Master Course
Computer Networks
IN2097

Prof. Dr.-Ing. Georg Carle

Chair for Network Architectures and Services
Department of Computer Science
Technische Universität München
http://www.net.in.tum.de

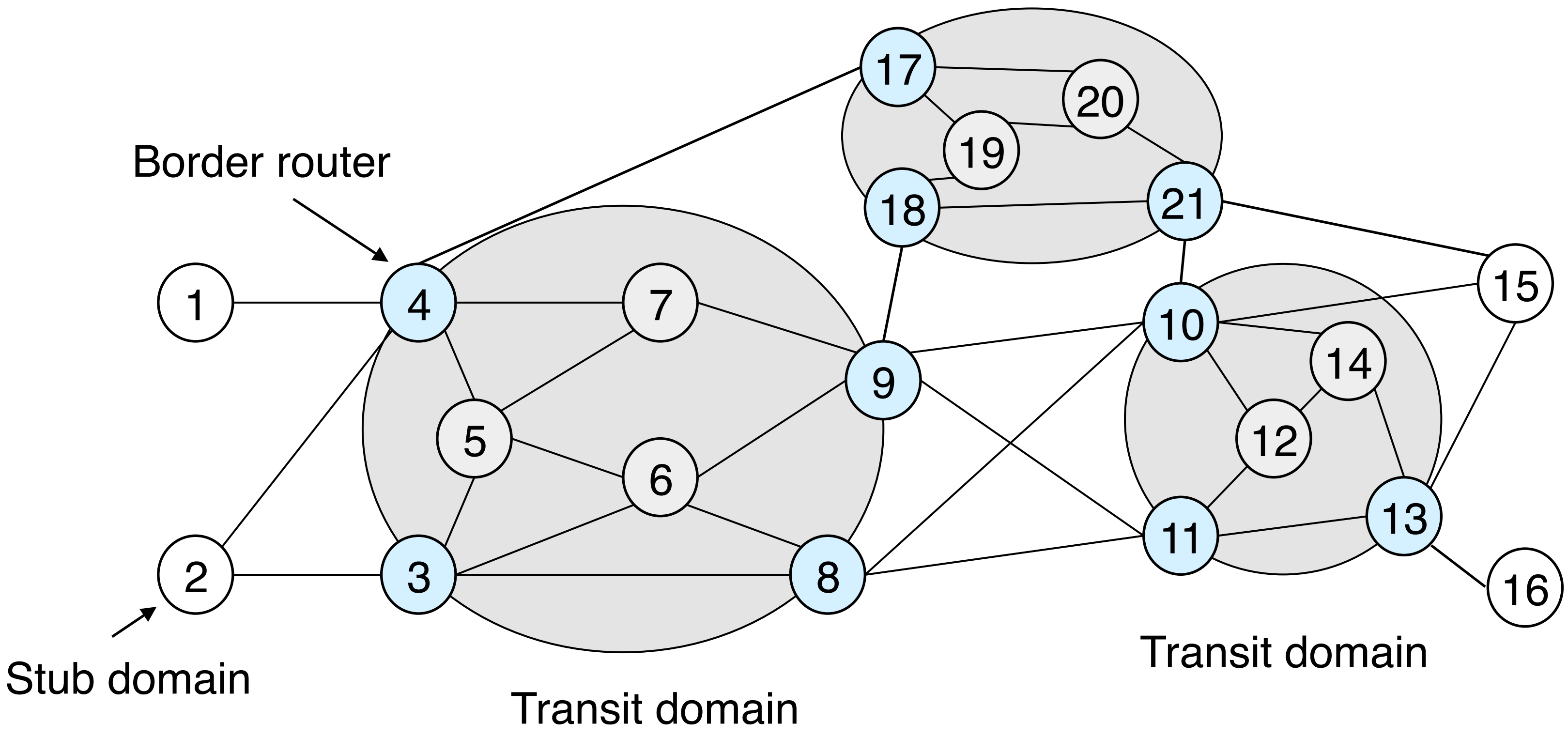
Chair for Network Architectures and Services – Prof. Carle
Department of Computer Science
TU München

