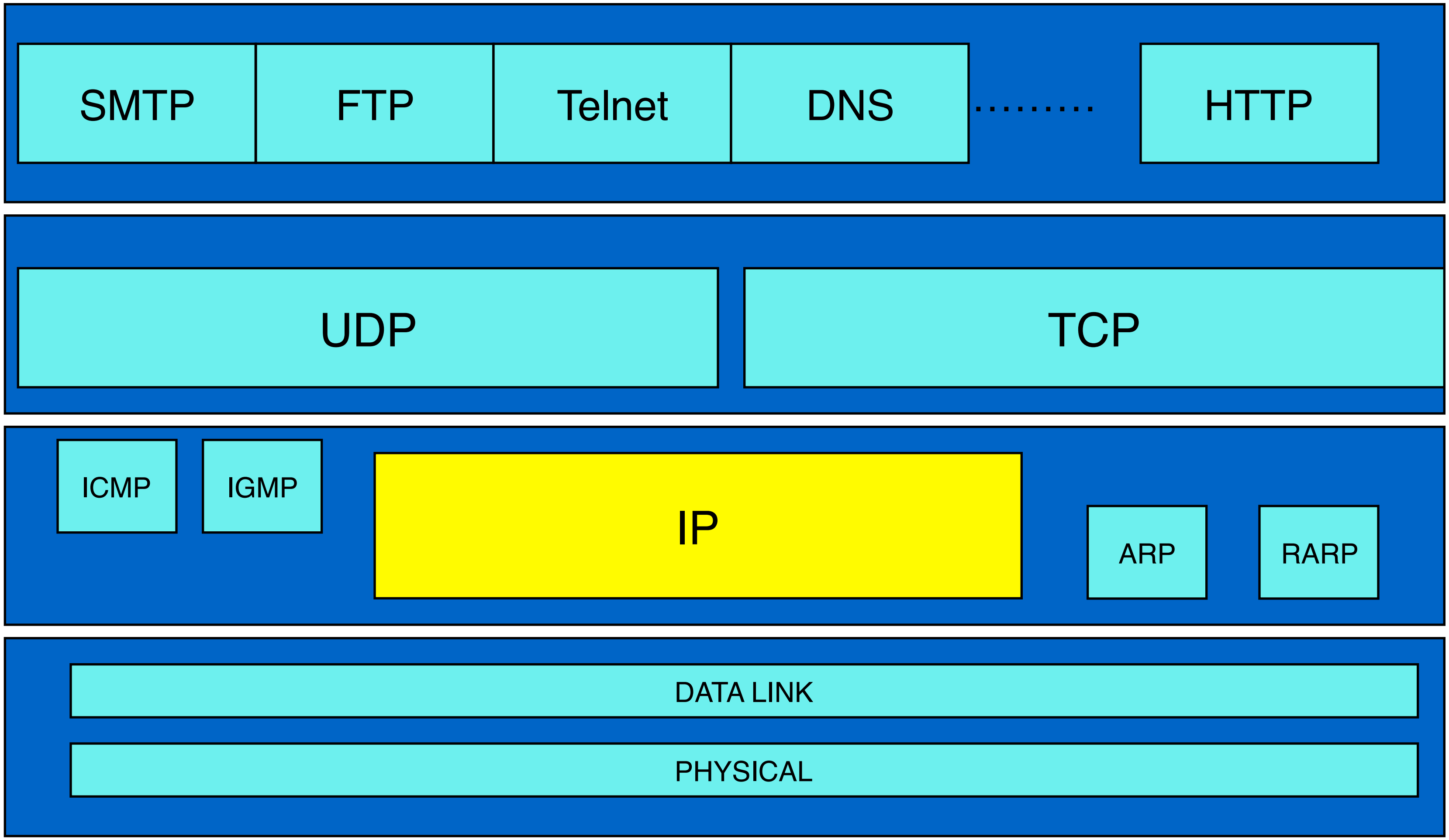


Source: www.cisco.com (ICND v1.0a—1-11)

Internet Protocol (IP)

- IP is an unreliable, connectionless delivery protocol
  - A best-effort delivery service
  - No error checking or tracking (no guarantees – Post Office)
  - Every packet treated independently
  - IP leaves higher level protocols to provide reliability services (if needed)
- IP provides three important definitions:
  - basic unit of data transfer
  - routing function
  - rules about delivery

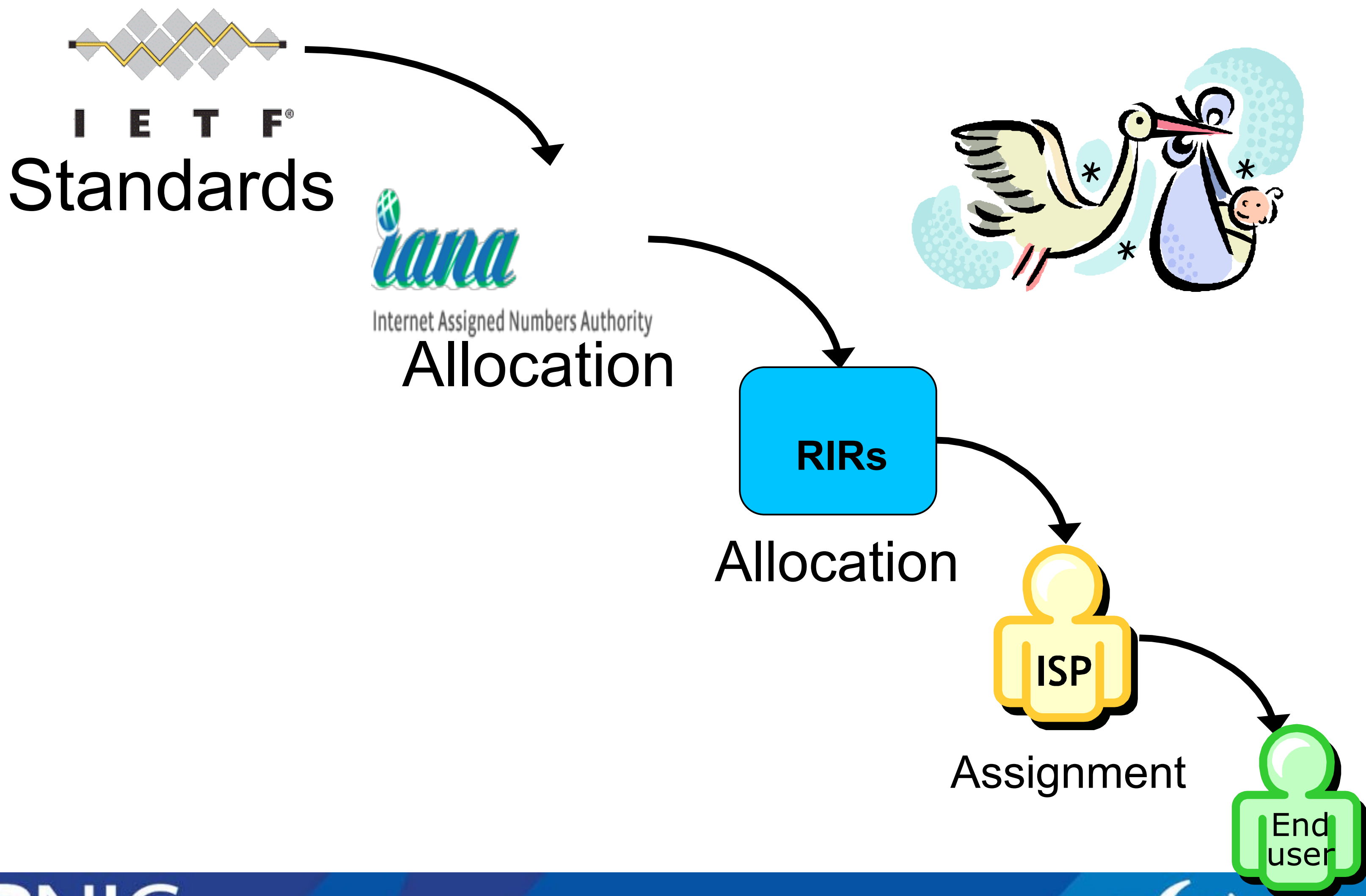
TCP/IP Protocol Structure



From Forouzan

IP Addressing Basics

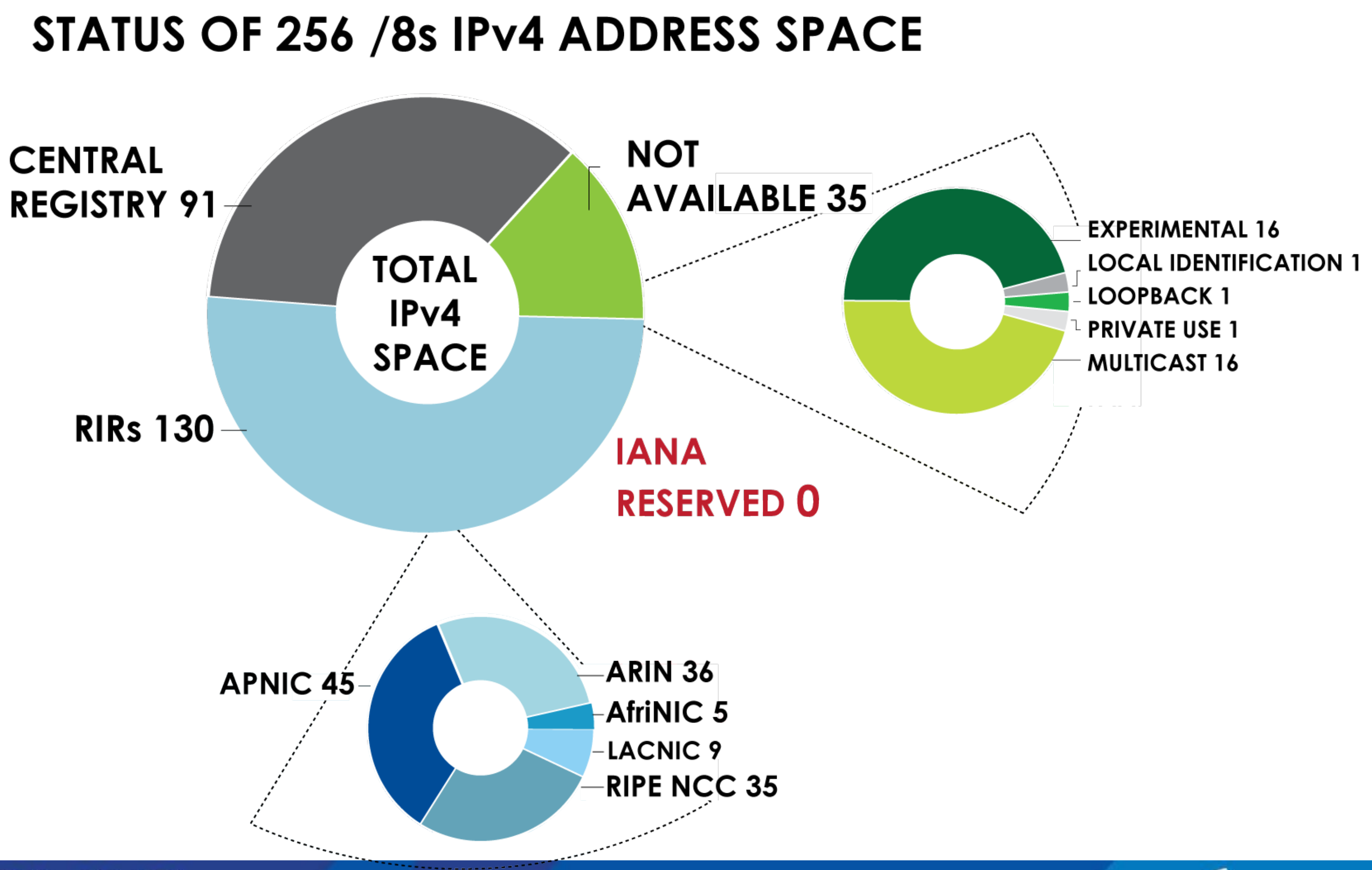
Where do IP addresses come from?



IP Addressing Issues

- Exhaustion of IPv4 addresses
  - Wasted address space in traditional subnetting
  - Limited availability of /8 subnets address
- Internet routing table growth
  - Size of the routing table due to higher number prefix announcement
- Tremendous growth of the Internet

How many IPv4 IANA pool available



IP Addressing Solutions

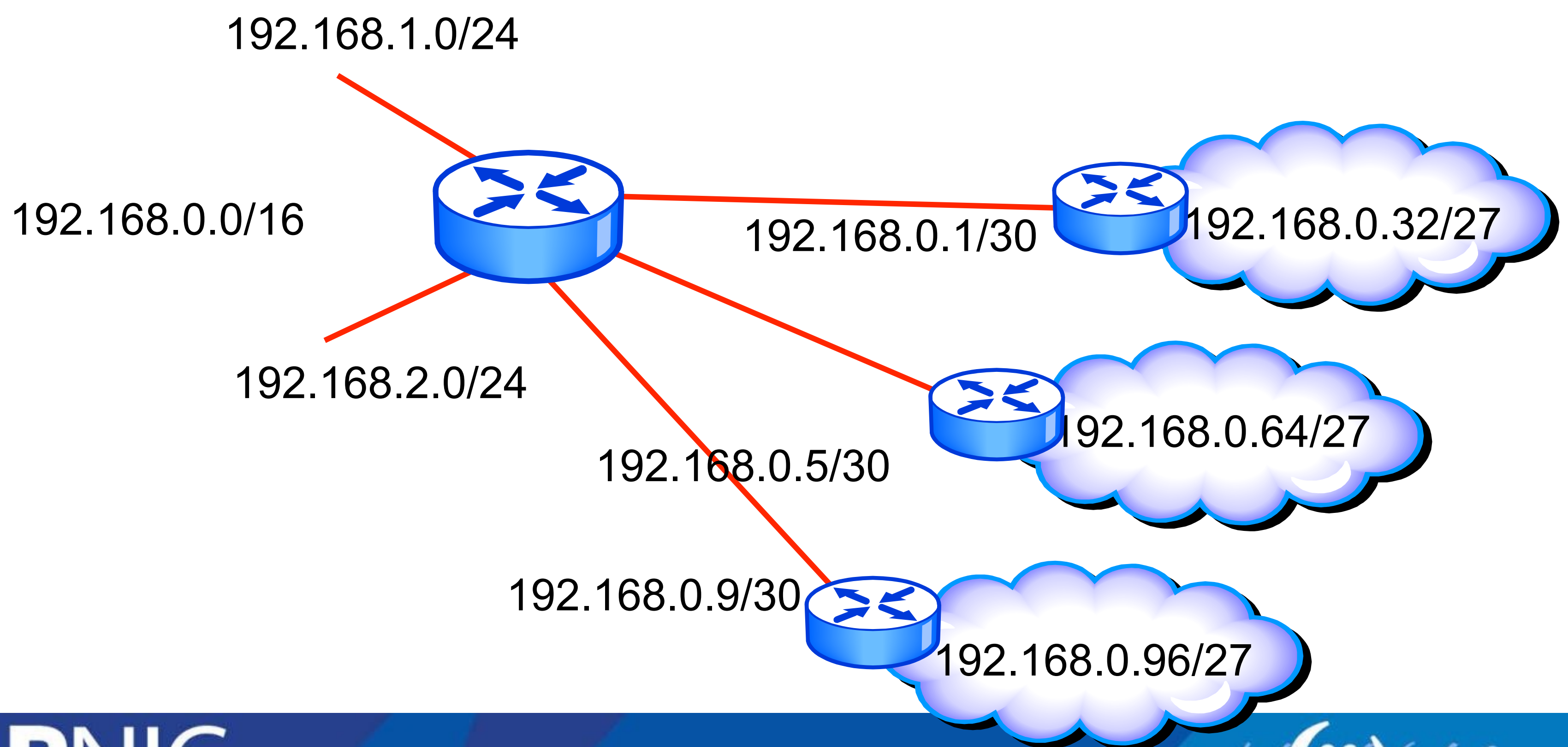
- Subnet masking and summarization
  - Variable-length subnet mask definition
  - Hierarchical addressing
  - Classless InterDomain Routing (CIDR)
  - Routes summarization (RFC 1518)
- Private address usage (RFC 1918)
  - Network address translation (NAT)
- Development of IPv6 address

Variable Length Subnet Mask (VLSM)

- Allows the ability to have more than one subnet mask within a network
- Allows re-subnetting
  - create sub-subnet network address
- Increase the routes capability
  - Addressing hierarchy
  - Summarisation

Calculating VLSM example

- Subnet 192.168.0.0/24 into smaller subnet
  - Subnet mask with /27 and /30 (point-to-point)



Calculating VLSM example (cont.)

