Lec2 – The Internet

50.012 Networks

Jit Biswas

(based on earlier version co-developed with Nils Tippenhauer)

Cohort 1: TT7&8 (1.409-10)

Cohort 2: TT24&25 (2.503-4)

Introduction

- Todays lecture: The internet
  - Brief history
  - Core components
  - Edge devices
  - Future of the Internet

Internet – Societal Impact

An Enabler of societal change

- Easy access to knowledge
- Electronic commerce
- Personal relationships
- Discussion without censorship

Internet – Economic impact

An engine of economic growth

- Advertising-sponsored search
- Online stores
- Online B2B marketplaces, eBay
- Crowdsourcing!

The Internet

Brief history

- Arpanet was first packet-switched network ('69-'90)
  - US military project, built for robustness and performance
  - Email ('71), FTP ('73), telnet ('73)
  - Not running TCP/IP (introduced '83)
  - DNS ('83)
- But no HTTP, no websites! Usenet ('80) as precursor of forums
  - IRC ('88) as chat protocol, still in use
- Academic researchers at CERN started the WWW ('89)
  - HTTP, HTML, first web server to share information
- Netscape ('94), flash ('95), mp3 ('95)
- P2P file-sharing started with Napster ('99)
- Skype as first usable major VoIP ('03)
- Ajax ('04) made website much more dynamic
- Mobile internet with rise of smartphones ('07+)

Comment on legacy systems

- The internet core is quite conservative
  - Up to layer 4, very limited changes in the last 20 years
- Exciting new services are almost always introduced on the edges
- 15 years ago, 90% of technology was the same (only physical layer has become a bit faster)
- It is unlikely that TCP/IP will be replaced any time soon
  - IPv6 was proposed 1981, still not widely in use …

What defines the Internet?

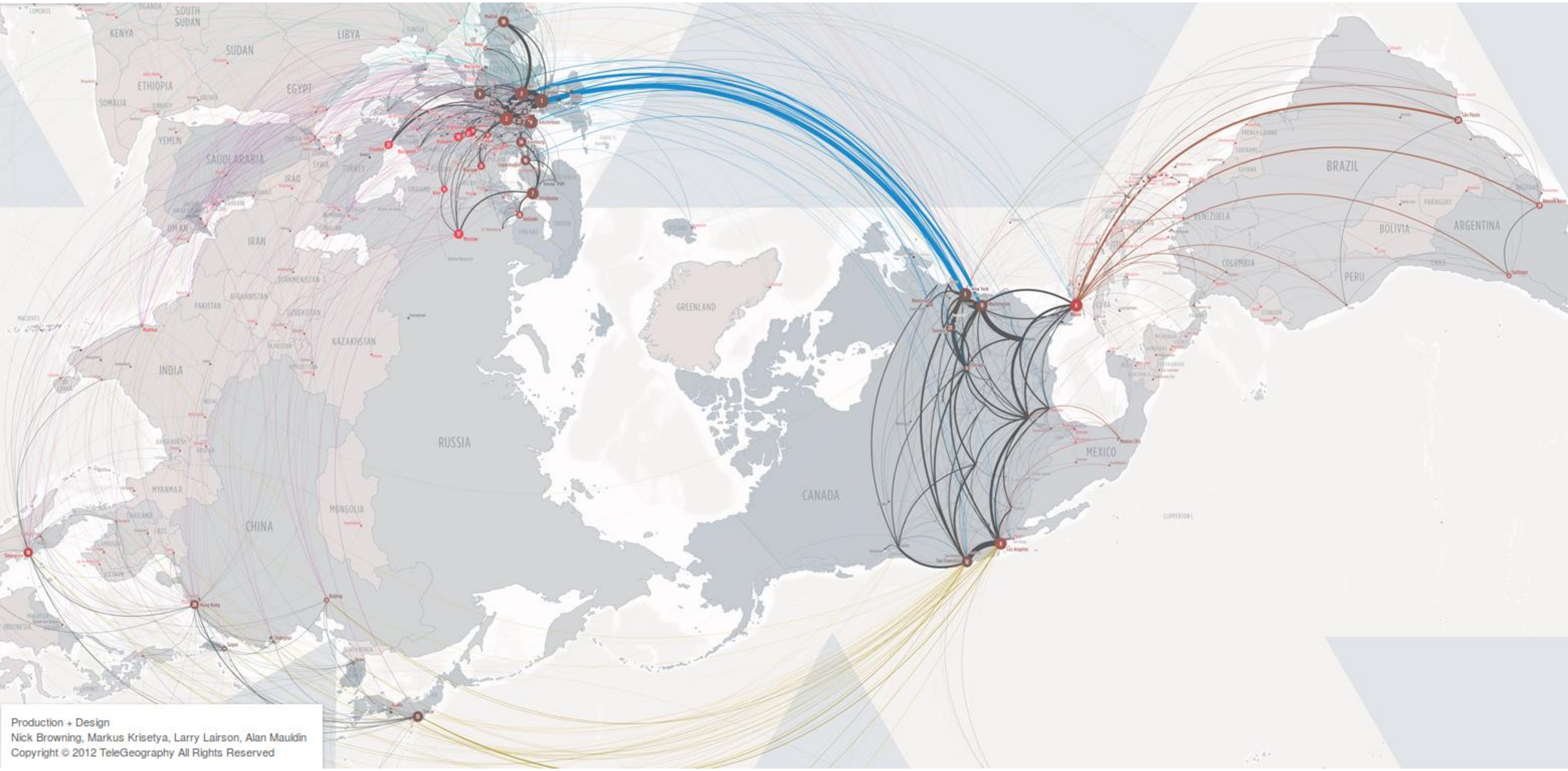
The Internet is

- A collection of connected autonomous systems (AS)
- The language/API between systems is IP (L3)+ UDP/TCP (L4)
- The Internet is more than the WWW
- The Internet is not every world-wide network

Autonomous Aystem (AS) :

- a collection IP routing prefixes
- under one or more network operators
- on behalf of a single administrative entity or domain
- a common, clearly defined routing policy to the Internet.

How is the Internet connected?