Internet Protocol

Technische Universität München

Multi-domain Structure of the Internet

- Hierarchical network structure



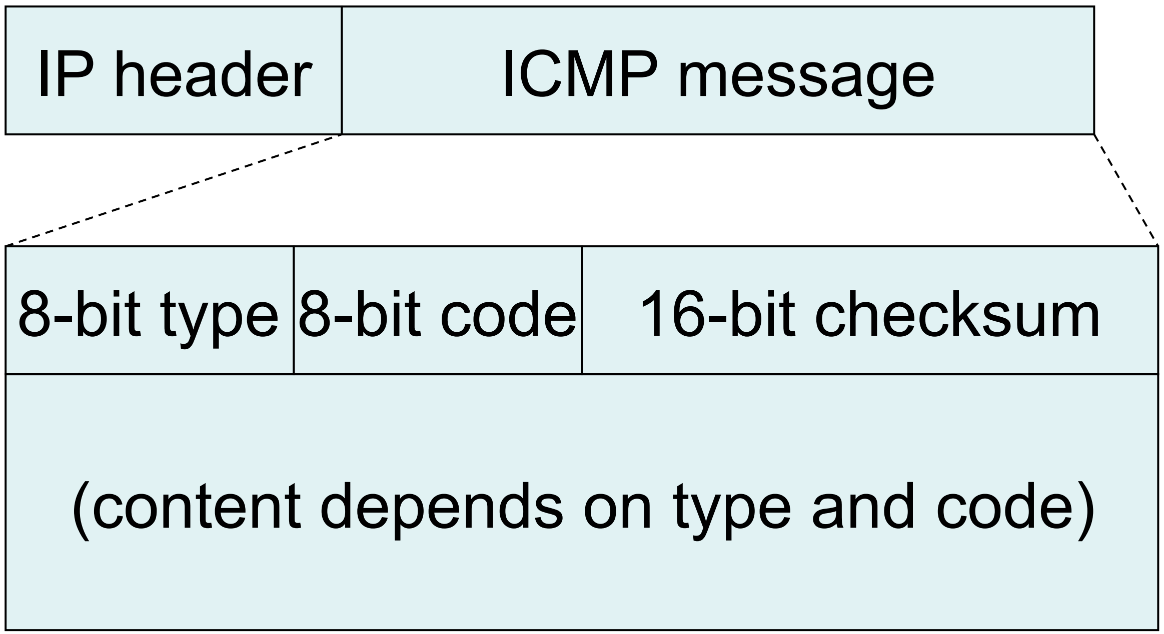
Border router: routers with direct connectivity to other domains
Stub domain: domain that originates or sinks traffic
Transit domain: domain that forwards transit traffic

ICMP

- Internet Control Message Protocol (ICMP)
  - C.f. RFC 792
  - Can be considered part of the IP layer.
  - Communicates error messages and other conditions that require attention
  - Error messages are acted on by either...
    - IP layer, or
    - TCP, or UDP
  - Some ICMP messages cause errors to be returned to user processes
  - network-layer "above" IP:
    - ICMP messages carried in IP datagrams

ICMP Message Format

- 15 different *types*.
- Some types use a *code* to further specify the condition.



ICMP Message Types

- Two classes of ICMP messages
  - Query messages
    - Only kind of ICMP messages that generate another ICMP message.
  - Error messages
    - Contain IP header and first 8 bytes of datagram that caused the ICMP error to be generated.
      - Allows receiving ICMP module to associate the message with a particular protocol and process (port number).

ICMP: Internet Control Message Protocol

0	echo reply (ping)
3	destination unreachable
4	source quench
5	**redirect**
8	echo request
9	**router advertisement**
10	**router solicitation**
11	time exceeded
12	parameter problem (bad IP header)
13	timestamp request
14	timestamp reply
15	info request
16	info reply
17	address mask request
18	address mask reply

ICMP: Internet Control Message Protocol

Type	Code	description
3	0	dest. network unreachable
3	1	dest host unreachable
3	2	dest protocol unreachable
3	3	dest port unreachable
3	6	dest network unknown
3	7	dest host unknown

- ICMP contents always contain IP header and first 8 bytes of IP contents that caused ICMP error message to be generated

20-byte standard IP header	8 bit ICMP type	8 bit ICMP code	16-bit checksum	contents of ICMP msg