- Subnet 192.168.0.0/24 into smaller subnet
  - Subnet mask with /30 (point-to-point)

Description	Decimal	Binary
Network Address	192.168.0.0/30	x.x.x.000000**00**
1st valid IP	192.168.0.1/30	x.x.x.000000**01**
2nd valid IP	192.168.0.2/30	x.x.x.000000**10**
Broadcast address	192.168.0.3/30	x.x.x.000000**11**

Calculating VLSM example (cont.)

- Subnet 192.168.0.0/24 into smaller subnet
  - Subnet mask with /27

Description	Decimal	Binary
Network Address	192.168.0.32/27	x.x.x.000**00000**
Valid IP range 192.168.0.33 - 192.168.0.62		x.x.x.000**00001**
		x.x.x.000**00010**
Broadcast address	192.168.0.63/30	x.x.x.000**11111**

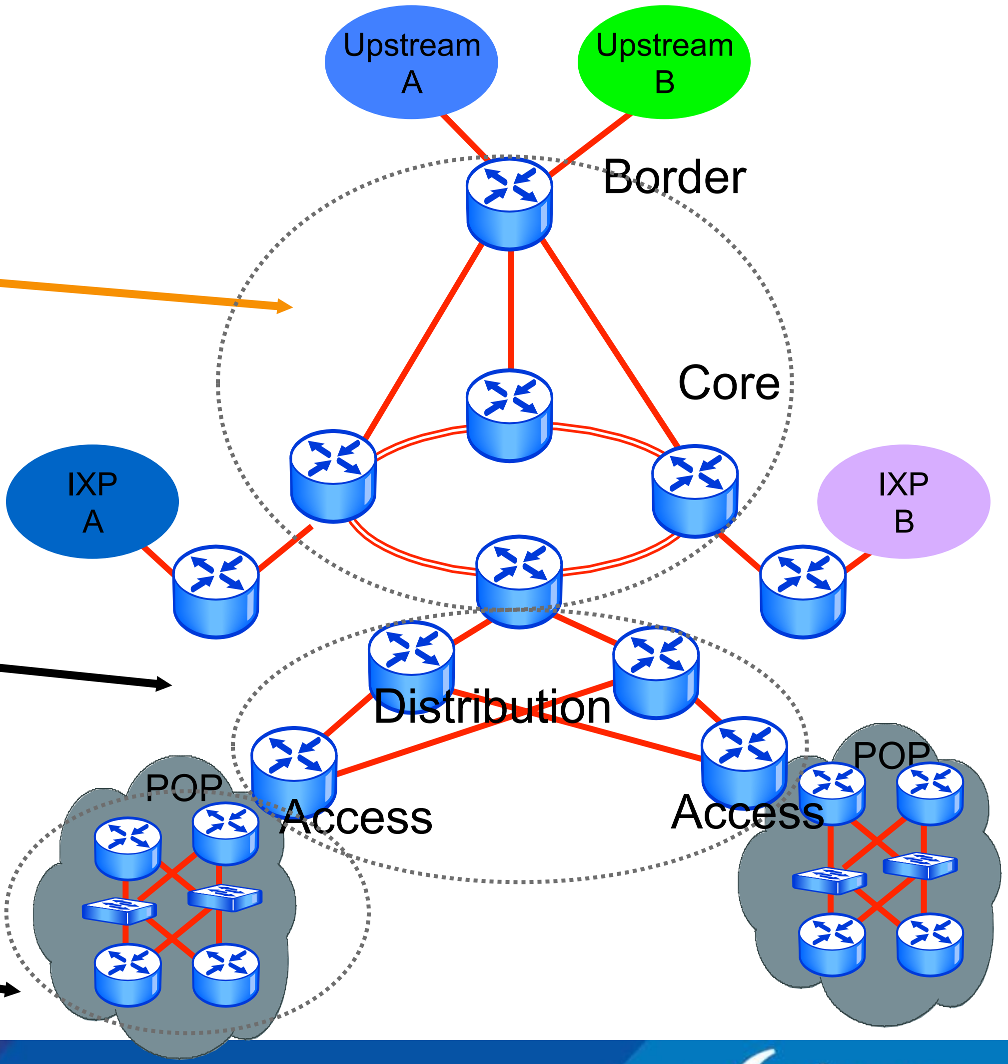
Addressing Hierarchy

Network Number
192.168.0.0/16

Core
192.168.32.0/19

Distribution/Core
192.168.32.0/21

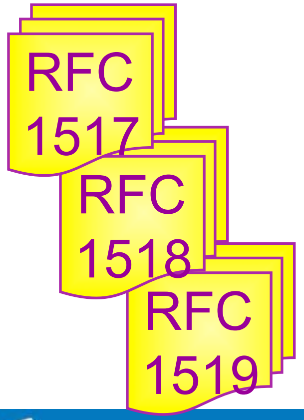
Access/Distribution
192.168.48.0/21



Classful and Classless

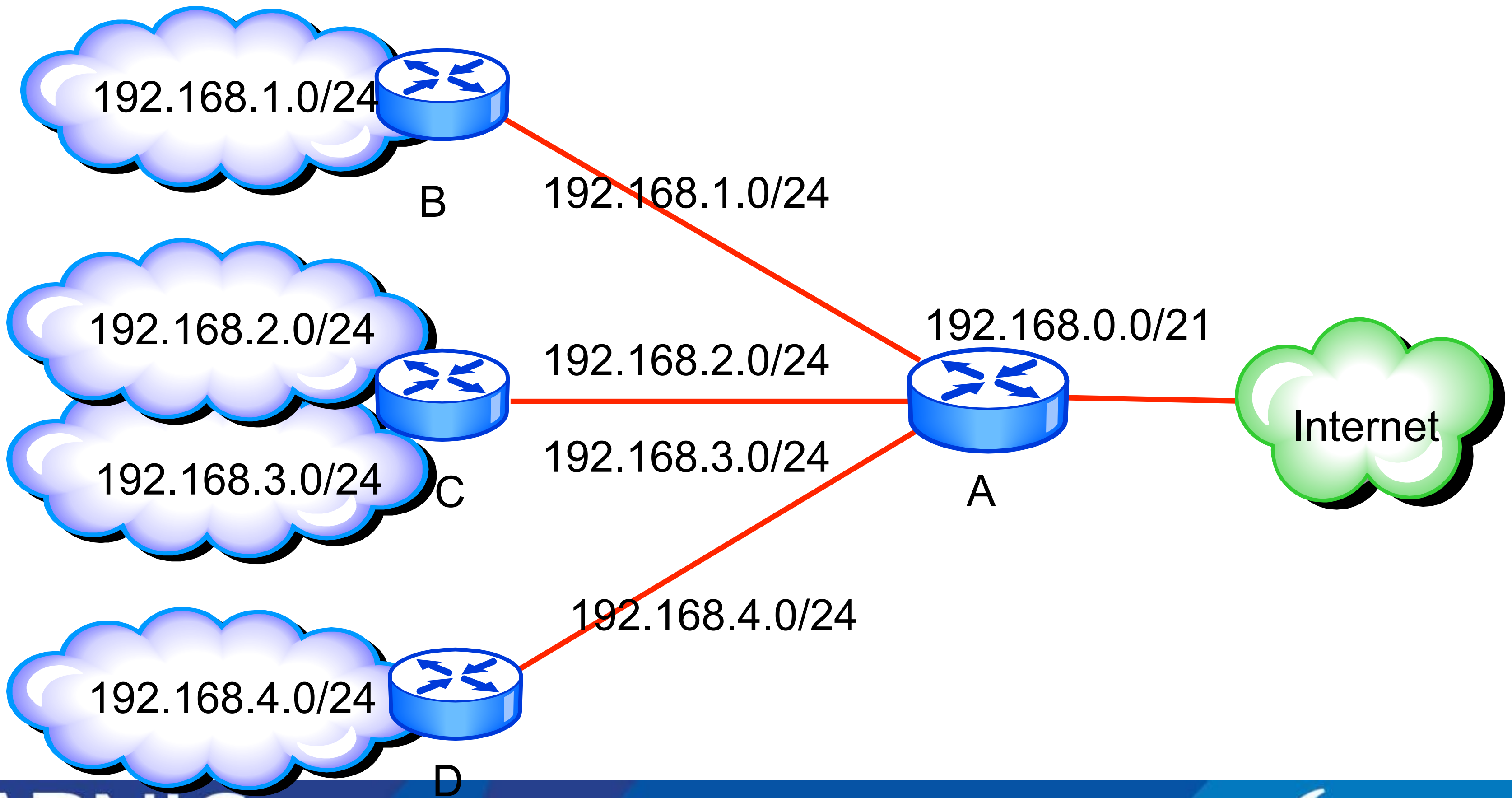
- Classful (Obsolete)
  - Wasteful address architecture
    - network boundaries are fixed at 8, 16 or 24 bits
    - (class A, B, and C)
- Classless
  - Efficient architecture
    - network boundaries may occur at any bit
    - (e.g. /12, /16, /19, /24 etc)
- CIDR
    - Classless Inter Domain Routing architecture
  - Allows aggregation of routes within ISPs infrastructure

Best Current Practice



Prefix Routing / CIDR

- CIDR offers the advantages reducing the routing table size of the network by summarising the ISP announcement in a single /21 advertisement



Route Summarisation

- Allows the presentation of a series of networks in a single summary address.
- Advantages:
  - Faster convergence
  - Reducing the size of the routing table
  - Simplification
  - Hiding Network Changes
  - Isolate topology changes

AS Numbers

APNIC

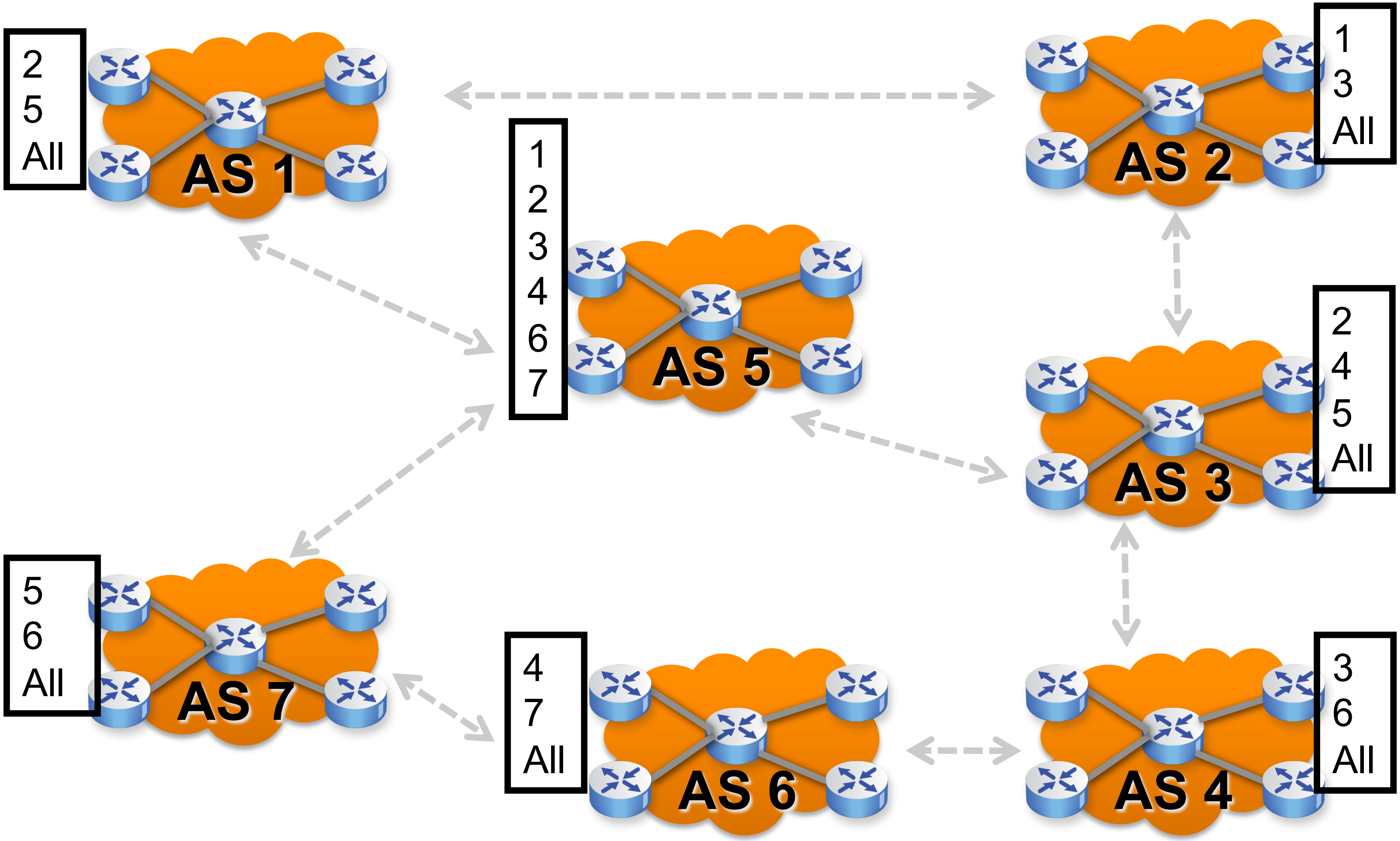
What is an Autonomous System Number?