http://global-internet-map-2012.telegeography.com

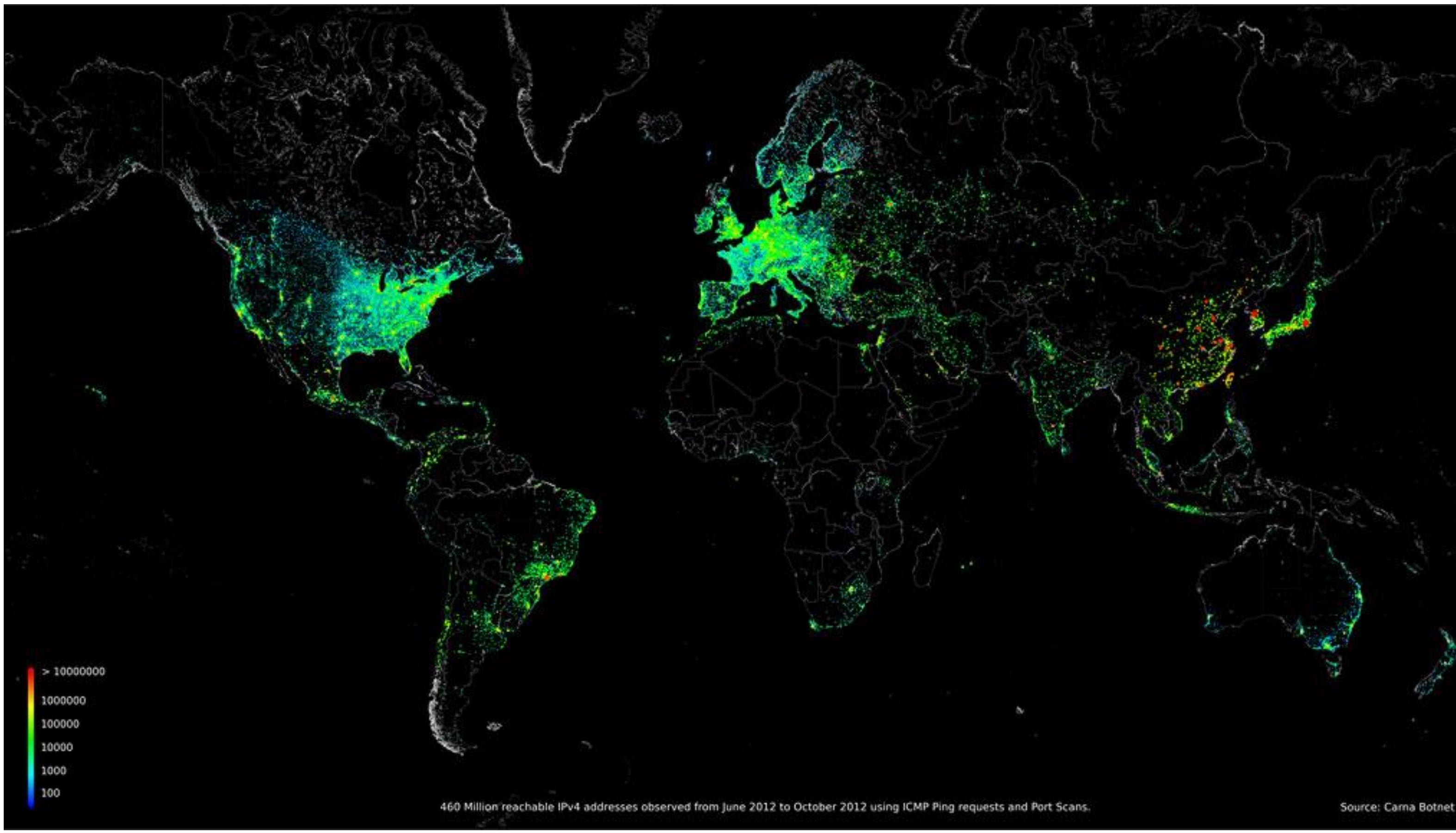
Activity 1 – Size and Growth of the Internet

According to one source there are fifty billion nodes predicted to be on the Internet by the year 2020. (See https://www.electronicsweekly.com/news/business/information-technology/fifty-billion-internet-nodes-predicted-by-2020-2013-01/)

It is estimated that the Internet is doubling approximately every 18 months. What is the estimated size of the Internet (in number of nodes) in the year 2029? Do you think this estimate is reasonable? Why or why not?

Mapping the Internet

- How many devices are connected to the internet?
  - IPv4 provides ~4.3 Billion addresses
  - More devices can be connected via NAT'ing
- Internet-wide ping scans have found ~460 Million reachable IPs

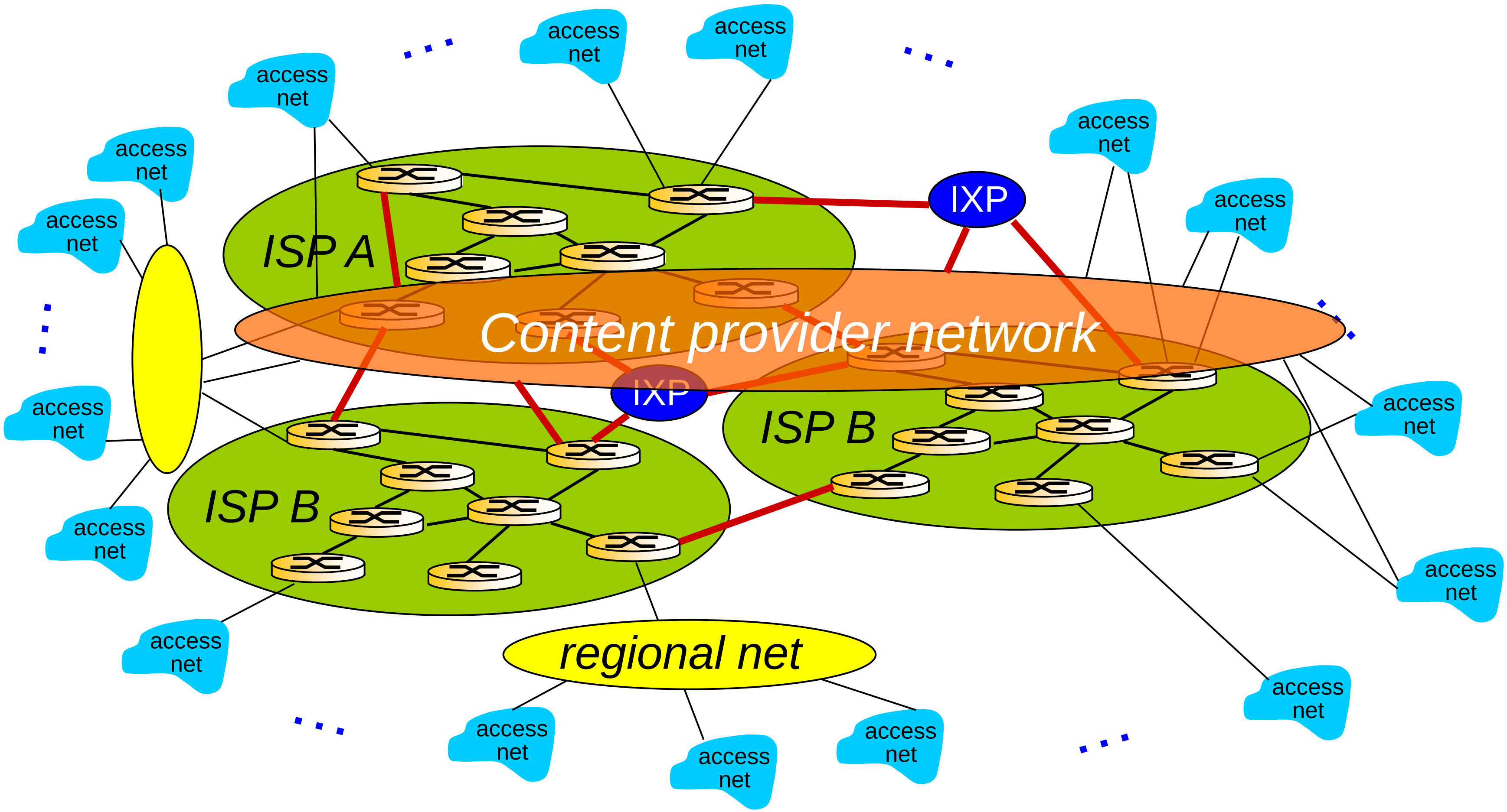


Internet Service Providers

- The internet core is consisting of Internet Service Providers and large corporate networks
- The ISPs forming the core are also called Tier 1 ISPs
- A country typically has several Tier 1 ISP, Singapore has 3:
  - Singtel
  - Starhub
  - Pacific Internet
- Below Tier 1, there are intermediate ISPs, and then Access ISP
- Typically, data exchange with higher tier ISPs will be billed
  - So Tier 1 ISPs never pay other ISPs for forwarded traffic
- Direct Data exchange with same-tier ISPs will be free
  - But lower tier ISPs/customers have to pay to access higher tiers