Ping

- checks if host is reachable, alive
- uses ICMP echo request/reply
- copy packet data request → reply

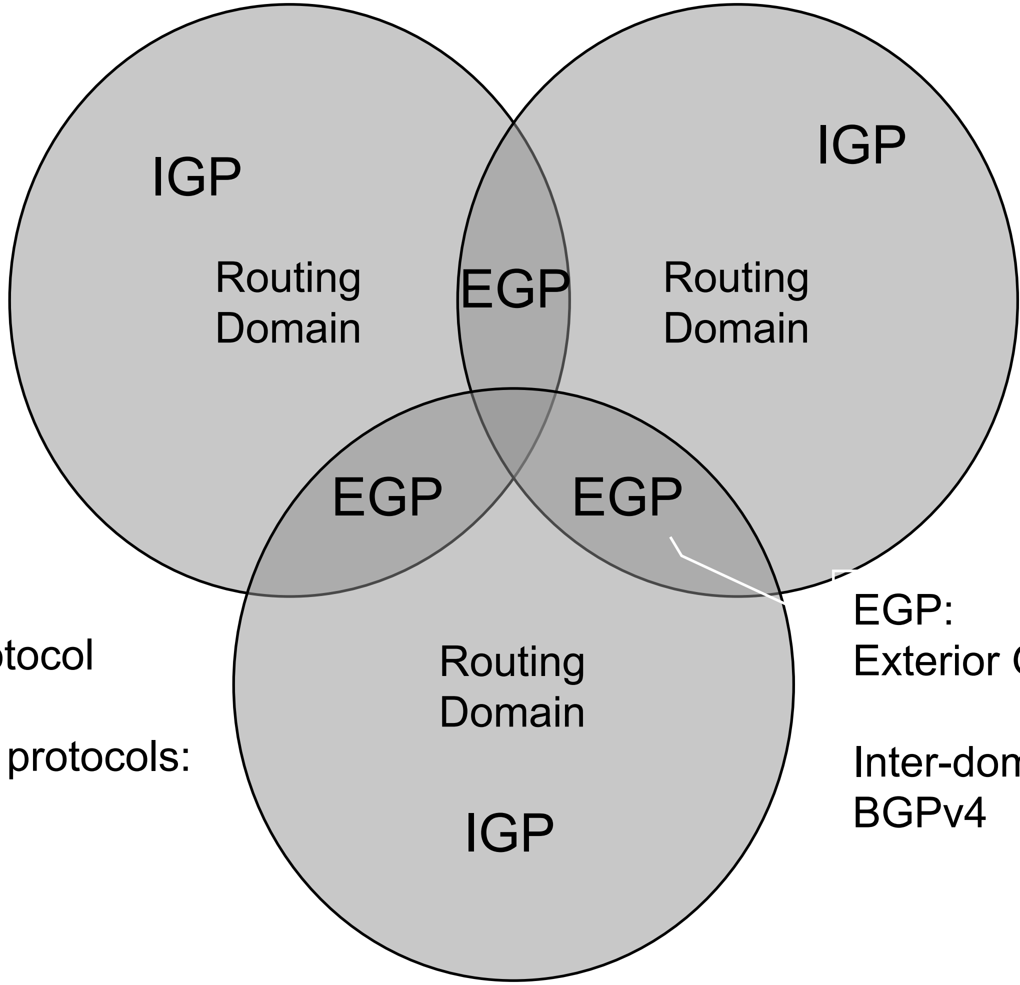
```
ping -s gaia.cs.umass.edu
PING gaia.cs.umass.edu: 56 data bytes
64 bytes from gaia.cs.umass.edu (128.119.40.186): icmp_seq=0
   time=276 ms
64 bytes from gaia.cs.umass.edu (128.119.40.186):
   icmp_seq=1 time=281 ms
64 bytes from gaia.cs.umass.edu (128.119.40.186): icmp_seq=2
   time=276 ms^C
----gaia.cs.umass.edu PING Statistics----
4 packets transmitted, 3 packets received, 25% packet loss
(ms) min/avg/max = 276/277/281
```