- Autonomous System Numbers (ASNs) are globally unique identifiers for IP networks
- ASNs are allocated to each Autonomous System (AS) for use in BGP routing
- AS numbers are important because the ASN uniquely identifies each network on the Internet

What Is An Autonomous System?

- Group of Internet Protocol-based networks with the same routing policy
- Usually under single ownership, trust or administrative control
- The AS is used both in the exchange of exterior routing information (between neighboring ASes) and as an identifier of the AS itself

How Do Autonomous Systems Work?



When Do I Need An ASN?

- An ASN is needed if you have a
  - Multi-homed network to different providers AND
  - Routing policy different to external peers

  - * For more information please refer to RFC1930: Guidelines for creation, selection and registration of an Autonomous System

Requesting an AS Number

- If a member requests an ASN from APNIC for own network infrastructure
  - AS number is “portable”
- If a member requests an ASN from APNIC for its downstream customer network
  - ASN is “non-portable”
  - ASN is returned if the customer changes provider
- Current Distribution
  - Previously 2 byte ASN (16 bits) runs into possibility of exhaustion
  - Currently 4 byte ASN distribution policy 32 bits
  - 2 byte ASN on request with documented justification

Aut-num Object Example

```
aut-num:     AS4777
as-name:     APNIC-NSPIXP2-AS
Descr:       Asia Pacific Network Information Centre
descr:       AS for NSPIXP2, remote facilities site
import:      from AS2500 action pref=100; accept ANY
import:      from AS2524 action pref=100; accept ANY
import:      from AS2514 action pref=100; accept ANY
export:      to AS2500 announce AS4777
export:      to AS2524 announce AS4777
export:      to AS2514 announce AS4777
default:     to AS2500 action pref=100; networks ANY
admin-c:     PW35-AP
tech-c:      NO4-AP
remarks:     Filtering prefixes longer than
mnt-by:      MAINT-APNIC-AP
changed:     paulg@apnic.net 19981028
source:      APNIC
```

POLICY
RPSL

AS Number Representation

- 2-byte only AS number range : 0 – 65535
- 4-byte only AS number range – represented in two ways
  - AS PLAIN: 65,536 - 4,294,967,295
  - AS DOT: 1.0 - 65535.65535
- Usages
  - 0 and 65535 Reserved
  - 1 to 64495 Public Internet
  - 64496 to 64511 Documentation –RFC5398
  - 64512 to 65534 Private use
  - 23456 represent 32 Bit range in 16 bit world
  - 65536 to 65551 Documentation – RFC 5398
  - 65552 to 4294967295 Public Internet

AS PLAIN

- IETF preferred standard notation RFC5396
- Continuation on how a 2-Byte AS number has been represented historically
- Notation: The 32 bit binary AS number is translated into a single decimal value
  - Example: AS 65546
- Total AS Plain range:
  2 byte: 0 – 65535 (original 16-bit range)
  4 byte: 65,536 - 4,294,967,295 (RFC4893)
  - APNIC region uses the AS PLAIN style of numbering

AS DOT

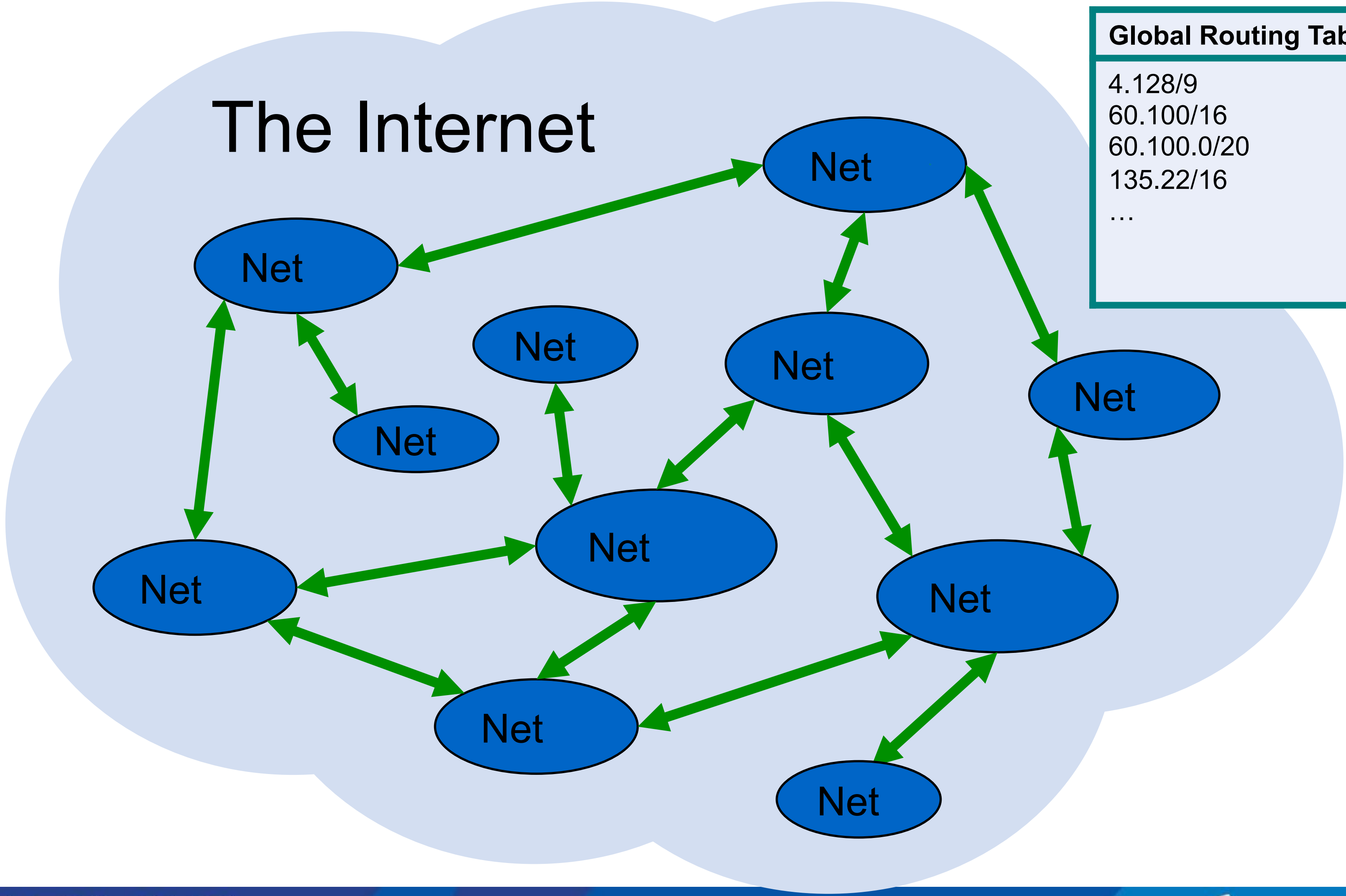
- Based upon 2-Byte AS representation
  - <Higher2bytes in decimal> . <Lower2bytes in decimal>
    - For example: AS 65546 is represented as 1.10
  - Easy to read, however hard for regular expressions
  - There is a meta character "." in regular expression
    - For example, a.c matches "abc", etc., but [a.c] matches only "a", "32 bit AS number representation
- Example: AS PLAIN Converted to AS DOT
  - AS PLAIN: 131072 ~ 132095
  - AS DOT: 2.0 ~ 2.1023

16 bit and 32 bit ASN - Working Together

- With the introduction of the “new” 32 bit AS Numbers, and the continuation of use of “old” 16 bit AS Numbers, a way had to be found to get them to work together
- The solution is known as AS23456, which allows BGP to either convert or truncate the AS number if it detects an “old” 16 bit number as part of the exchange

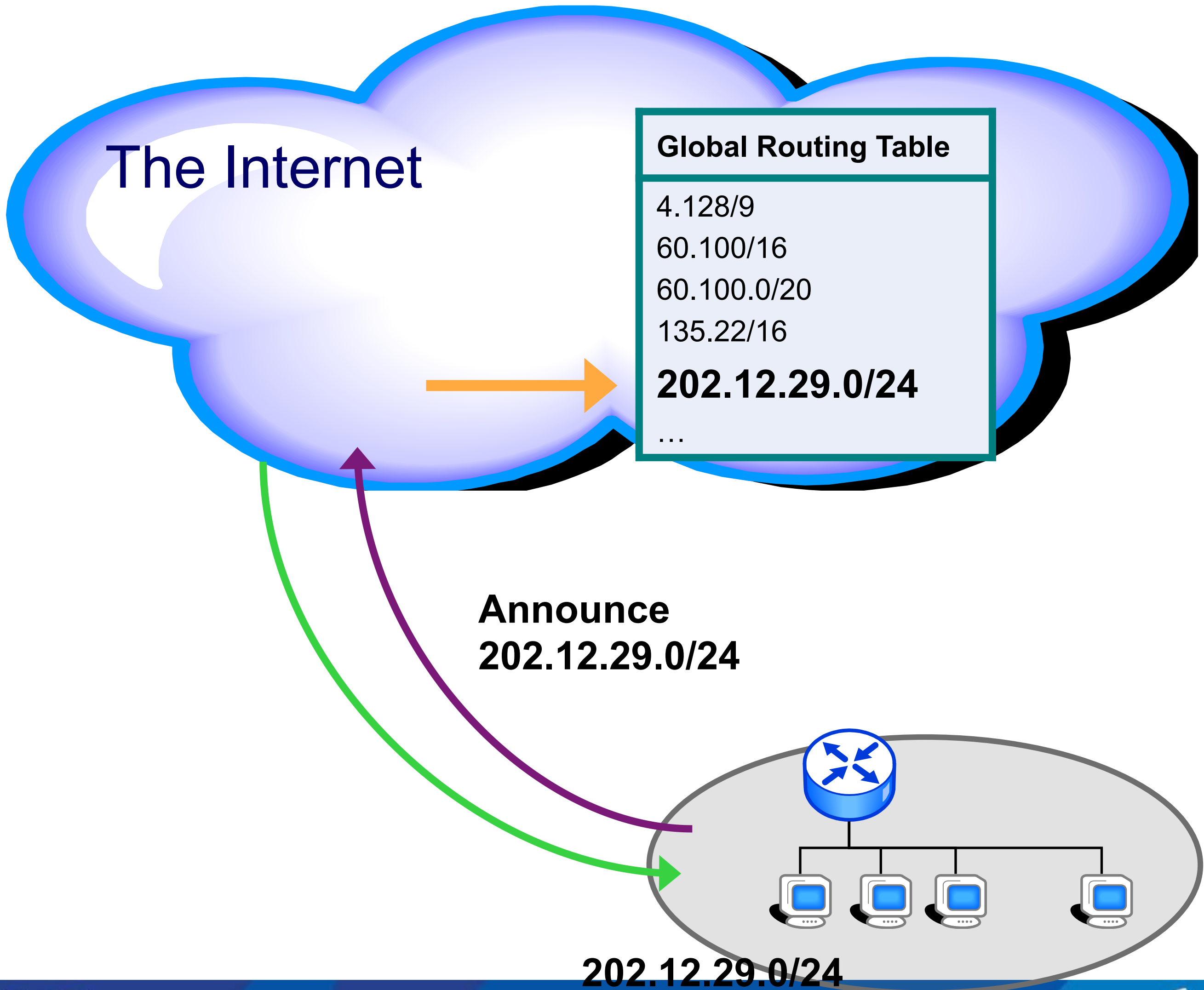
IP Routing Basics

Internet Routing

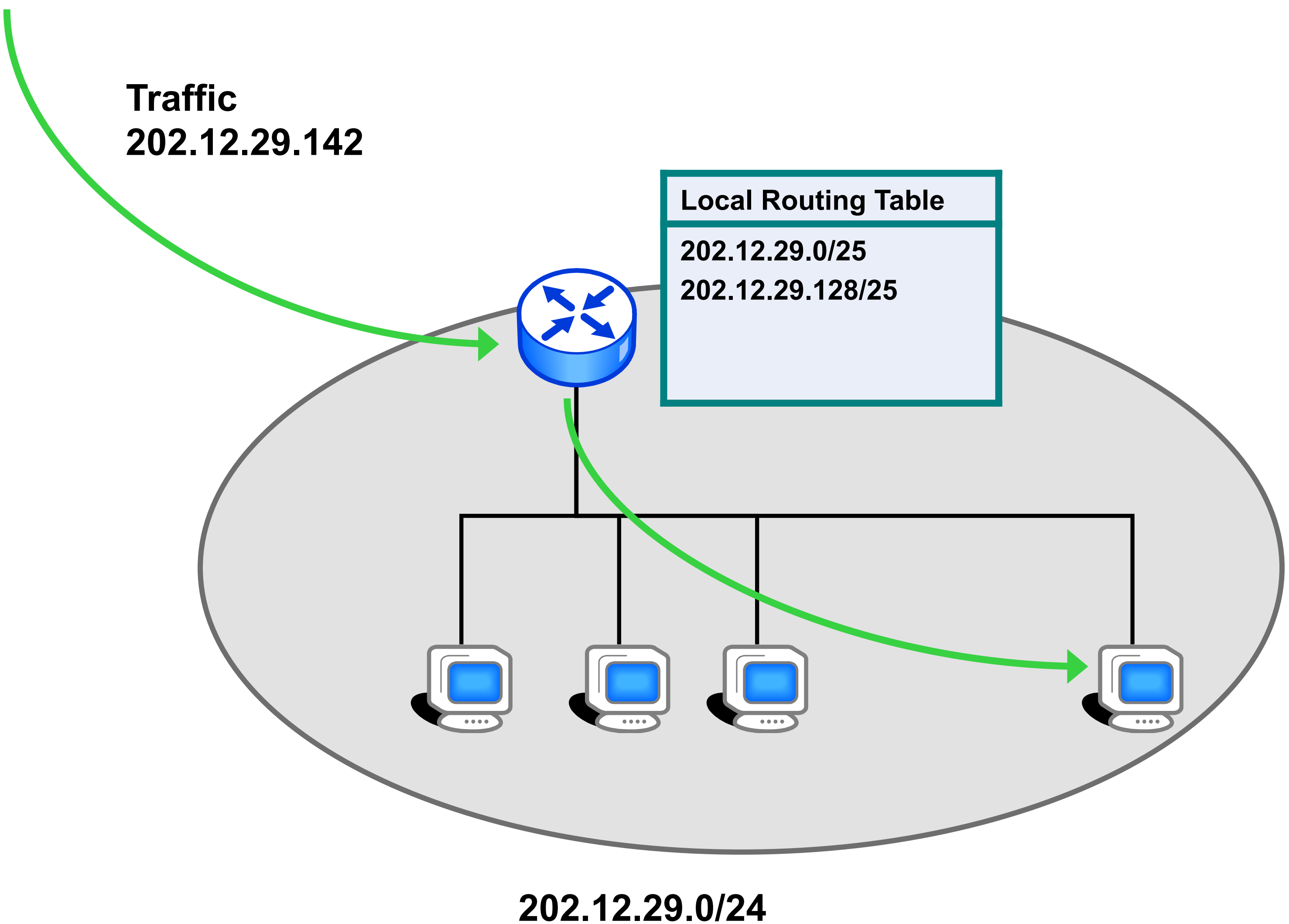


Global Routing Table
4.128/9
60.100/16
60.100.0/20
135.22/16
…

Internet Routing

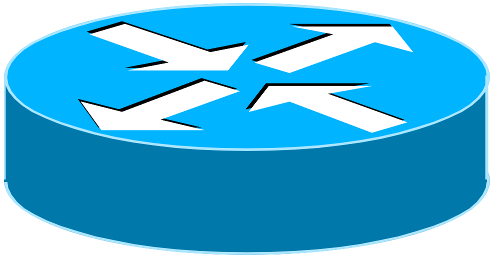


Internet Routing



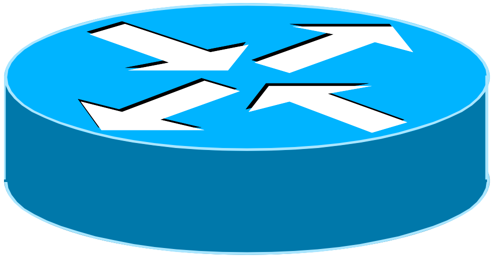
What does a router do?