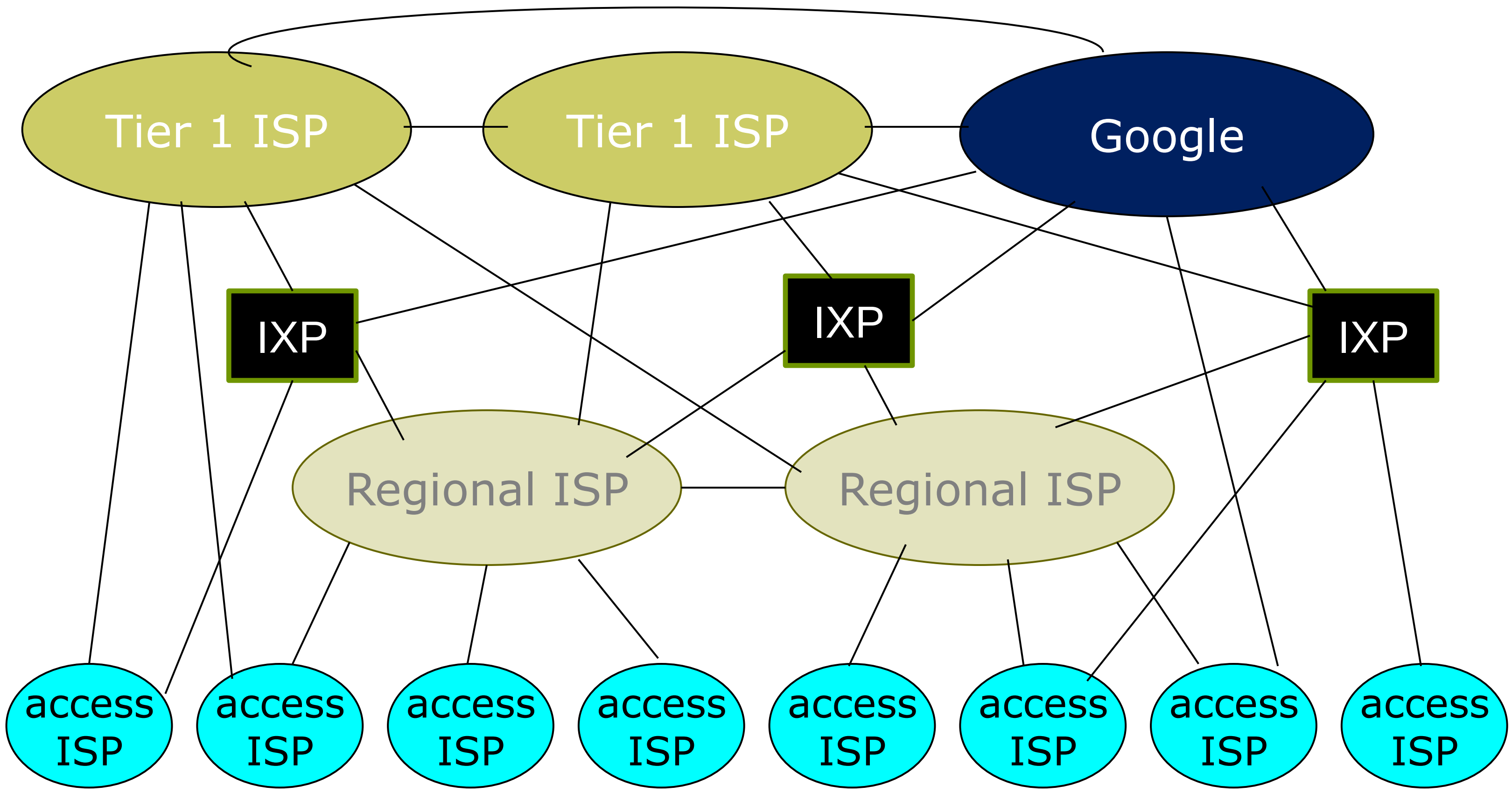
Internet structure: network of networks

… and content provider networks (e.g., Google, Microsoft, Akamai) may run their own network, to bring services, content close to end users