Traceroute

- allows to follow path taken by packet
- send UDP packets with increasing TTL to unlikely port
- ICMP replies: ’time exceeded’; last ICMP message: ’port unreachable’

```
$ traceroute gaia.cs.umass.edu
1 gmdbgate (192.35.149.248) 6 ms 2 ms 2 ms
2 188.1.132.142 (188.1.132.142) 263 ms 178 ms 188 ms
3 gmdisgate.gmd.de (192.54.35.68) 153 ms 187 ms 151 ms
4 icm-bonn-1.gmd.de (192.76.246.17) 226 ms 207 ms 242 ms
5 icm-dc-1-S2/6-512k.icp.net (192.157.65.209) 320 ms 315 ms 393 ms
6 icm-mae-e-H1/0-T3.icp.net (198.67.131.9) 372 ms 297 ms 354 ms
7 mae-east (192.41.177.180) 456 ms 537 ms 401 ms
8 borderx.Washington.mci.net (204.70.74.117) 529 ms 385 ms 340 ms
9 core-fddi-1.Washington.mci.net (204.70.3.1) 437 ms 554 ms 581 ms
10 core-hssi-3.NewYork.mci.net (204.70.1.6) 418 ms 547 ms 492 ms
11 core-hssi-3.Boston.mci.net (204.70.1.2) 453 ms 595 ms 724 ms
12 border1-fddi-0.Boston.mci.net (204.70.2.34) 789 ms 404 ms 354 ms
13 nearnet.Boston.mci.net (204.70.20.6) 393 ms 323 ms 346 ms
14 mit3-gw.near.net (192.233.33.10) 340 ms 465 ms 399 ms
15 umass1-gw.near.net (199.94.201.66) 557 ms 316 ms 369 ms
16 lgrc-gw.gw.umass.edu (192.80.83.1) 396 ms 309 ms 389 ms
17 cs-gw.cs.umass.edu (128.119.44.1) 276 ms 490 ms 307 ms
18 gaia.cs.umass.edu (128.119.40.186) 335 ms 317 ms 350 ms
```