- ?



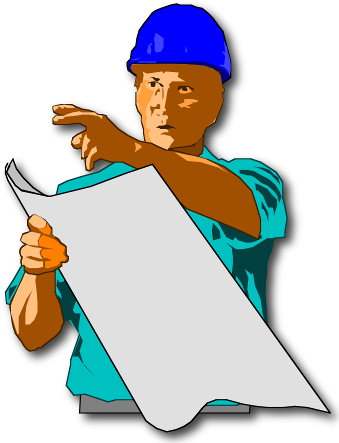
A day in a life of a router

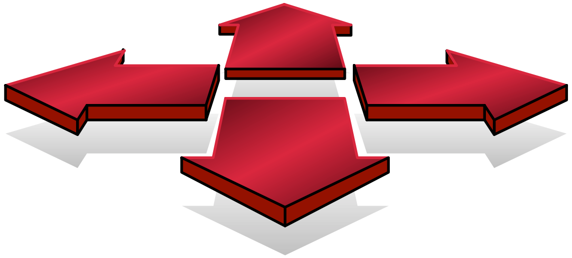
- find path
- forward packet, forward packet, forward packet, forward packet...
- find alternate path
- forward packet, forward packet, forward packet, forward packet…
- repeat until powered off



Routing versus Forwarding

- Routing = building maps and giving directions
- Forwarding = moving packets between interfaces according to the "directions"





IP Routing – finding the path

- Path derived from information received from a routing protocol
- Several alternative paths may exist
  - best path stored in forwarding table
- Decisions are updated periodically or as topology changes (event driven)
- Decisions are based on:
  - topology, policies and metrics (hop count, filtering, delay, bandwidth, etc.)

Metric field

- To determine which path to use if there are multiple paths to the remote network
- Provide the value to select the best path
- But take note of the administrative distance selection process ☺

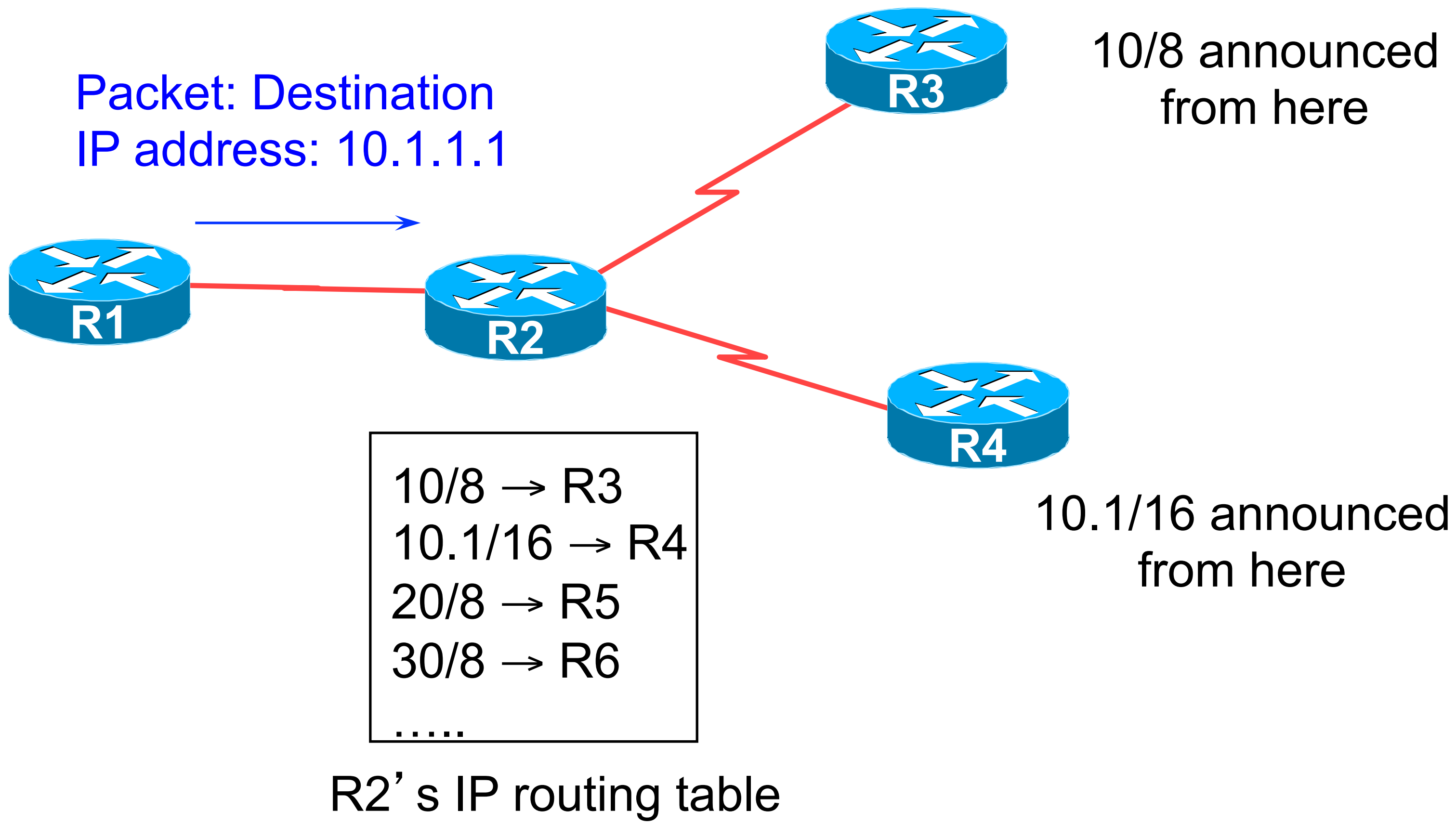
Routing Protocol	Metric
RIPv2	Hop count
EIGRP	Bandwidth, delay, load, reliability, MTU
OSPF	Cost (the higher the bandwidth indicates a lower cost)
IS-IS	Cost

IP route lookup

- Based on destination IP address
- “longest match” routing
  - More specific prefix preferred over less specific prefix
  - Example: packet with destination of 10.1.1.1/32 is sent to the router announcing 10.1/16 rather than the router announcing 10/8.

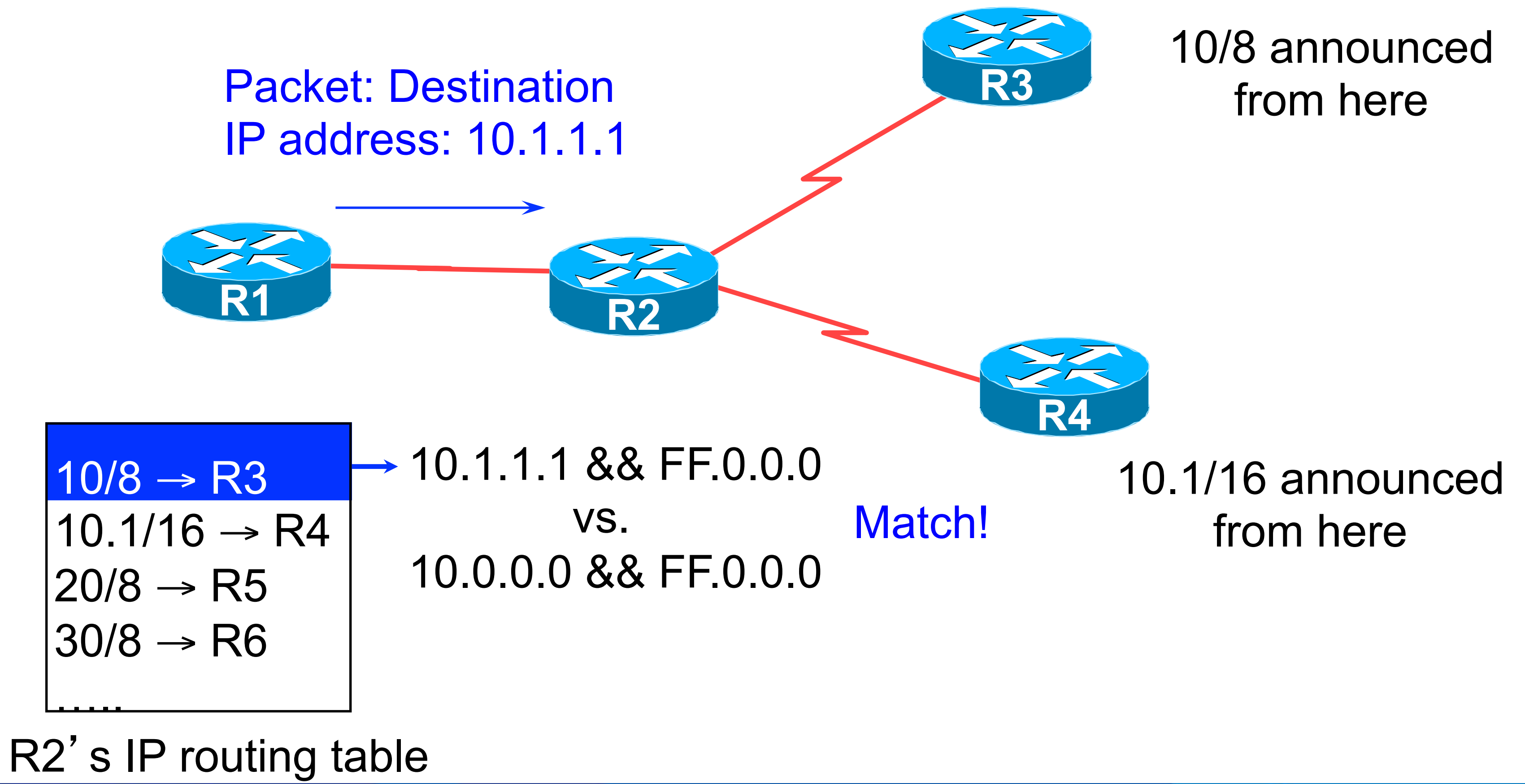
IP route lookup

- Based on destination IP address



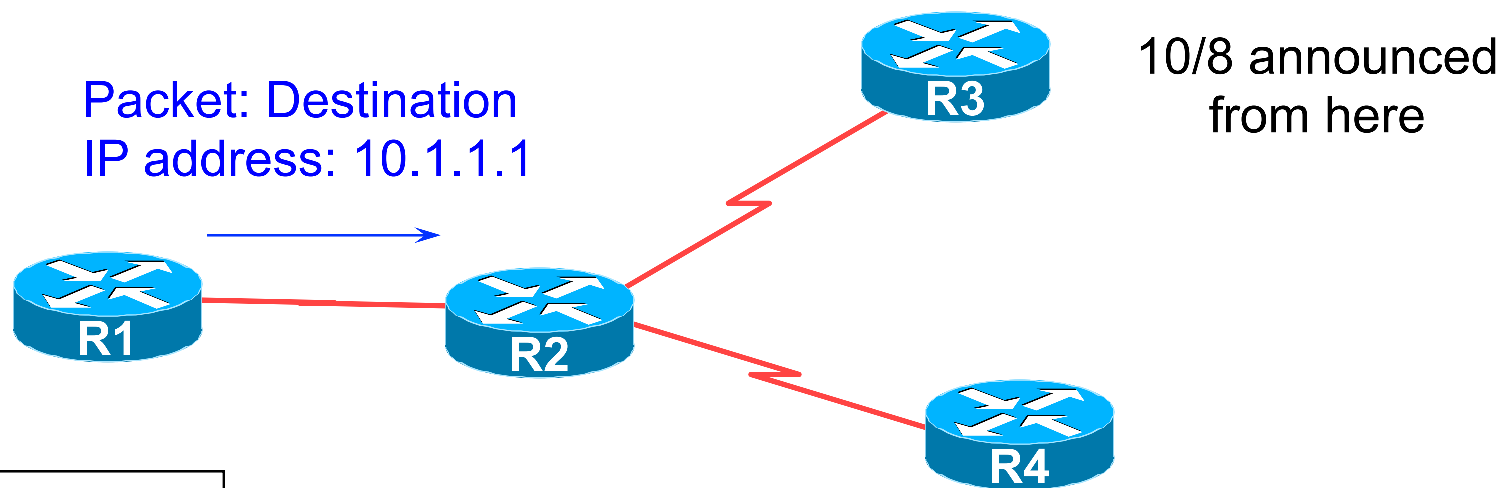
IP route lookup: Longest match routing

- Based on destination IP address



IP route lookup: Longest match routing

- Based on destination IP address



10.1/16 announced from here

10/8 → R3
10.1/16 → R4
20/8 → R5
30/8 → R6
…..