Source: Kurose & Ross Chapter 1 slide set

Internet structure: network of networks

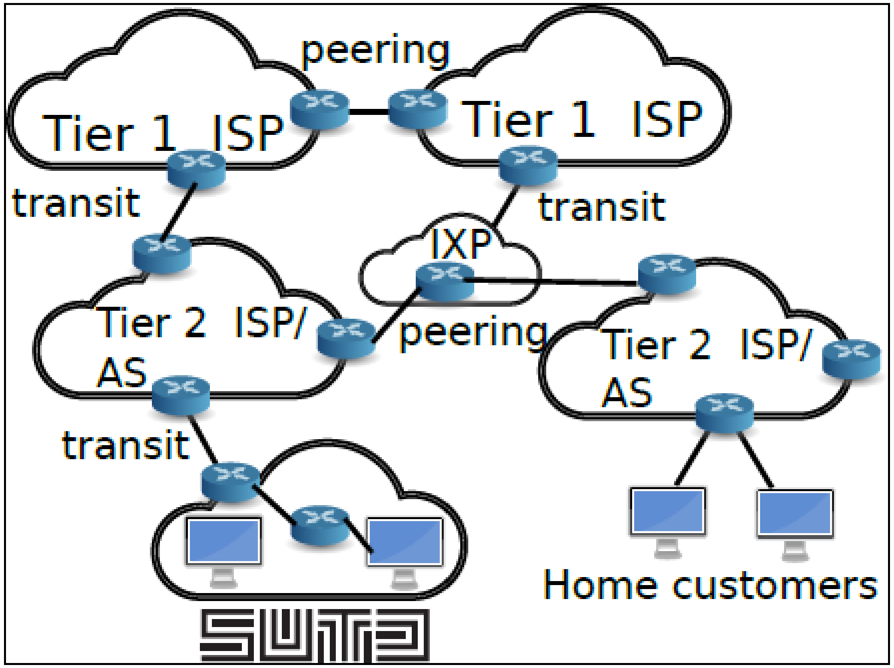


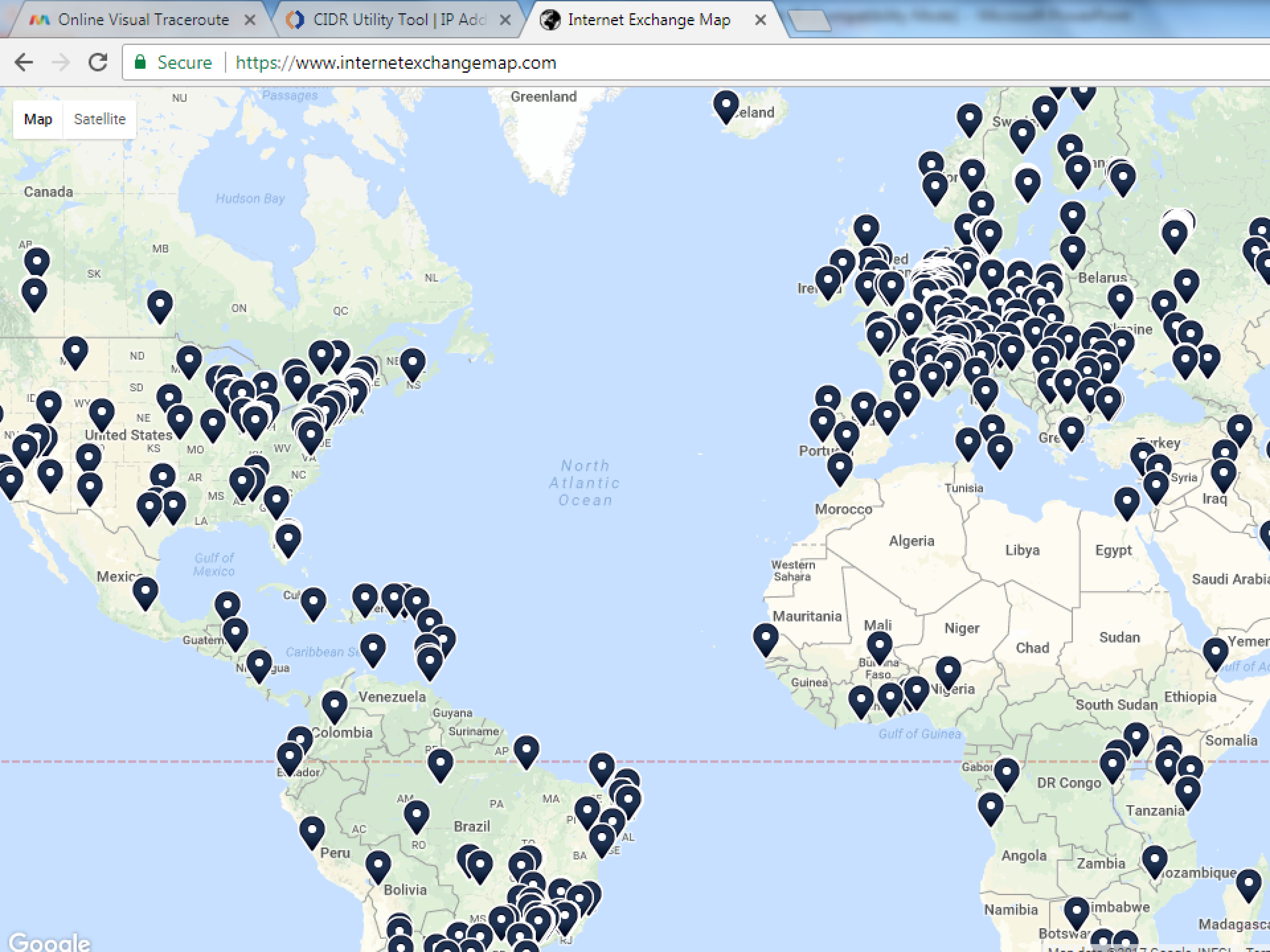
- at center: small # of well-connected large networks
  - "tier-1" commercial ISPs (e.g., Level 3, Sprint, AT&T, NTT), national & international coverage (*tier 1* represents the *largest* global ISPs)
  - content provider network (e.g, Google): private network that connects it data centers to Internet, often bypassing tier-1, regional ISPs

Source: Kurose & Ross Chapter 1 slide set

Internet Exchange Points

- Internet Exchange Points (IXPs) provide fast connections to Tier1 ISPs and important backbone connections
- They are often used to connect multiple Tier 1 ISPs together
- Closest large one to Singapore: Hong-Kong (~200 participants/200+GBit/s)
- http://www.internetexchangemap.com/




Online Visual Traceroute
CIDR Utility Tool | IP Add
Internet Exchange Map
Secure
https://www.internetexchangemap.com
Map
Satellite
Greenland
Canada
Hudson Bay
United States
Gulf of Mexico
Mexico
Caribbean Sea
Venezuela
Colombia
Guyana
Suriname
Peru
Brazil
Bolivia
North Atlantic Ocean
Belarus
Morocco
Algeria
Libya
Egypt
Tunisia
Western Sahara
Mauritania
Mali
Niger
Chad
Sudan
Saudi Arabia
Iraq
Syria
Ethiopia
South Sudan
Somalia
DR Congo
Tanzania
Angola
Zambia
Namibia
Gulf of Guinea
Guinea

Design Challenge: Domain names

- IP addresses are not very human friendly
- They also directly contain network-specific parts, so they can change when a service is relocated
- Domain Names offer a convenient abstraction
  - Strings, so often human-readable and easy to remember
  - Higher layer of abstraction, don’t change when location of service changes
- But how can we route traffic to a domain name?
- The Domain Name Service translates from names to IP addresses
  - First, client contacts DNS server for IP of domain
  - Then, traffic is sent to IP address

DNS Hierarchy

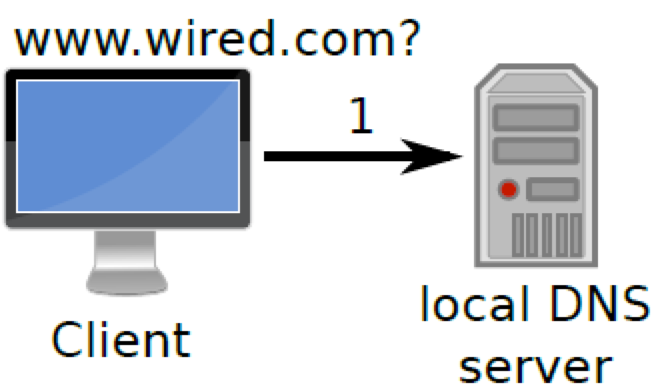
- DNS to IP assignments can change frequently
- One single DNS server for the internet is not going to work
- How do we know which server to connect to?
  - DNS names are hierarchical: thirdLevel.secondLevel.topLevel
  - So, .com, .sg, are top level domains
  - For each, there is a DNS server that has either:
    - A direct mapping to an IP
    - A reference to a lower level DNS server
- Example: foo.bar.com
  - .com server might not have entry
  - Will ask bar.com domain DNS server, that has "foo" entry

DNS bootstrapping

- But how to find the .com DNS server?
- Every network will have a dedicated DNS server
- DNS works recursively, i.e. queries are forwarded
- Local server handles local lookups, or forward to higher servers

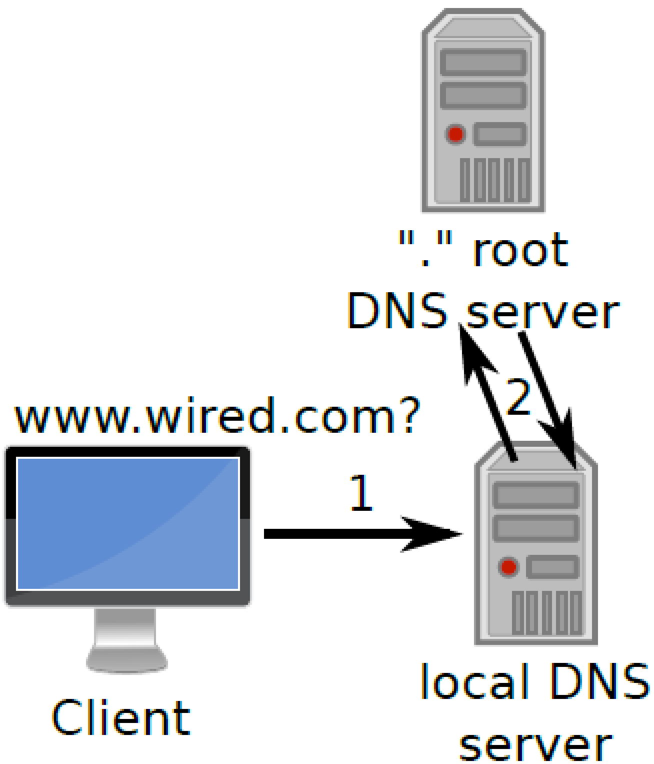
DNS bootstrapping

- But how to find the .com DNS server?
- Every network will have a dedicated DNS server
- DNS works recursively, i.e. queries are forwarded
- Local server handles local lookups, or forward to higher servers