Two levels of Routing Protocols



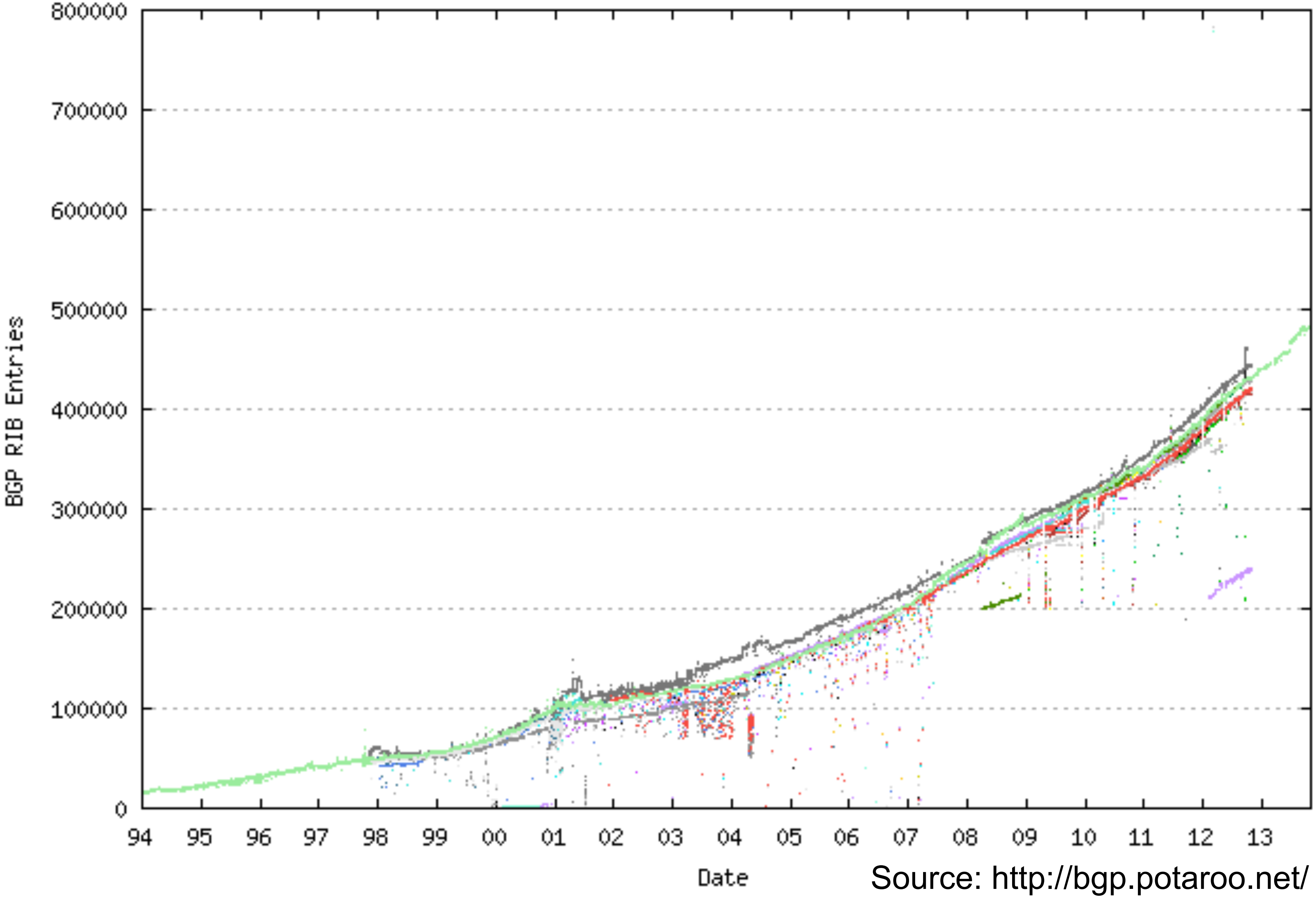
IGP:
Interior Gateway Protocol

Intra-domain routing protocols:
RIP
OSPF
IS-IS

EGP:
Exterior Gateway Protocol

Inter-domain routing protocol:
BGPv4

Number of Prefixes in Forwarding Table



Source: http://bgp.potaroo.net/

IP addresses: how to get one?

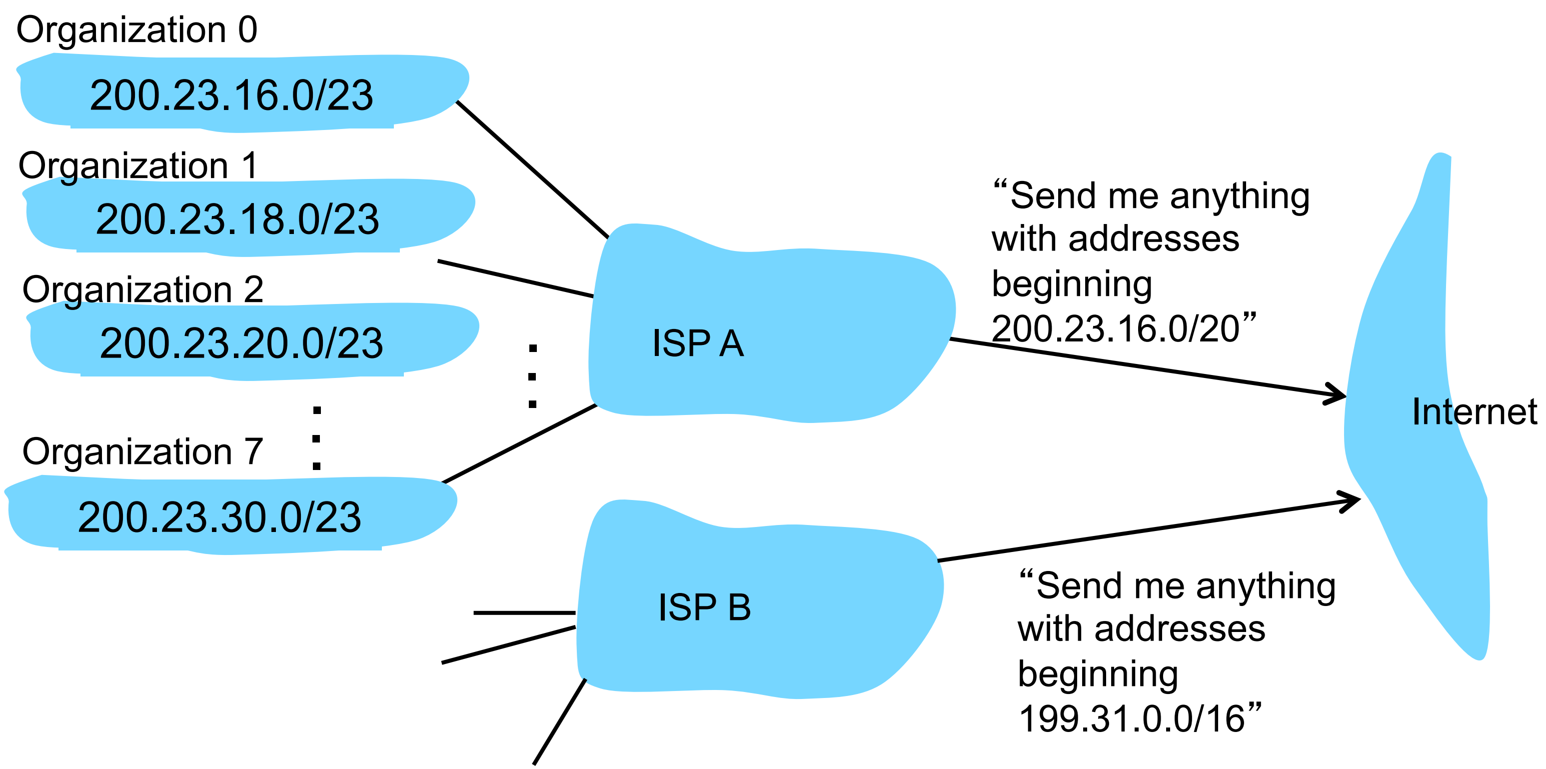
Q: How does *network* get subnet part of IP addr?