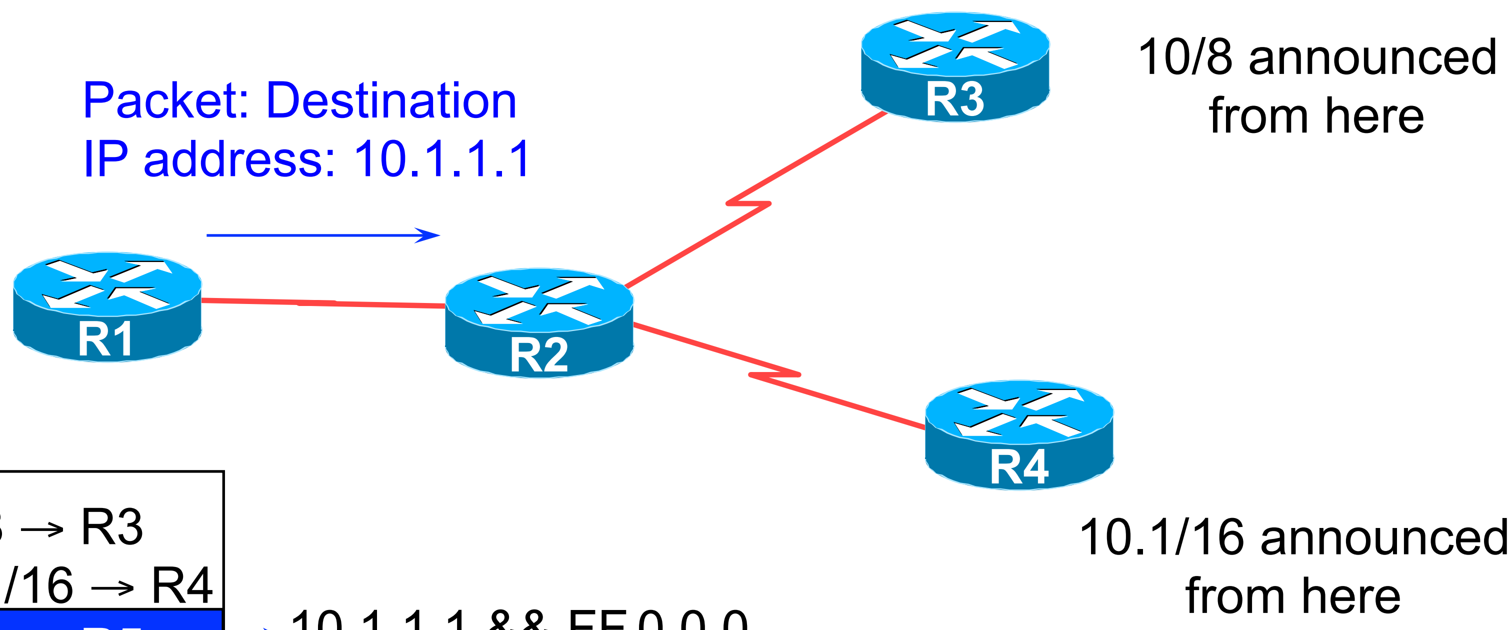
R2's IP routing table

→ 10.1.1.1 && FF.FF.0.0
vs.
10.1.0.0 && FF.FF.0.0

Match as well!

IP route lookup: Longest match routing

- Based on destination IP address



10/8 → R3
10.1/16 → R4
20/8 → R5
30/8 → R6
…..

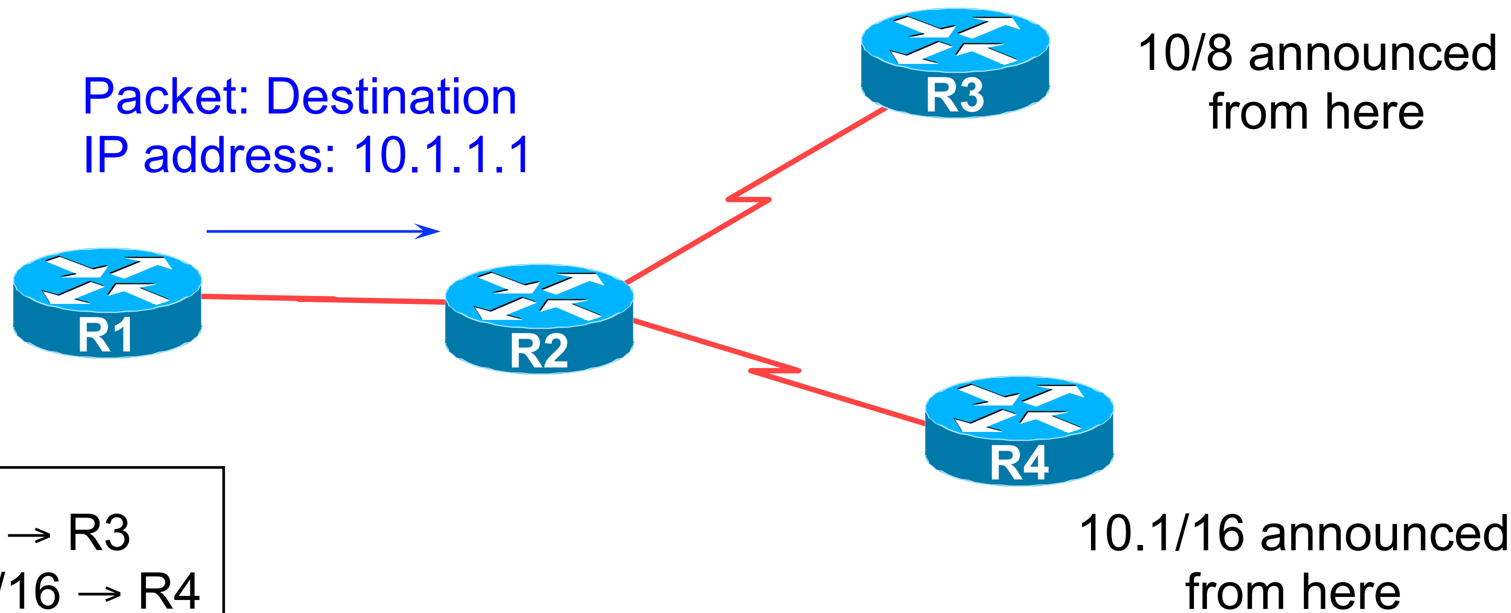
R2’s IP routing table

10.1.1.1 && FF.0.0.0
vs.
20.0.0.0 && FF.0.0.0

Does not match!

IP route lookup: Longest match routing

- Based on destination IP address



10/8 → R3
10.1/16 → R4
20/8 → R5
30/8 → R6
.....

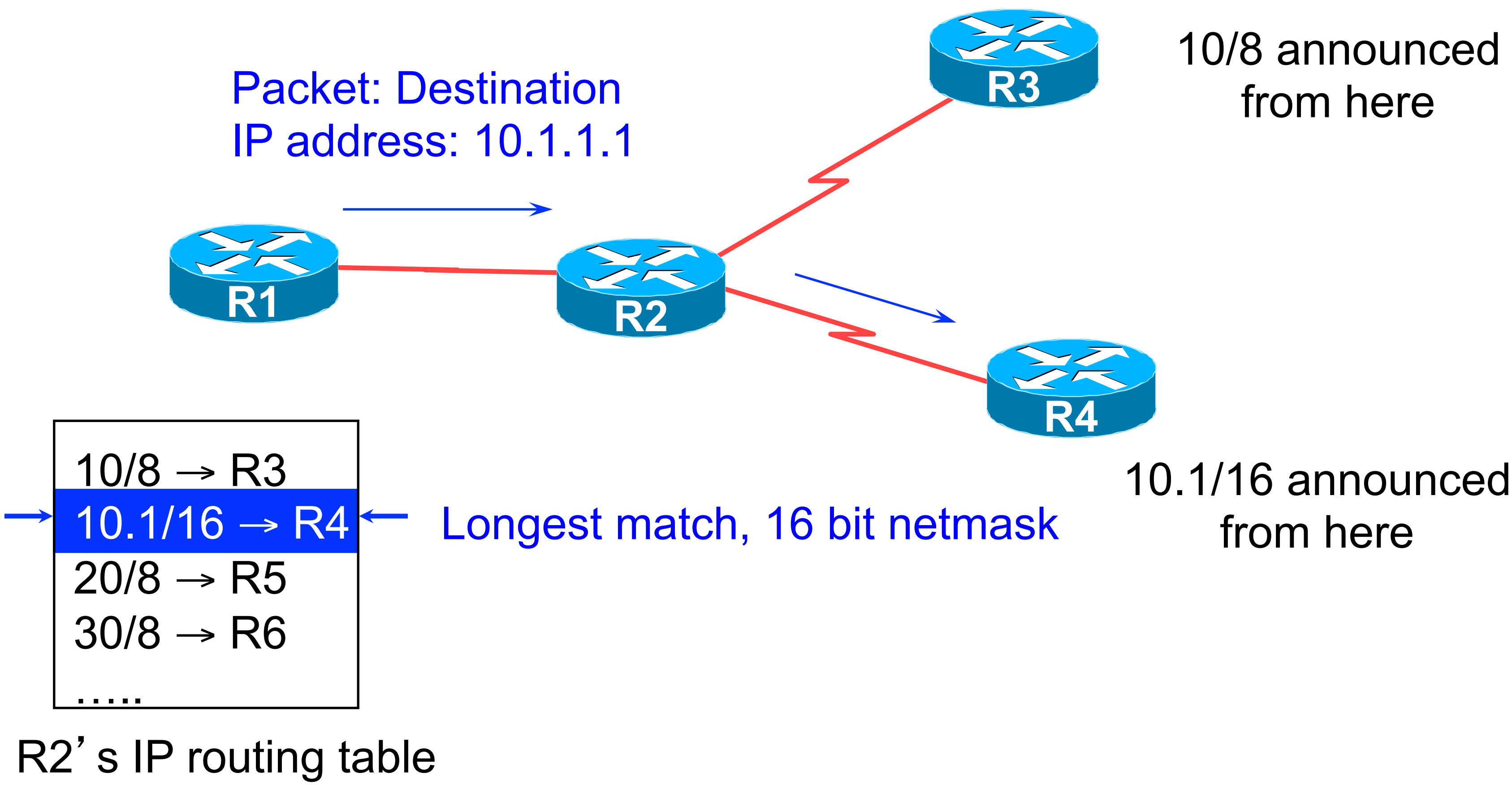
R2's IP routing table

10.1.1.1 && FF.0.0.0
vs.
30.0.0.0 && FF.0.0.0

Does not match!

IP route lookup: Longest match routing

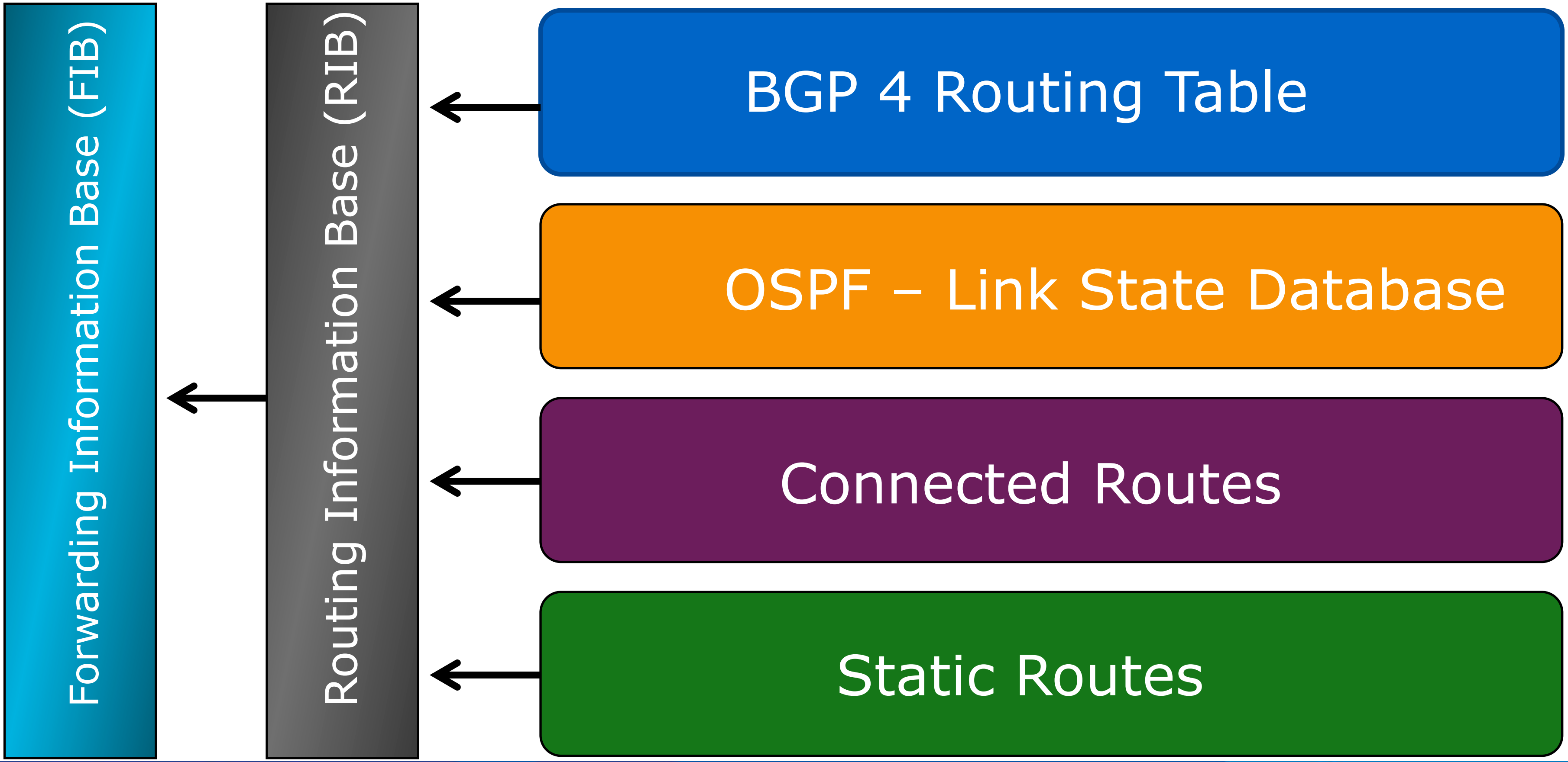
- Based on destination IP address



RIBs and FIBs

- FIB is the Forwarding Table
  - It contains destinations and the interfaces to get to those destinations
  - Used by the router to figure out where to send the packet
  - Careful! Some people still call this a route!
- RIB is the Routing Table
  - It contains a list of all the destinations and the various next hops used to get to those destinations – and lots of other information too!
  - One destination can have lots of possible next-hops – only the best next-hop goes into the FIB