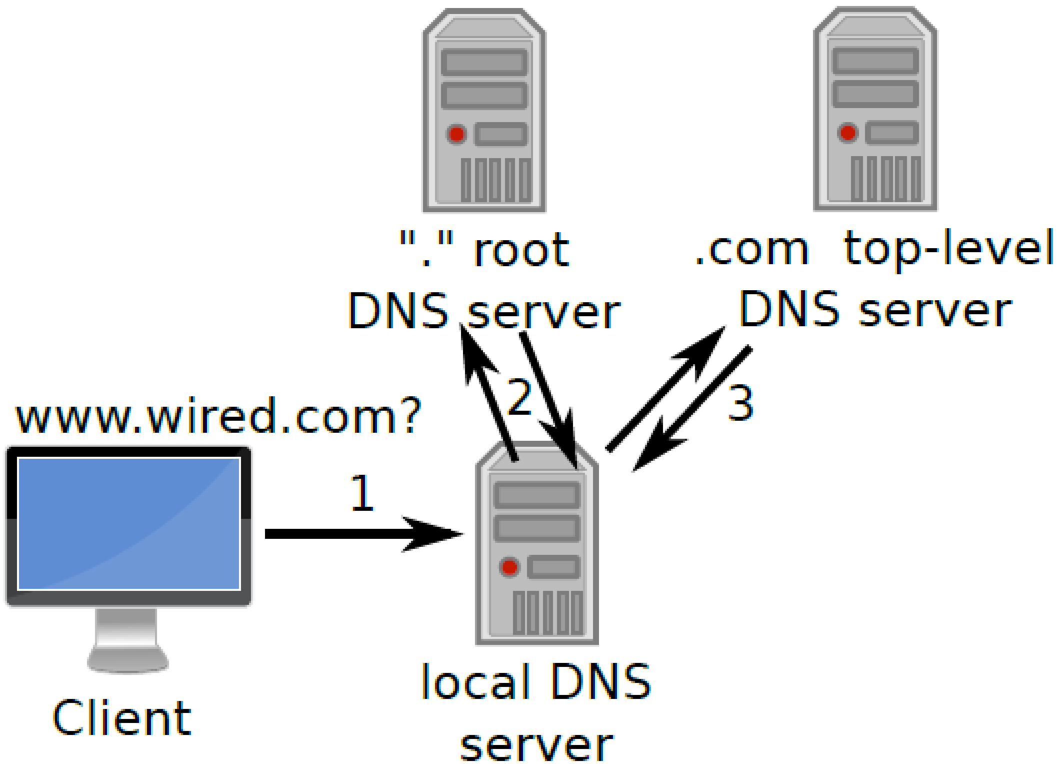
DNS bootstrapping

- But how to find the .com DNS server?
- Every network will have a dedicated DNS server
- DNS works recursively, i.e. queries are forwarded
- Local server handles local lookups, or forward to higher servers



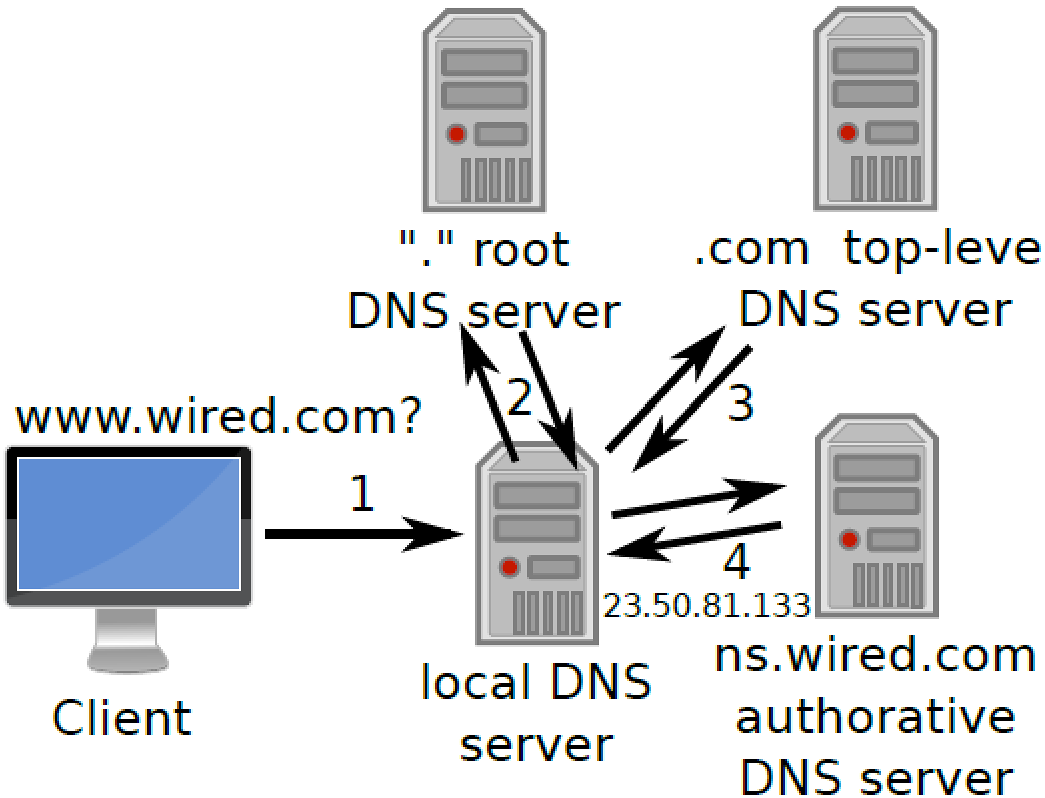
DNS bootstrapping

- But how to find the .com DNS server?
- Every network will have a dedicated DNS server
- DNS works recursively, i.e. queries are forwarded
- Local server handles local lookups, or forward to higher servers



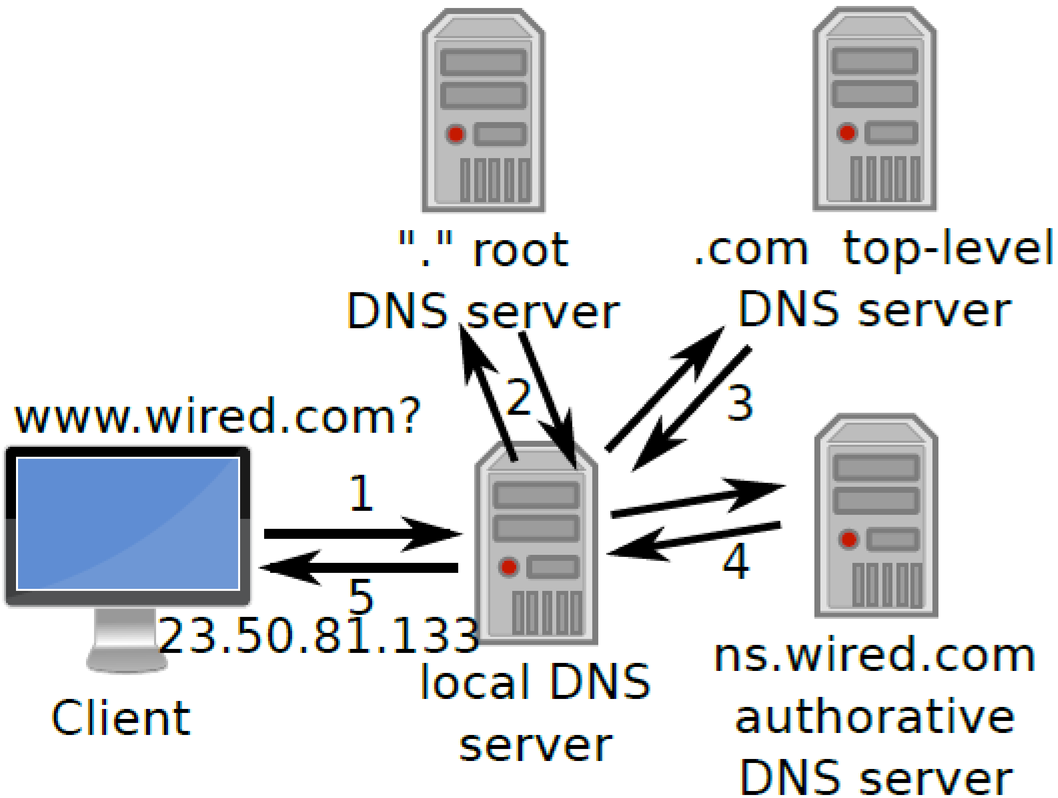
DNS bootstrapping

- But how to find the .com DNS server?
- Every network will have a dedicated DNS server
- DNS works recursively, i.e. queries are forwarded
- Local server handles local lookups, or forward to higher servers