A: gets allocated portion of its provider ISP's address space

ISP's block	11001000	00010111	00010000	00000000	200.23.16.0/20
Organization 0	11001000	00010111	00010000	00000000	200.23.16.0/23
Organization 1	11001000	00010111	00010010	00000000	200.23.18.0/23
Organization 2	11001000	00010111	00010100	00000000	200.23.20.0/23
...	…..			….	….
Organization 7	11001000	00010111	00011110	00000000	200.23.30.0/23

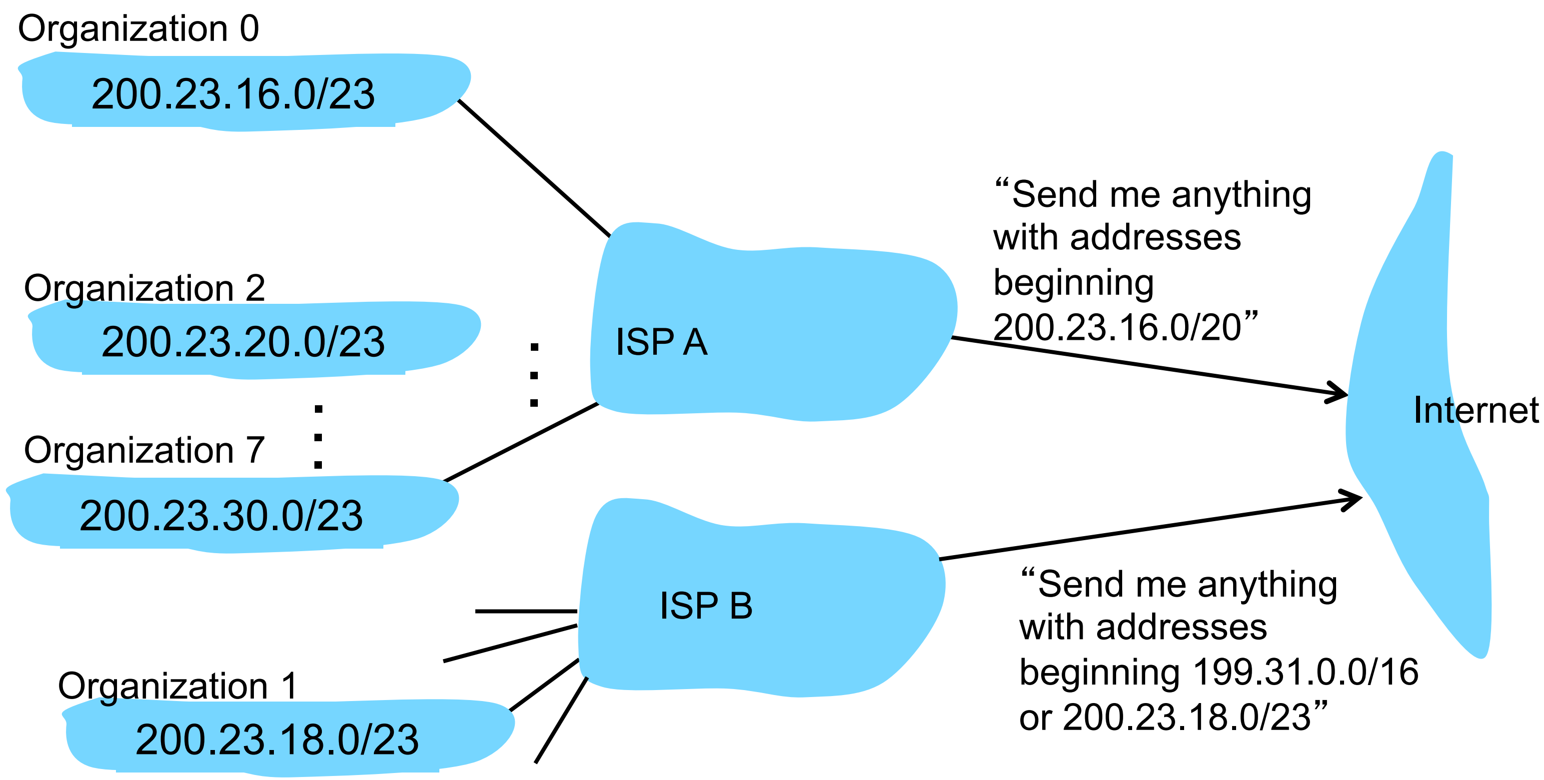
Hierarchical addressing: route aggregation