Routing Tables Feed the Forwarding Table



Explicit versus Default Routing

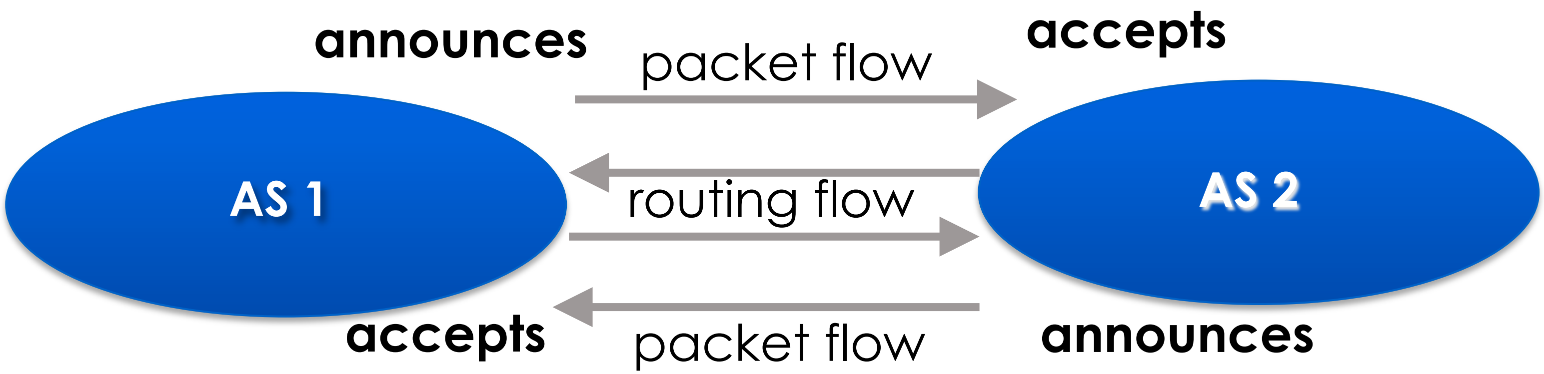
- Default:
  - simple, cheap (cycles, memory, bandwidth)
  - low granularity (metric games)
- Explicit (default free zone)
  - high overhead, complex, high cost, high granularity
- Hybrid
  - minimise overhead
  - provide useful granularity
  - requires some filtering knowledge

Routing Policy

- Used to control traffic flow in and out of an ISP network
- ISP makes decisions on what routing information to accept and discard from its neighbours
  - Individual routes
  - Routes originated by specific ASes
  - Routes traversing specific ASes
  - Routes belonging to other groupings
    - Groupings which you define as you see fit

Representation of Routing Policy

- Routing and packet flows

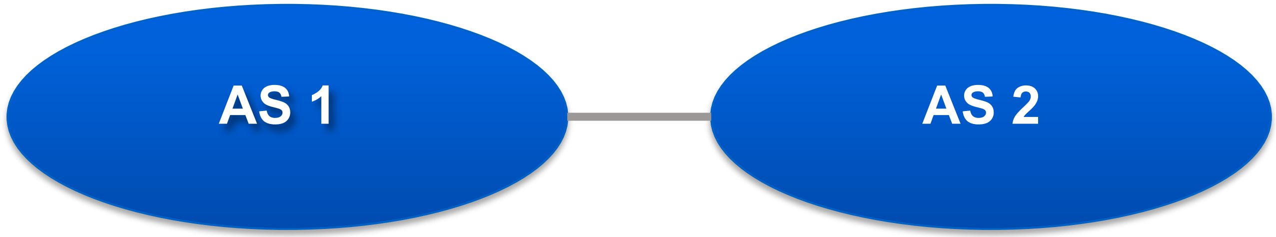


For AS1 and AS2 networks to communicate

- AS1 must announce to AS2
- AS2 must accept from AS1
- AS2 must announce to AS1
- AS1 must accept from AS2

Representation of Routing Policy

Basic concept



"action pref" - the lower the value, the more preferred the route

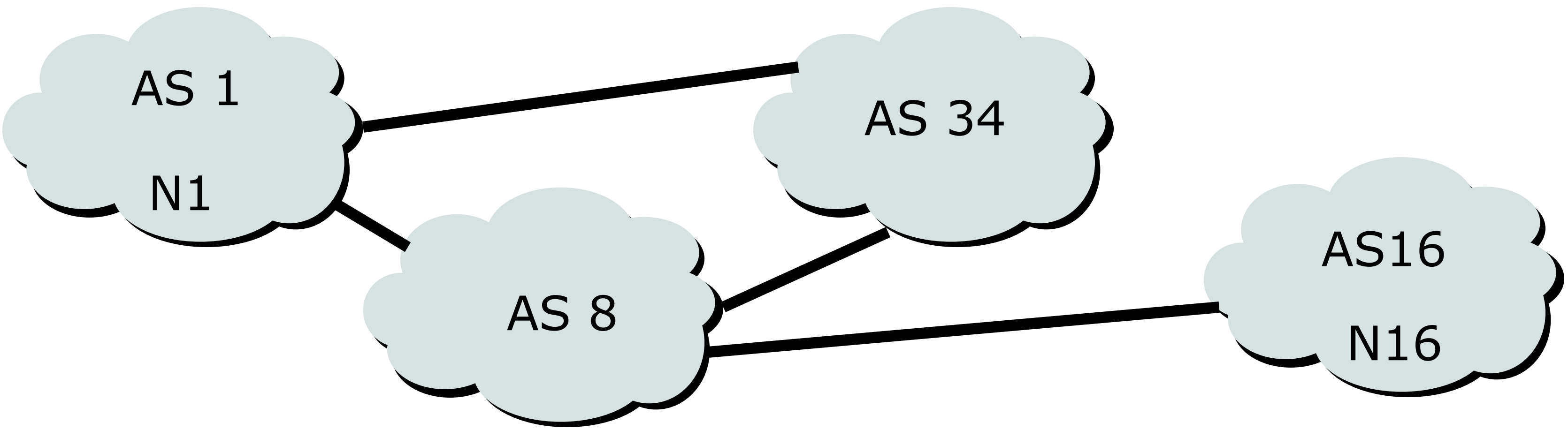
```
aut-num: AS1
…
import:   from AS2
          action pref=100;
          accept AS2
export:   to AS2 announce AS1
```

```
aut-num: AS2
…
import:   from AS1
          action pref=100;
          accept AS1
export:   to AS1 announce AS2
```

Routing flow and Traffic flow

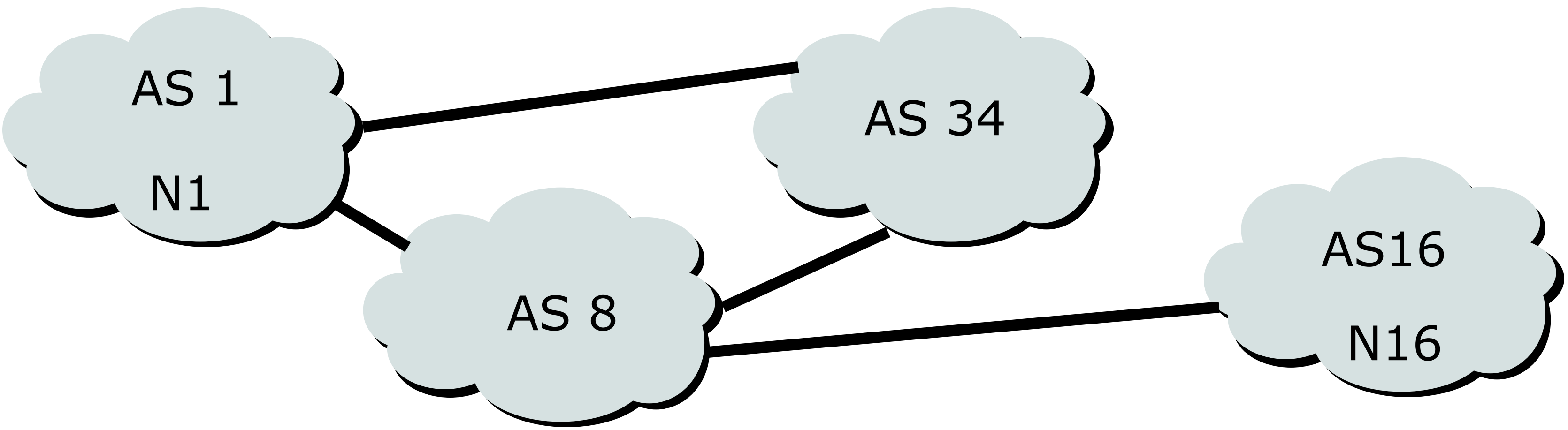
- Traffic flow is always in the opposite direction of the flow of Routing information
  - Filtering outgoing routing information inhibits traffic flow inbound
  - Filtering inbound routing information inhibits traffic flow outbound

Routing Flow/Packet Flow: With multiple ASes



- For net N1 in AS1 to send traffic to net N16 in AS16:
  - AS16 must originate and announce N16 to AS8.
  - AS8 must accept N16 from AS16.
  - AS8 must forward announcement of N16 to AS1 or AS34.
  - AS1 must accept N16 from AS8 or AS34.
- For two-way packet flow, similar policies must exist for N1

Routing Flow/Packet Flow: With multiple ASes



- As multiple paths between sites are implemented it is easy to see how policies can become quite complex.

Routing Protocols

- Routers use “routing protocols” to exchange routing information with each other
  - IGP is used to refer to the process running on routers inside an ISP’s network
  - EGP is used to refer to the process running between routers bordering directly connected ISP networks

What Is an IGP?

- Interior Gateway Protocol
- Within an Autonomous System
- Carries information about internal infrastructure prefixes
- Two widely used IGPs in service provider network:
  - OSPF
  - ISIS

Why Do We Need an IGP?

- ISP backbone scaling
  - Hierarchy
  - Limiting scope of failure
  - Only used for ISP's infrastructure addresses, not customers or anything else
  - Design goal is to minimise number of prefixes in IGP to aid scalability and rapid convergence

What Is an EGP?

- Exterior Gateway Protocol
- Used to convey routing information between Autonomous Systems
- De-coupled from the IGP
- Current EGP is BGP

Why Do We Need an EGP?

- Scaling to large network
  - Hierarchy
  - Limit scope of failure
- Define Administrative Boundary
- Policy
  - Control reachability of prefixes
  - Merge separate organisations
  - Connect multiple IGPs

Administrative Distance

- method used for selection of route priority of IP routing protocol, the lowest administrative distance is preferred
  - Manually entered routes are preferred from dynamically learned routes
    - Static routes
    - Default routes
  - Dynamically learned routes depend on the routing protocol metric calculation algorithm and default metrics values the smallest metric value are preferred

Administrative Distance Chart (Cisco)

Routed Sources	Default Distance
Connected interface	0
Static route out an interface	0
Static route to a next hop	1
External BGP	20
IGRP	100
OSPF	110
IS-IS	115
RIP v1, v2	120
EGP	140
Internal BGP	200
Unknown	255

DNS and Reverse DNS

APNIC

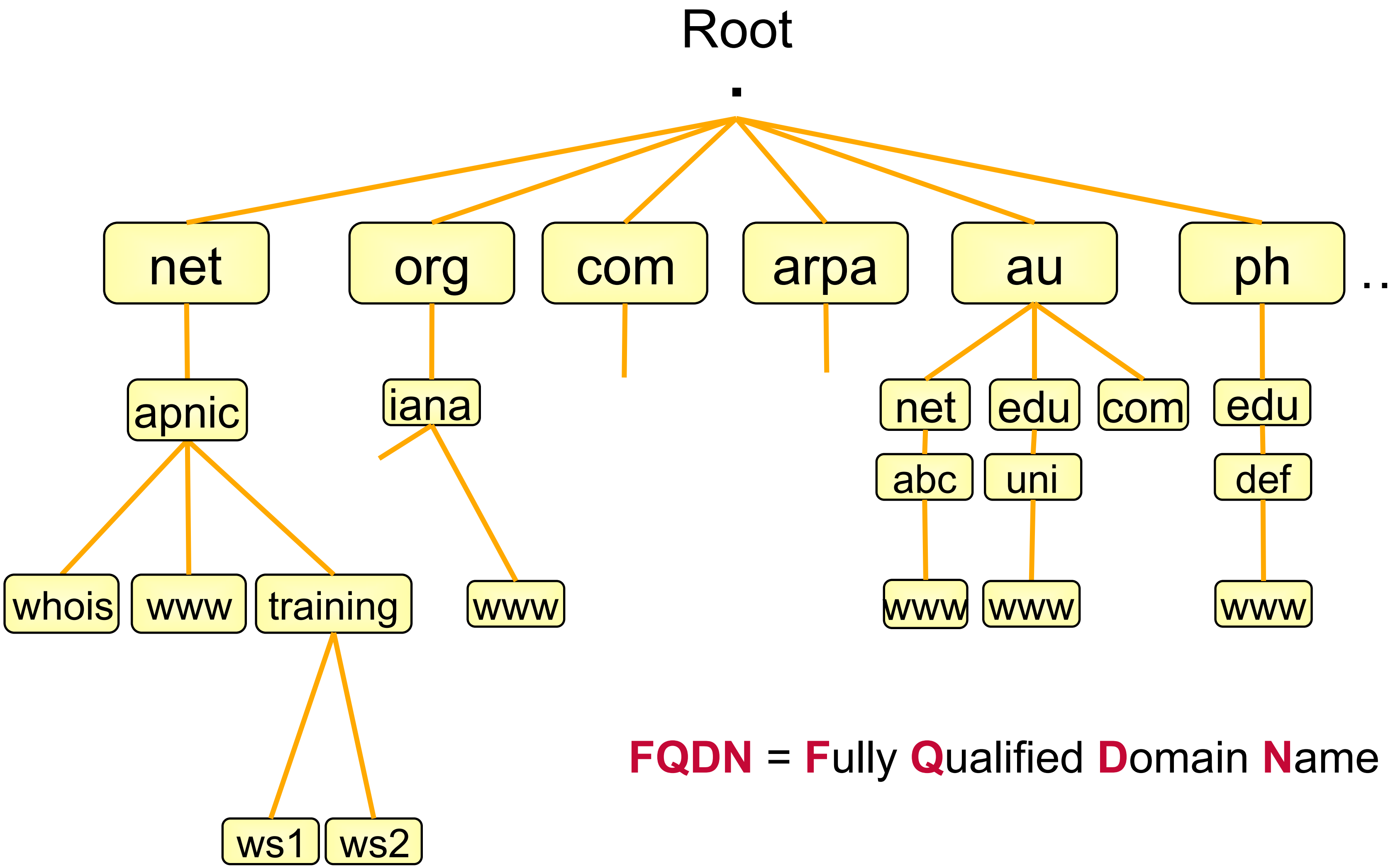
Domain Name System

- A lookup mechanism for translating objects into other objects
  - Mapping <u>names</u> to <u>numbers</u> and vice versa
- A globally distributed, loosely coherent, scalable, reliable, dynamic database
- Comprised of three components
  - A “name space”
  - Servers making that name space available
  - Resolvers (clients) which query the servers about the name space
- A critical piece of the Internet infrastructure