DNS bootstrapping

- But how to find the .com DNS server?
- Every network will have a dedicated DNS server
- DNS works recursively, i.e. queries are forwarded
- Local server handles local lookups, or forward to higher servers



DNS Record Types

- A – Mapping of name to IPv4 Address(es)
- AAAA – Mapping of name to IPv6 Address(es)
- MX – Mail Exchange for name
- CNAME – Canonical name (alias)
- TXT – Text record associated with name
- NS – Name server responsible for domain

Design Challenge: Addressing

Consider the following setting:

- N nodes are connected in network
- Each has unique address
- Every node can reach every other through multiple hops
- Highly connected nodes form "core", edge is sparsely connected
- What addressing scheme should be chosen to enable quick routing to edges?

Analogy – PSTN (or POTS)

In the United States, for example, there are 10-digit phone numbers.

- The first three digits are the **area code** or **national destination code (NDC)**, which helps route the call to the right regional switching station.
- The next three digits are the **exchange**, which represents the smallest amount of circuits that can be bundled on the same switch. In other words, when you make a call to another user in your same exchange -- maybe a neighbor around the corner -- the call doesn't have to be routed onto another switch.
- The last four digits of the phone number represent the **subscriber number**, which is tied to your specific address and phone lines.

Visual Tracepath

`http://www.monitis.com/traceroute`

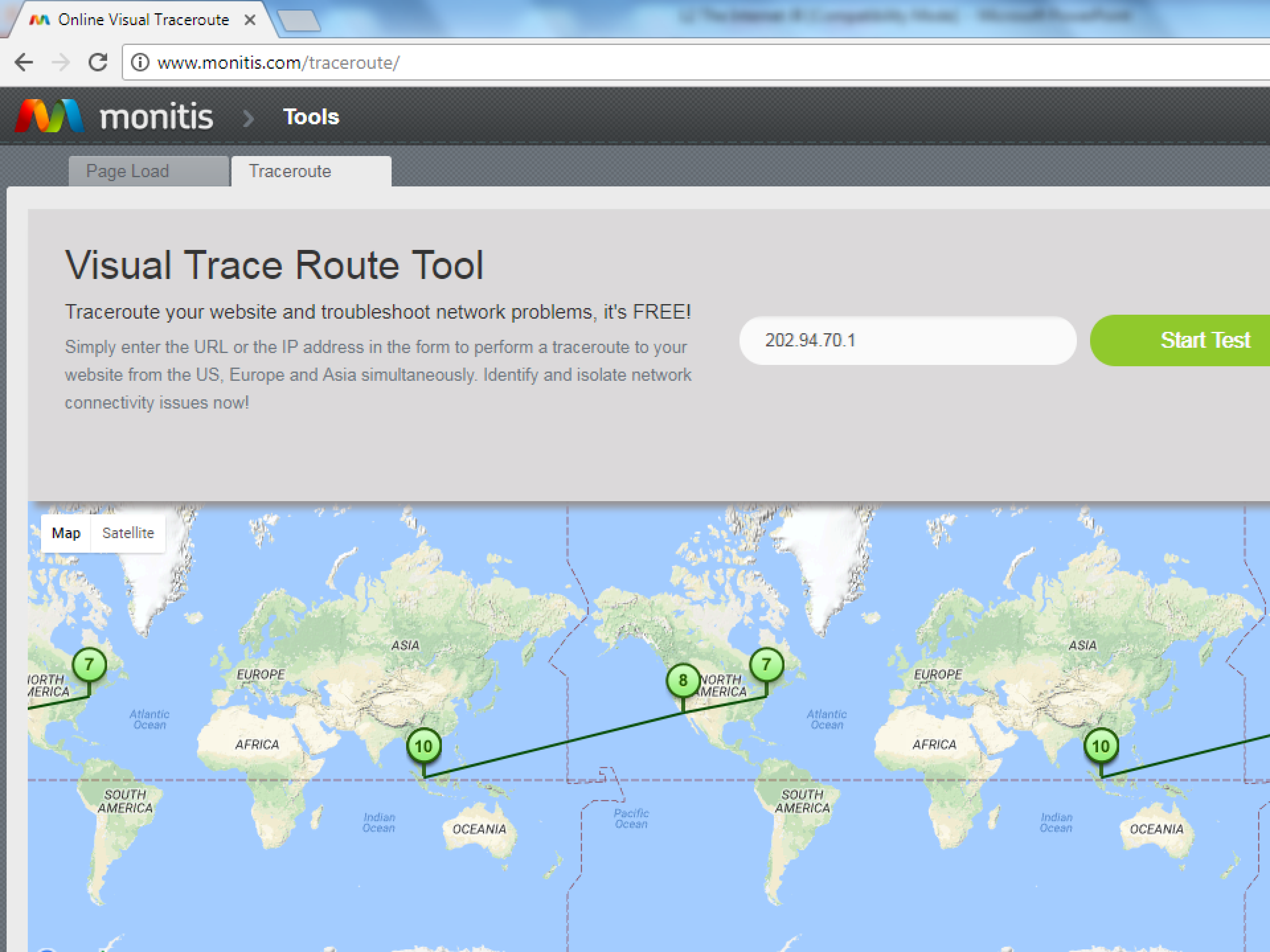
- Example: 202.94.70.1 from US/Europe
  - Both routes seem to go via North America?

Activity 2 – DNS Lookup

Find out information about the company abracadabra.com using mxlookup, monitis or software on your computer such as dig / nslookup. Find out the following details:

- IP address
- Using the IP Address use mxlookup to do an Reverse DNS lookup – do you get the same domain name? Can you explain your answer?
- Which is their primary DNS server
- Which is their authoritative name server?
- Could you find more than one? How can you explain it?
- Find out the ASN of the AS that the domain belongs to
- Find out email gateway(s) if possible
- Find out security related information (keys, Hashing algorithms etc.)