Hierarchical addressing allows efficient advertisement of routing information:



Hierarchical addressing: more specific routes

ISP B has a more specific route to Organization 1



IP addressing

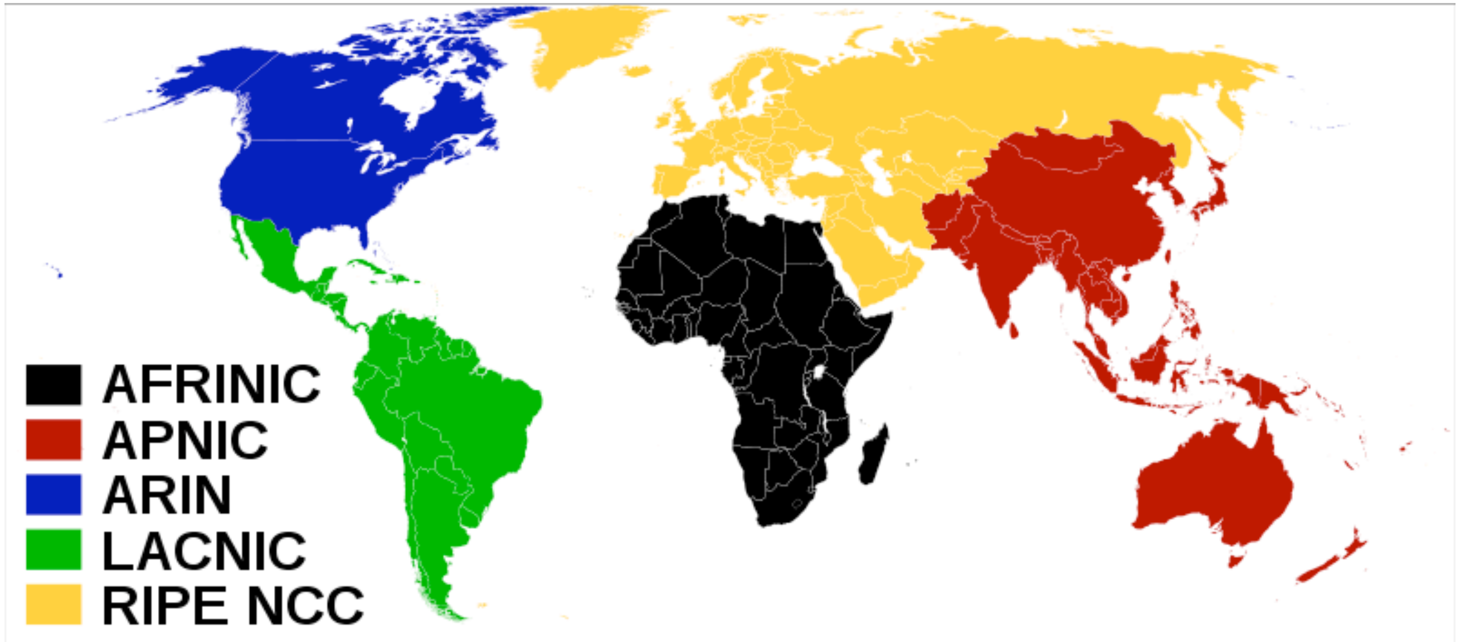
Q: How does an ISP get block of addresses?