DNS Features

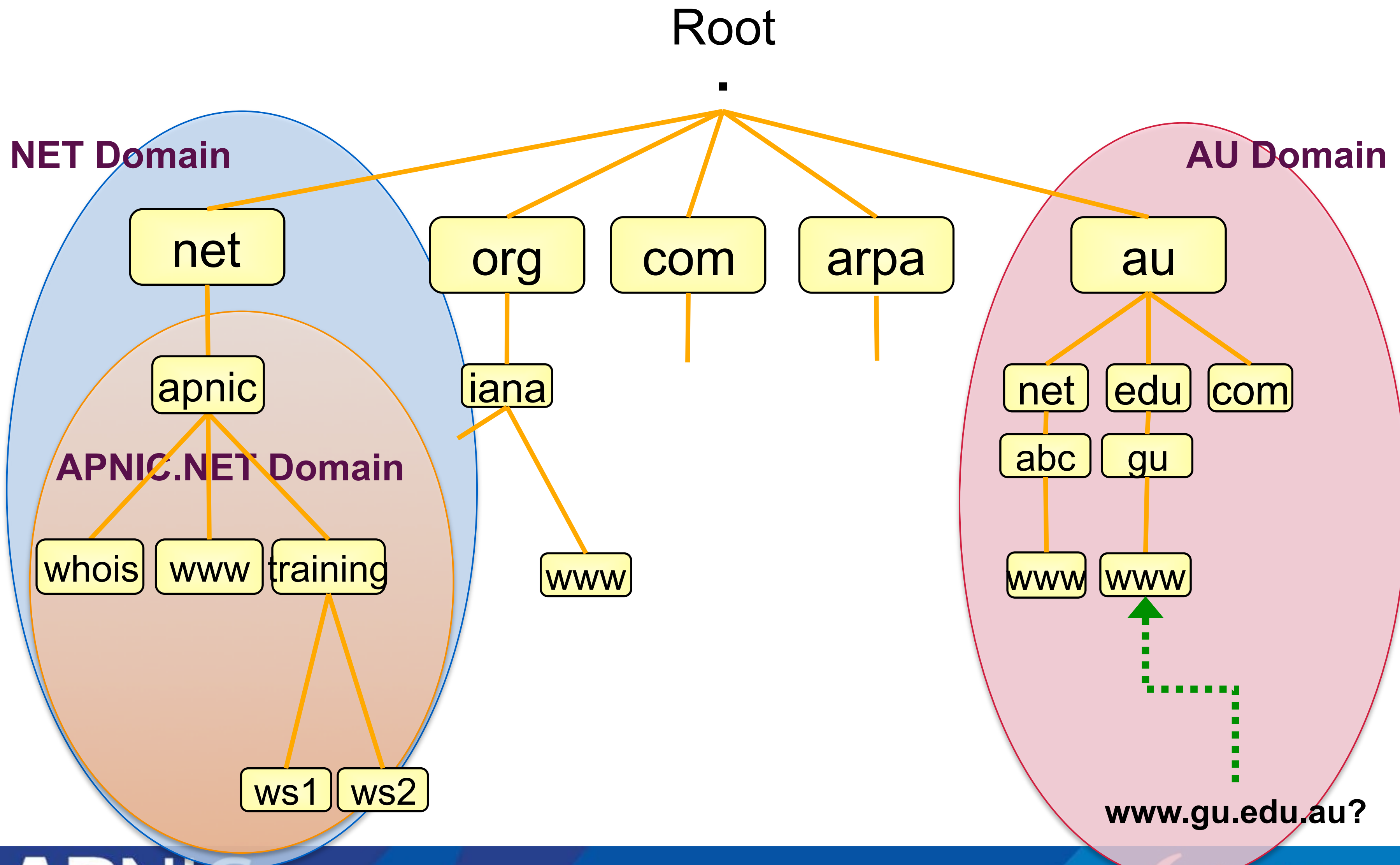
- Global distribution
  - Shares the load and administration
- Loose Coherency
  - Geographically distributed, but still coherent
- Scalability
  - can add DNS servers without affecting the entire DNS
- Reliability
- Dynamicity
  - Modify and update data dynamically

DNS Tree Hierarchy



FQDN = **F**ully **Q**ualified **D**omain **N**ame

Domains


Root
.
NET Domain
AU Domain
net
org
com
arpa
au
apnic
iana
net
edu
com
APNIC.NET Domain
abc
gu
whois
www
training
www
www
www
ws1
ws2
www.gu.edu.au?

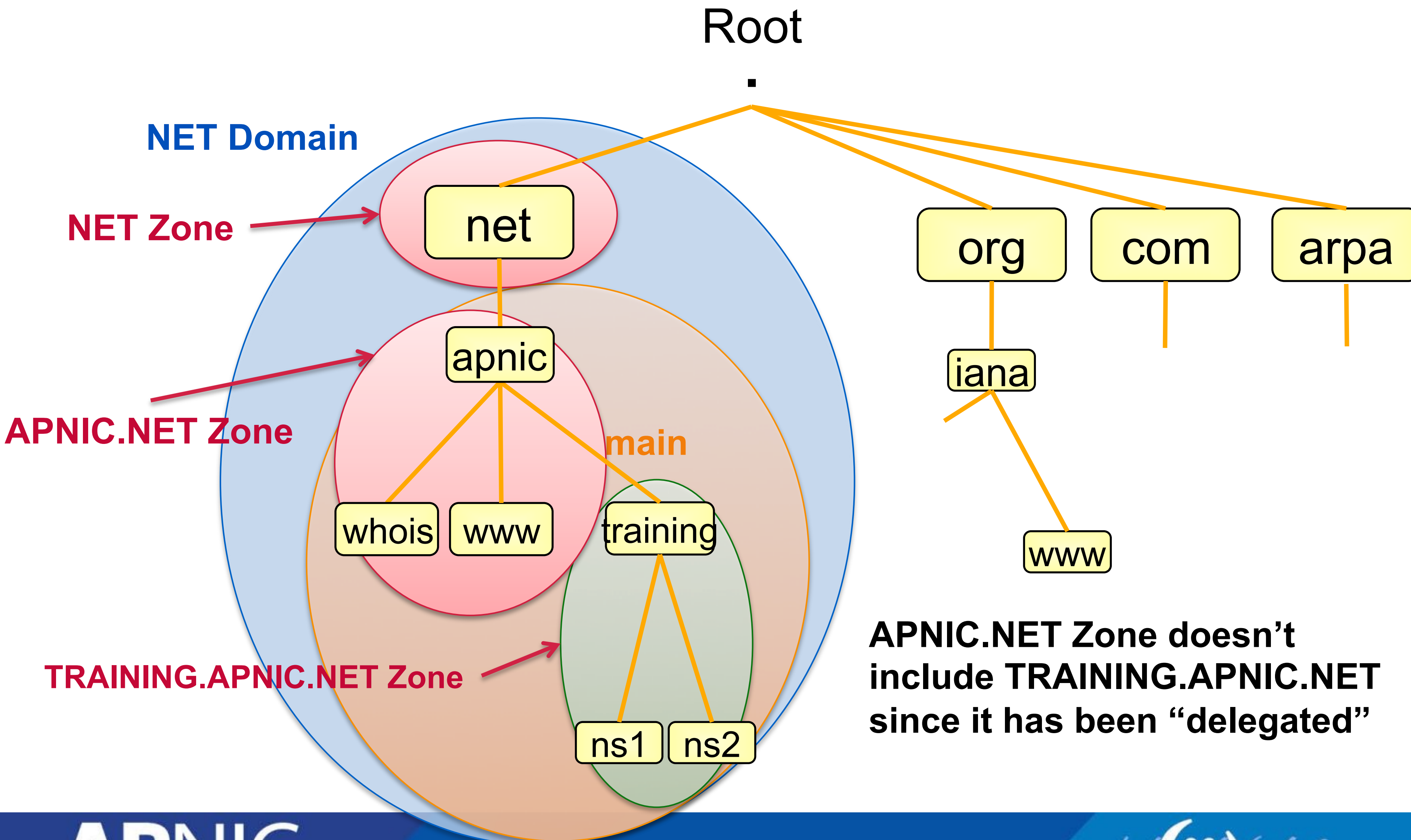
Delegation

- Administrators can create subdomains to group hosts
  - According to geography, organizational affiliation or any other criterion
- An administrator of a domain can delegate responsibility for managing a subdomain to someone else
- The parent domain retains links to the delegated subdomain
  - The parent domain “remembers” to whom it delegated the subdomain

Zones and Delegations

- Zones are “administrative spaces”
- Zone administrators are responsible for portion of a domain’s name space
- Authority is delegated from parent to child

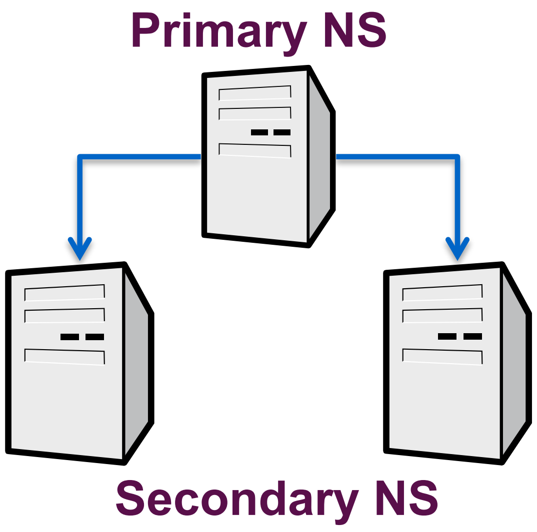
Zones



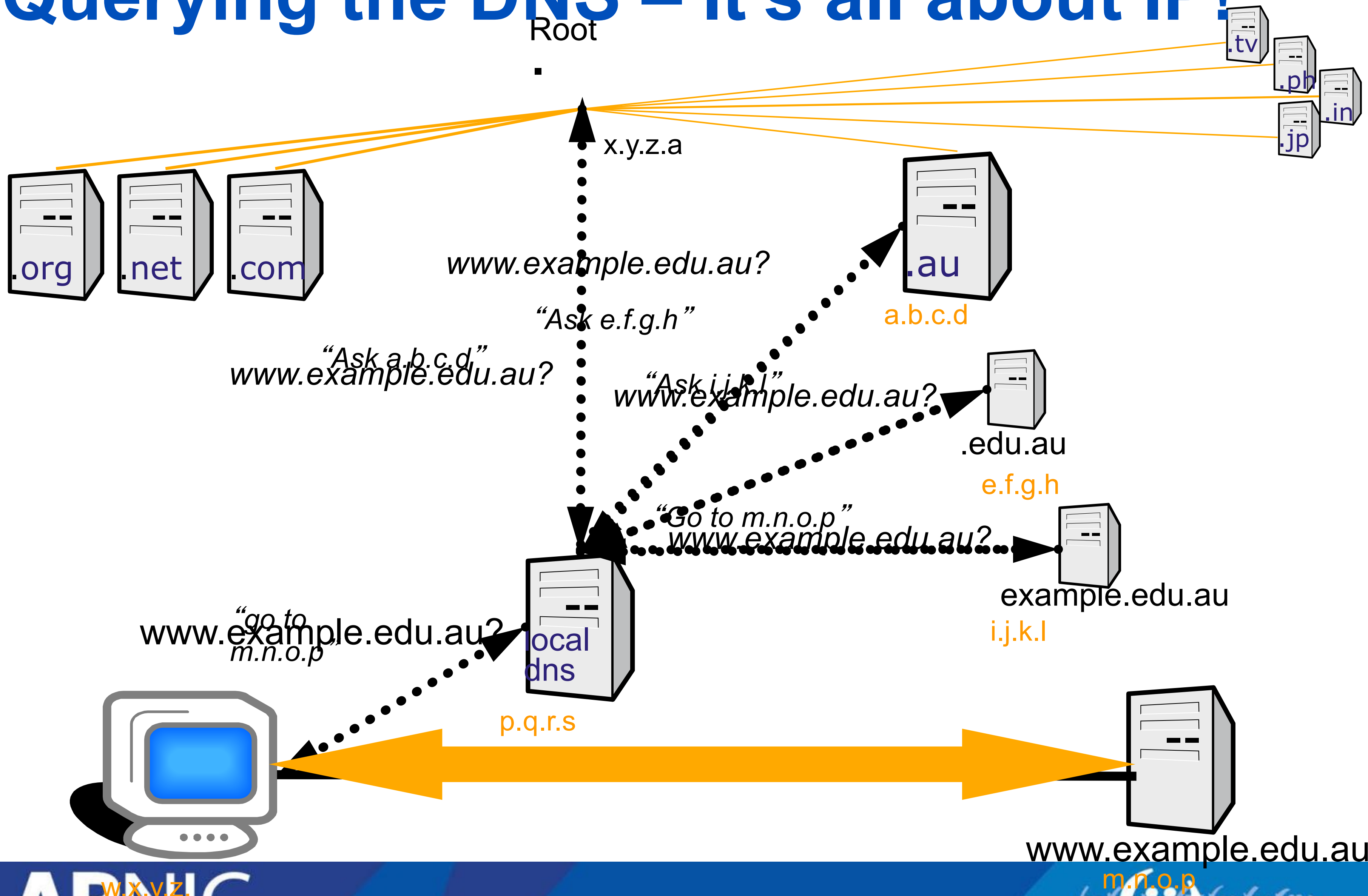
APNIC.NET Zone doesn't include TRAINING.APNIC.NET since it has been "delegated"

Name Servers

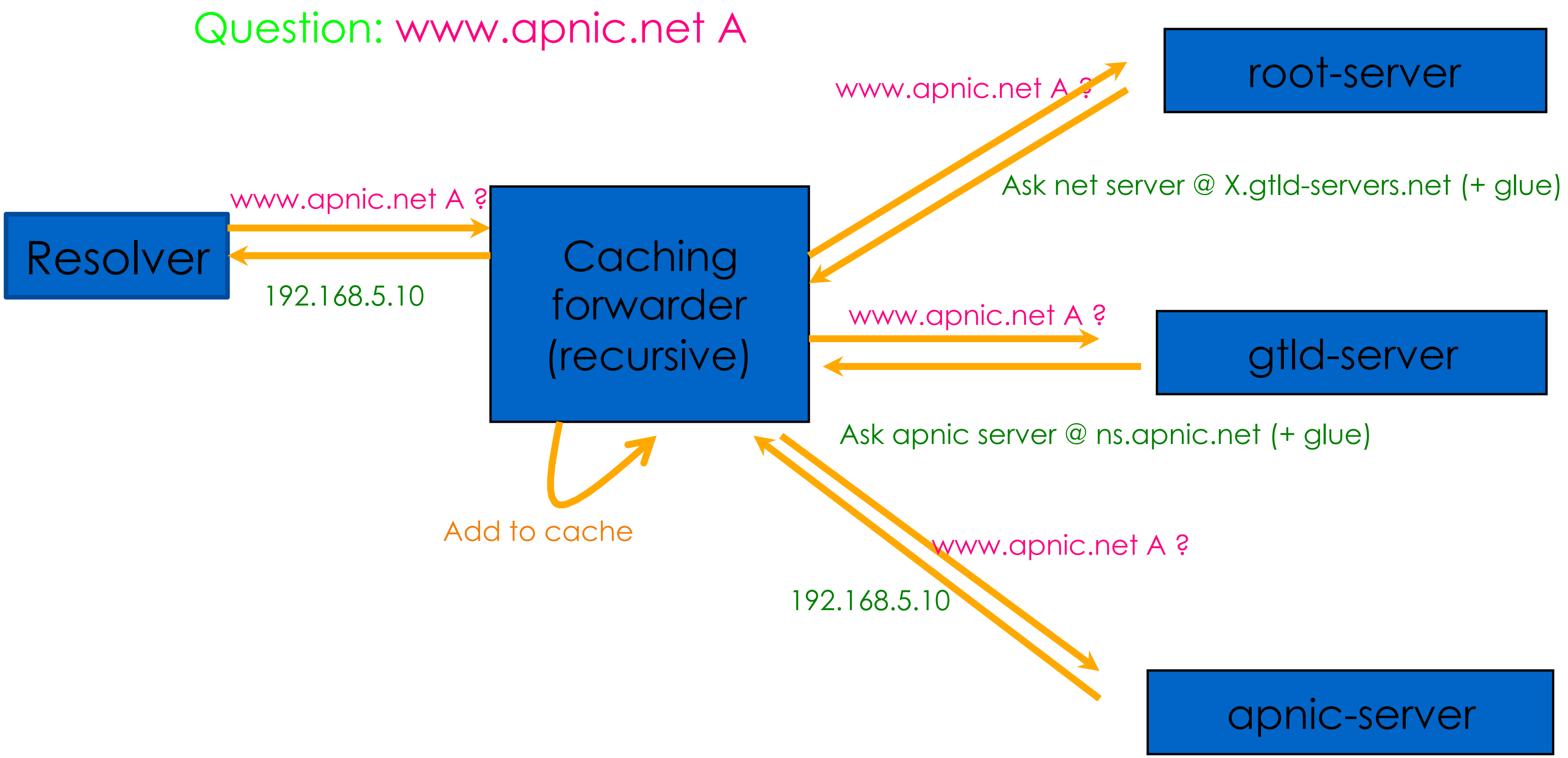
- Name servers answer ‘DNS’ questions
- Several types of name servers
  - Authoritative servers
    - Master / primary
    - Slave / secondary
  - Caching or recursive servers
    - also caching forwarders
- Mixture of functions



Querying the DNS – It's all about IP!