Submit on eDimension


Online Visual Traceroute
www.monitis.com/traceroute/
monitis
Tools
Page Load
Traceroute
Visual Trace Route Tool
Traceroute your website and troubleshoot network problems, it's FREE!
Simply enter the URL or the IP address in the form to perform a traceroute to your website from the US, Europe and Asia simultaneously. Identify and isolate network connectivity issues now!
202.94.70.1
Start Test
Map
Satellite
ASIA
EUROPE
AFRICA
NORTH AMERICA
SOUTH AMERICA
OCEANIA
Atlantic Ocean
Indian Ocean
Pacific Ocean
7
8
10

Online Visual Traceroute | Network Tools: DNS,IP,Er

Secure | https://mxtoolbox.com/SuperTool.aspx?action=ptr%3a202.94.70.1&run=toolpage

MX TOOLBOX®

MX Lookup | Blacklists | Diagnostics | Domain Health | Analyze Headers | Free Monitoring | Investigator | DNS Lookup

SuperTool Beta7

202.94.70.1 | Reverse Lookup

ptr:202.94.70.1 Find Problems | ptr | Share Resu

	Test	Result	
✖	DNS Record Published	DNS Record not found	More Inf

smtp diag | blacklist | port scan | subnet tool | dns propagation

Reported by **secdns2.starhub.net.sg** on 9/11/2017 at **7:42:54 PM (UTC 0)**, just for you. (History) Trans

ABOUT THE SUPERTOOL!

All of your MX record, DNS, blacklist and SMTP diagnostics in one integrated tool. Input a **domain name** or **IP Address** or **Host Name.** Links in the results will guide you to other relevant tools and information. And you'll have a chronological history of your results.

If you already know exactly what you want, you can force a particular test or lookup. Try some of these examples:

Global addressing in IPv4

- IP addresses consist of 4 8-bit numbers
  - e.g. 192.168.0.1, 10.0.4.134 (in decimal)
  - In total the 256^4=4.3 Billion addresses mentioned earlier
- How can we find the route from one IP address to another?
  - Store a route for each target IP? 4.3 Billion routes
  - Maybe a central server that stores all routes?
- How would graphs for both cases look like?

Internet Topology

- The actual topology of the Internet is a hybrid between a fully connected graph, and a star graph
- core: few highly connected nodes with many edges
- edge: consists of many poorly connected nodes with only one edge
- Addressing reflects that
  - "Divide and conquer" is used: divide address space into local subgroups (subnets)
  - Routing is done between subnets
  - Within a subnet, routing is only local
  - In general: "left part" of IP address denotes network, "right part“ denotes host

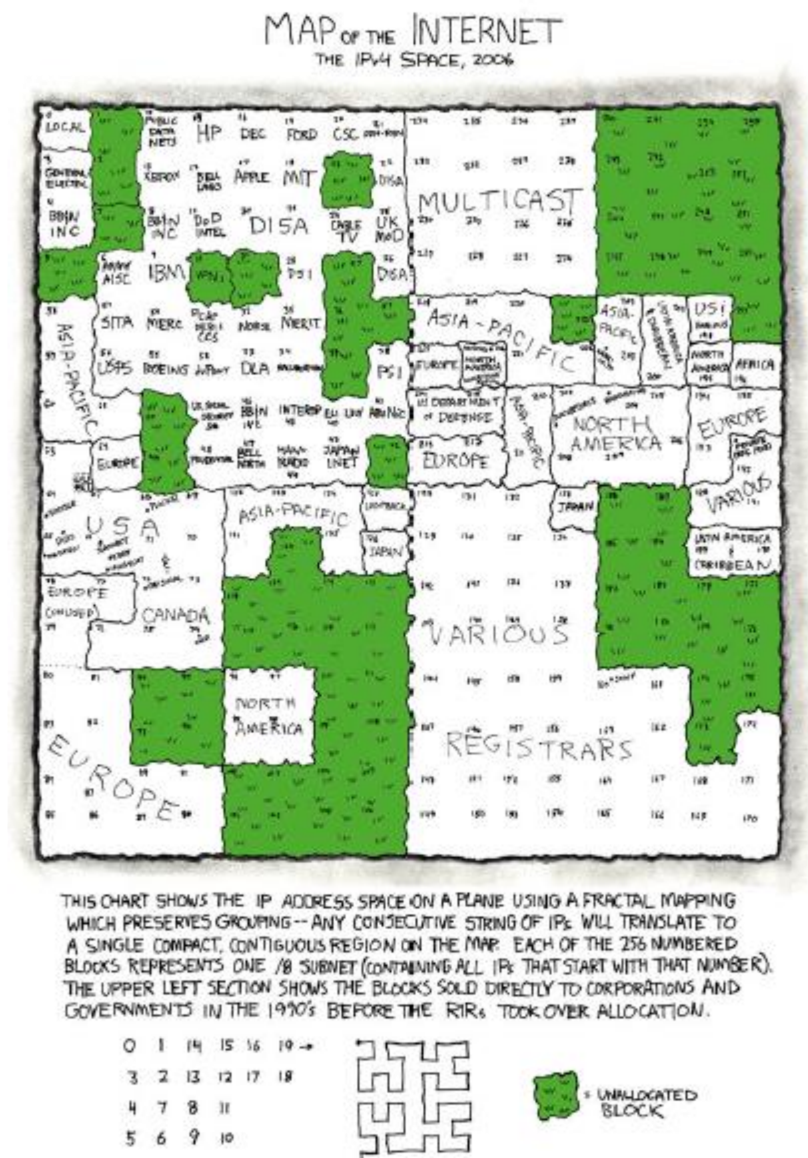
Class-based Subnets

- Originally, there were three possible classes of subnets: A, B, C
  - Up to 256 Class A networks with each 256^3 addresses
  - Up to 256^2 Class B networks with 256^2 addresses
  - Up to 256^3 Class C networks with 256 addresses
  - Compared to our tree view: two layers, nodes in l1 have up to $256^{1|2|3}$ children
- Companies or providers would get a whole Class A,B,C network
- But these classes were not well matched to realistic use:
  - Class C was too small for companies or universities
  - Class B was too big for almost all users (65k addresses)
  - So nowadays, subnets can have 2^n addresses for variable n

Routing between Subnets

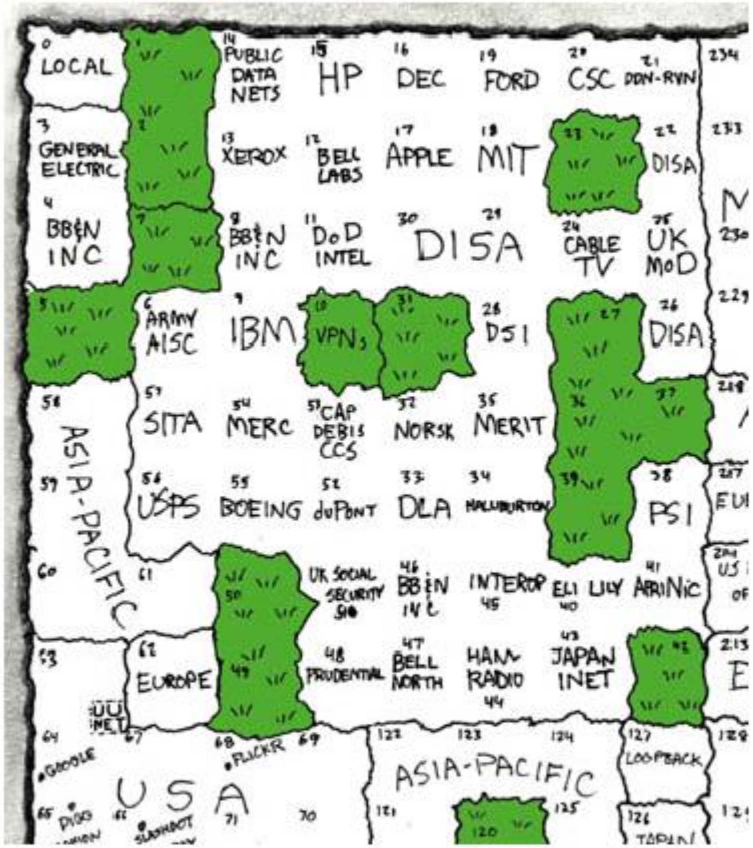
- Routes on the internet are advertised for subnets (prefixes)
  - Subnets always have their host part set to 0 (e.g. 1.2.0.0)
  - "to reach any of 1.2.0.0, come here"
  - To simplify routing, small subnets are aggregated if possible
    - Instead of "to reach 1.2.0.0 or 1.2.1.0 or 1.2.2.0 or 1.2.3.0, come here", you advertise "to reach 1.2.[0-3].0, come here"
    - We will introduce the exact notation for this soon
- If specific routes to target subnet is not available, you can always route towards parent subnet
  - This should get you closer to the target
- Vague similarity to country/area code in phone calls