Concept: Resolving process & Cache



Resource Records

- Entries in the DNS zone file
- Components:

Resource Record	Function
Label	Name substitution for FQDN
TTL	Timing parameter, an expiration limit
Class	IN for Internet, CH for Chaos
Type	RR Type (A, AAAA, MX, PTR) for different purposes
RDATA	Anything after the Type identifier; Additional data

Common Resource Record Types

RR Type	Name	Functions
A	Address record	Maps domain name to IP address `www.apnic.net. IN A 203.176.189.99`
AAAA	IPv6 address record	Maps domain name to an IPv6 address `www.apnic.net. IN AAAA 2001:db8::1`
NS	Name server record	Used for delegating zone to a nameserver `apnic.net. IN NS ns1.apnic.net.`
PTR	Pointer record	Maps an IP address to a domain name `99.189.176.203.in-addr.arpa. IN PTR www.apnic.net.`
CNAME	Canonical name	Maps an alias to a hostname `web IN CNAME www.apnic.net.`

MX	Mail Exchanger	Defines where to deliver mail for user @ domain
		`apnic.net. IN MX 10 mail01.apnic.net.`
		`IN MX 20 mail02.apnic.net.`

Start of Authority (SOA) record

```
Domain_name. CLASS  SOA  hostname.domain.name. mailbox.domain.name (
                                        Serial Number
                                        Refresh
                                        Retry
                                        Expire
                                        Minimum TTL )
```

- **Serial Number** – must be updated if any changes are made in the zone file
- **Refresh** – how often a secondary will poll the primary server to see if the serial number for the zone has increased
- **Retry** - If a secondary was unable to contact the primary at the last refresh, wait the retry value before trying again
- **Expire** - How long a secondary will still treat its copy of the zone data as valid if it can't contact the primary.
- **Minimum TTL** - The default TTL (time-to-live) for resource records

TTL Time Values

- The right value depends on your domain
- Recommended time values for TLD (based on RFC 1912)

Refresh	86400 (24h)
Retry	7200 (2h)
Expire	2592000 (30d)
Min TTL	345600 (4d)

- For other servers – optimize the values based on
  - Frequency of changes
  - Required speed of propagation
  - Reachability of the primary server
  - (and many others)

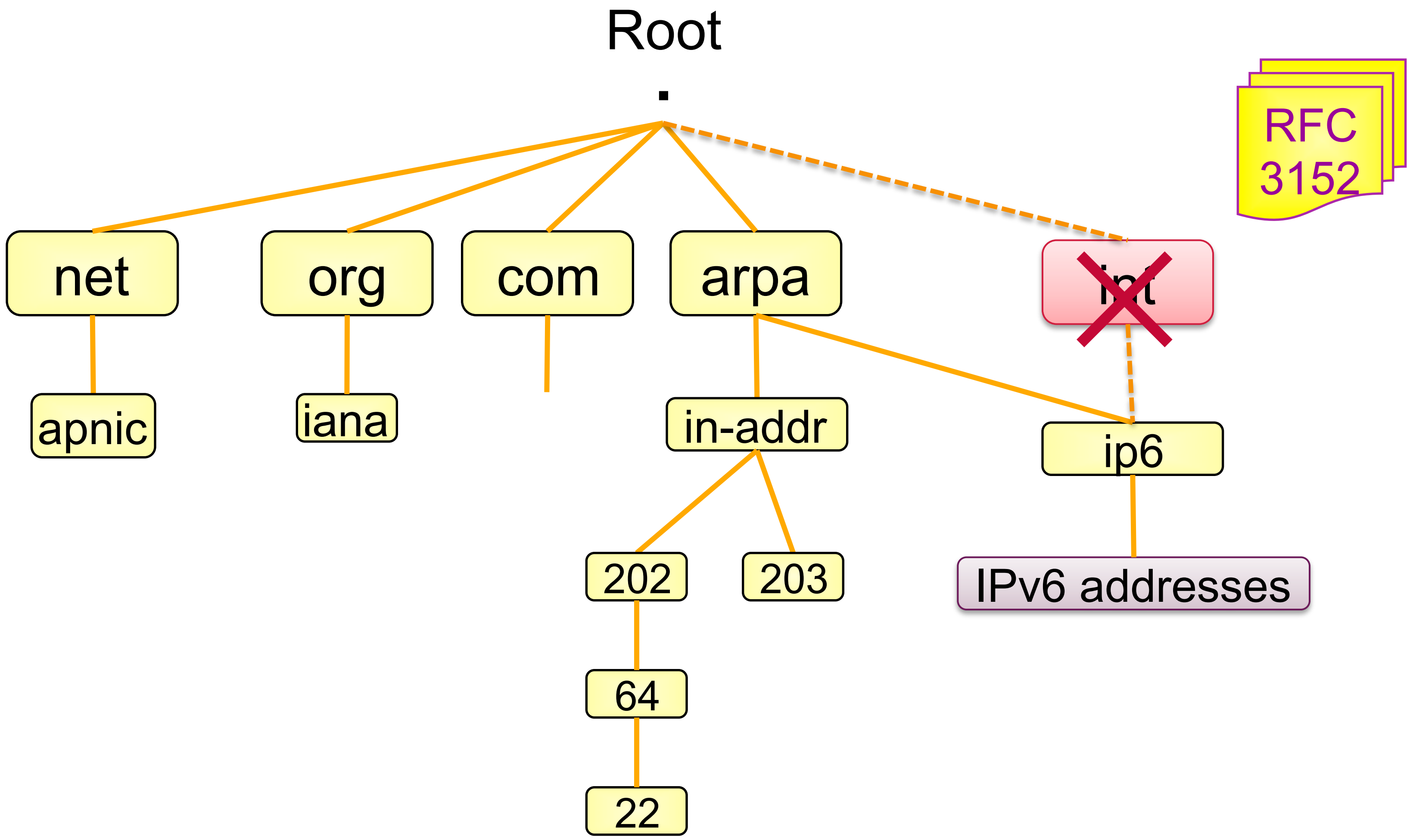
Example: RRs in a Zone File

```
apnic.net. 7200 IN    SOA    ns.apnic.net. admin.apnic.net. (
                20130522          ; Serial
                12h               ; Refresh 12 hours
                4h                ; Retry 4 hours
                4d                ; Expire 4 days
                2h                ; Negative cache 2 hours )

apnic.net.          7200   IN   NS   ns.apnic.net.
apnic.net.          7200   IN   NS   ns.ripe.net.
whois.apnic.net.    3600   IN   A    193.0.1.162
www.apnic.net       3600   IN   A    192.0.3.25
```

Label	TTL	Class	Type	Rdata
www.apnic.net	3600	IN	A	192.0.3.25

Reverse DNS Tree – with IPv6



Pointer (PTR) records

- Create pointer (PTR) records for each IP address

```
131.28.12.202.in-addr.arpa. IN PTR svc00.apnic.net.
```

or

```
131            IN      PTR          svc00.apnic.net.
```

IPv6 Reverse Lookups – PTR records

- Similar to the IPv4 reverse record

```
b.a.9.8.7.6.5.0.4.0.0.0.3.0.0.0.2.0.0.0.1.0.0.0.0.0.0.0.1.2.3.4.ip6.arpa.
        IN     PTR  test.ip6.example.com.
```

- Example: reverse name lookup for a host with address 3ffe:8050:201:1860:42::1

```
$ORIGIN 0.6.8.1.1.0.2.0.0.5.0.8.e.f.f.3.ip6.arpa.

1.0.0.0.0.0.0.0.0.0.0.0.2.4.0.0  14400  IN PTR host.example.com.
```

Reverse Delegation Requirements

- /24 Delegations
  - Address blocks should be assigned/allocated
  - At least two name servers
- /16 Delegations
  - Same as /24 delegations
  - APNIC delegates entire zone to member
  - Recommend APNIC secondary zone
- < /24 Delegations
  - Read “Classless IN-ADDR.ARPA delegation” (RFC 2317)

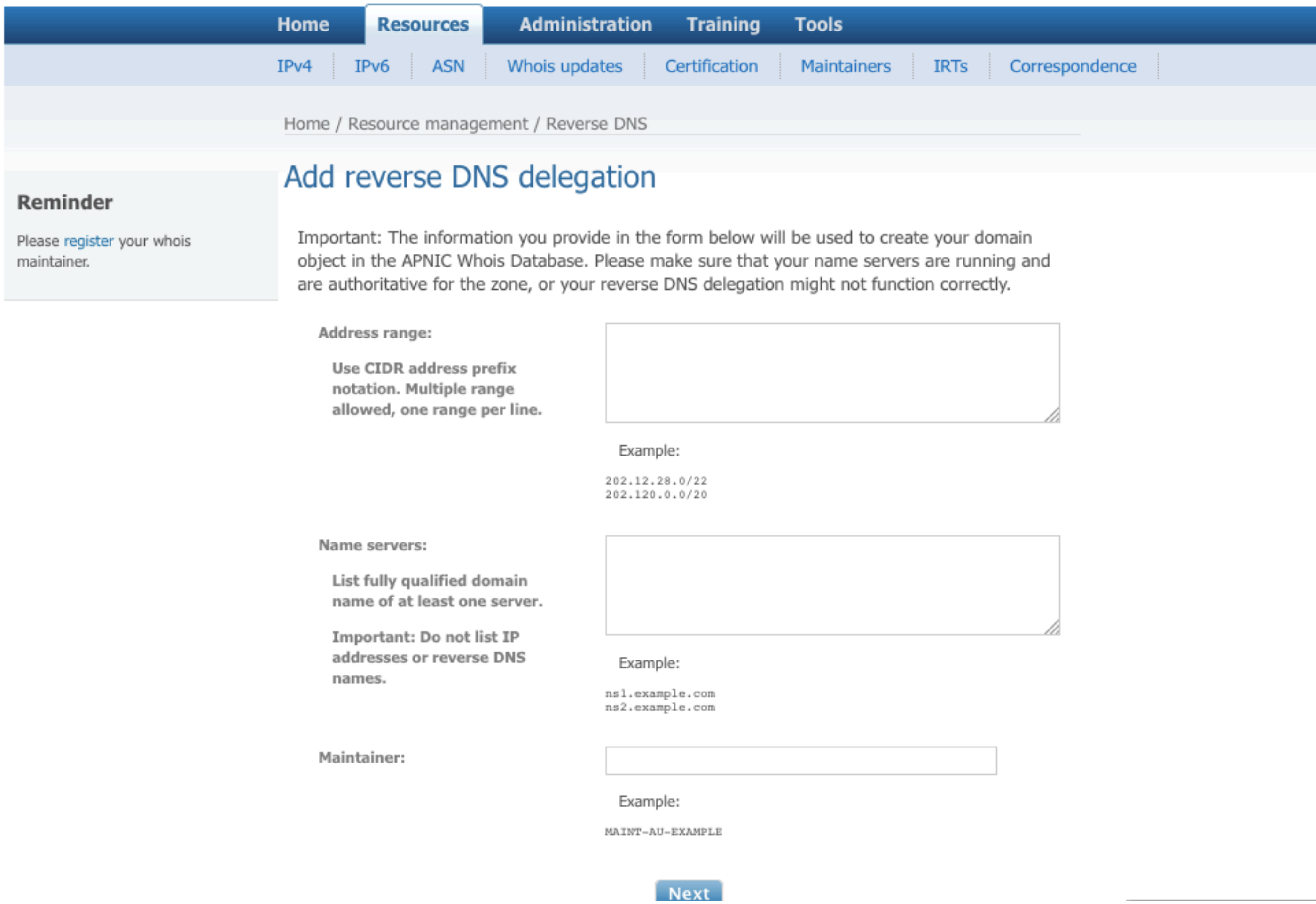
APNIC & ISPs Responsibilities

- APNIC
  - Manage reverse delegations of address block distributed by APNIC
  - Process organisations requests for reverse delegations of network allocations
- Organisations
  - Be familiar with APNIC procedures
  - Ensure that addresses are reverse-mapped
  - Maintain nameservers for allocations
  - Minimise pollution of DNS

Reverse Delegation Procedures

- Standard APNIC database object,
  - can be updated through myAPNIC
- Nameserver/domain set up verified before being submitted to the database.
- Protection by maintainer object
  - (current auths: CRYPT-PW, PGP).
- Any queries
  - Contact helpdesk@apnic.net

Reverse Delegation Procedures



Whois domain object

Reverse Zone

```
domain:     28.12.202.in-addr.arpa
Descr:      in-addr.arpa zone for 28.12.202.in-addr.arpa
admin-c:    NO4-AP
tech-c:     AIC1-AP
zone-c:     NO4-AP
nserver:    cumin.apnic.net
nserver:    tinnie.apnic.net
nserver:    tinnie.arin.net
mnt-by:     MAINT-APNIC-AP
mnt-lower:  MAINT-AP-DNS
changed:    inaddr@apnic.net 20021023
changed:    inaddr@apnic.net 20040109
changed:    hm-changed@apnic.net 20091007
changed:    hm-changed@apnic.net 20111208
source:     APNIC
```

Contacts

Nameservers

Maintainers

Thank You

APNIC