Map of the Internet



Randall Munroe/XKCD (CC BY-NC 2.5)

Map of the Internet



Randall Munroe/XKCD
(CC BY-NC 2.5)

Private Subnets

- Most IP addresses are globally unique
- Three private range(s) were defined
  - 10.0.0.0 - 10.255.255.255
  - 172.16.0.0 - 172.31.255.255
  - 192.168.0.0 - 192.168.255.255
- These IP addresses will not be globally unique
- Internet edge routers will not forward traffic to these
  - This is why your 192.168.0.1 IP is not reachable from Internet
- To allow private IPs to communicate with Internet:
  - Network Address Translation (NAT) is required (more later)

Classless-Interdomain-Routing (CIDR)

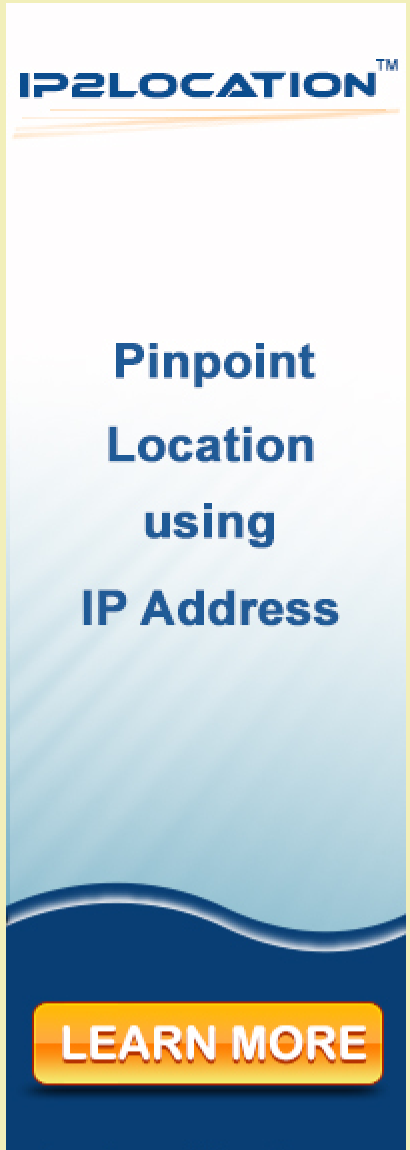
- The CIDR format allows finer-grained division of IP space
- Arbitrary cut-off point between network and individual address
- Allows subnets smaller than Class B, larger than Class C
- Out of the 4x8 bits of an IP address, leftmost n are used to describe the subnet, and 32-n are used for the host
  - Notation: v.x.y.z/n, with n the length of the subnet
  - Example: 1.2.3.4/24 denotes host 4 in subnet 1.2.3.0/24

Link

Online Visual Traceroute | CIDR Utility Tool | IP Add

www.ipaddressguide.com/cidr

Sponsor



CIDR

CIDR is the short for Classless Inter-Domain Routing, an IP addressing schen
the older system based on classes A, B, and C. A single IP address can be us
many unique IP addresses with CIDR. A CIDR IP address looks like a normal
that it ends with a slash followed by a number, called the IP network prefix.
reduce the size of routing tables and make more IP addresses available withi

CIDR to IP range

This tool converts CIDR to IP range.

CIDR Range	10.0.2.1/30
Netmask	255.255.255.252
Wildcard Bits	0.0.0.3
First IP	10.0.2.0
Last IP	10.0.2.3
Total Host	4

CIDR

10.0.2.1/30

Calculate Cancel

IP Range to CIDR

www.ip2location.com/?utm_source=hexasoft&utm_medium=banner&utm_term=ipaddressguide&utm_content=skyscraper&utm_campaign=hexasoft banner

Subnet Masks

- Example: 1.2.65.4/22
  - What is the subnet address?
  - What is the host address?
- To easily find both, create subnet mask with n Ones and 32-n zeros
  - AND this subnet mask with binary representation of IP address
  - Result is the subnet address
  - Invert mask + apply with AND to get host address

```
IP:   0000 0001 0000 0010 0100 0001 0000 0100
Mask: 1111 1111 1111 1111 1111 1100 0000 0000
Net : 0000 0001 0000 0010 0100 0000 0000 0000
```

So subnet is 1.2.64.0/22, host part is 01 0000 0100, i.e. the 260th host in the sub-range.

Broadcasting

- What if you want to send a message to everyone on the local network?
  - For n recipients, you send n messages? Inefficient!
  - We need a target address that everyone listens to!
- The broadcast address is the highest address in the subnet
  - So, for 192.168.0.0/24 the broadcast address is
    - 192.168.0.255
  - While for 192.168.0.0/23 the broadcast address is
    - 192.168.1.255

Other reserved addresses

- Multicast: 224.0.0.0/4 (various uses, HUGE range)
- Loopback: 127.0.0.0/8 (interface to local machine)

Activity 3: Subnet and host addressing

If I want just one device to be on the network, can I set the subnet mask to 255.255.255.255? For example, consider the following configuration:

IP address 10.0.0.10
Subnet mask 255.255.255.255
Gateway 10.0.0.1

10.0.0.10 will be the only device in the network talking to the 10.0.0.1 gateway? Is there a problem with this configuration? If no, explain. If yes, can you fix the problem.

Outlook: IPv6

- The 4.3 Billion IP addresses are actually not enough
  - (Also: .0 and .255 and others can't be used, there are slightly fewer addresses)
- What limits number of addresses? length (32 bit)
- IPv6 increases length to 128 bit
- More optimizations "under the hood", more on that later
- Addresses are not as easy to remember any more
  - To simplify, use addresses with Zero fields
  - To distinguish from IPv4 Addr: use ":" instead of "."
  - "::" denotes a sequence of Zero fields

The Future of the Internet?

On lower layers:

- Easy to predict: more devices!
- Higher bandwidth requirements
- More mobile users, so more wireless PHY

On application layer:

- AaaS, Iaas, Cloud, PaaS, web2.0
- Smart automation, M2M communication

Homework 2 – IP addressing (due 16th Sept 23:59)

Q1. The subnet mask for a particular network is 255.255.31.0 Which of the following pairs of IP addresses could belong to this network ?

(a) 172.57.88.62 and 172.56.87.23.2

(b) 10.35.28.2 and 10.35.29.4

(c) 191.203.31.87 and 191.234.31.88

(d) 128.8.129.43 and 128.8.161.55

Conclusion

- We discussed Tier 1 and access ISPs, and how they are connected
- We revised how DNS and DNS queries work
- The length of the subnet part of the IP address is indicated in bits after a slash:
  - 10.0.2.1/30 = first 30 bits (of 32) are subnet part, 4 local IPs
  - 10.0.2.1/24 = first 24 bits (of 32) are subnet part, 256 local IPs
- 10.0.2.1/24 identifies a host+netmask, 10.0.2.0/24 is a network