A: ICANN: Internet Corporation for Assigned Names and Numbers

- allocates addresses to RIRs
- manages DNS
- assigns domain names, resolves disputes

How does one get an IP network address?

- IP address allocation is managed by five Regional Internet Registries (RIR)



- Each RIR manages ranges of addresses:

  http://www.iana.org/assignments/ipv4-address-space/ipv4-address-space.xml

Special IP Addresses

- **Special addresses:**

  Loopback interfaces
  - all addresses 127.0.0.1-127.0.0.255 are reserved for loopback interfaces
  - Most systems use 127.0.0.1 as loopback address
  - loopback interface is associated with name "localhost"

  IP address of a network
  - Host number is set to all zeros, e.g., 128.143.**0.0**

  Broadcast address
  - Host number is all ones, e.g., 128.143.**255.255**
  - Broadcast goes to all hosts on the network
  - Often ignored due to security concerns

- **Private Address Space**

  Certain address ranges are reserved for "private Internets" and "experimental use". Packets with with private source or destination addresses received on inter-AS links should be dropped (see RFC 1918):

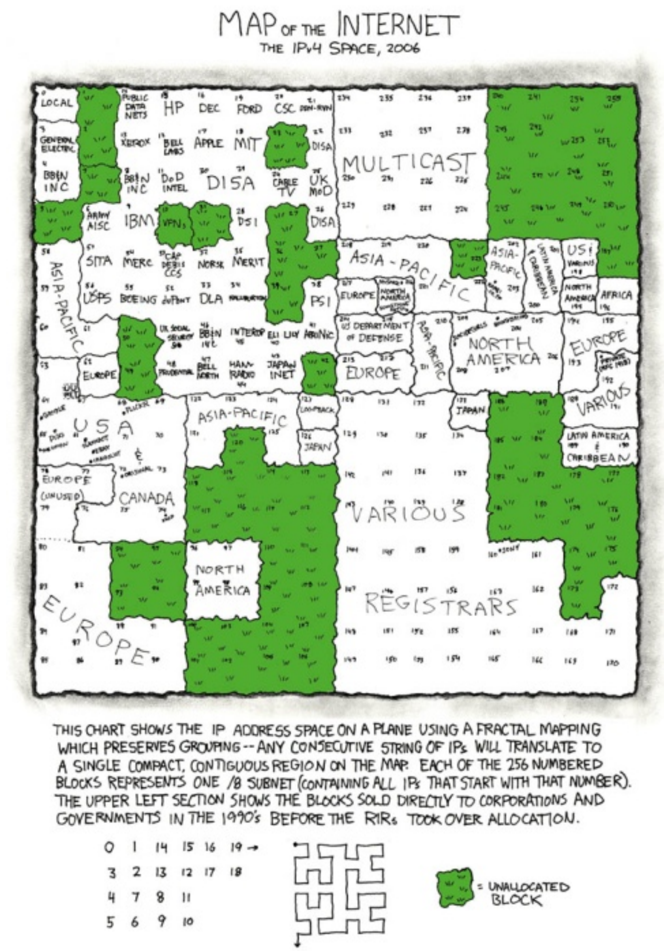
  10.0.0.0 - 10.255.255.255
  172.16.0.0 - 172.31.255.255
  192.168.0.0 - 192.168.255.255

- **Convention (but not a reserved address)**

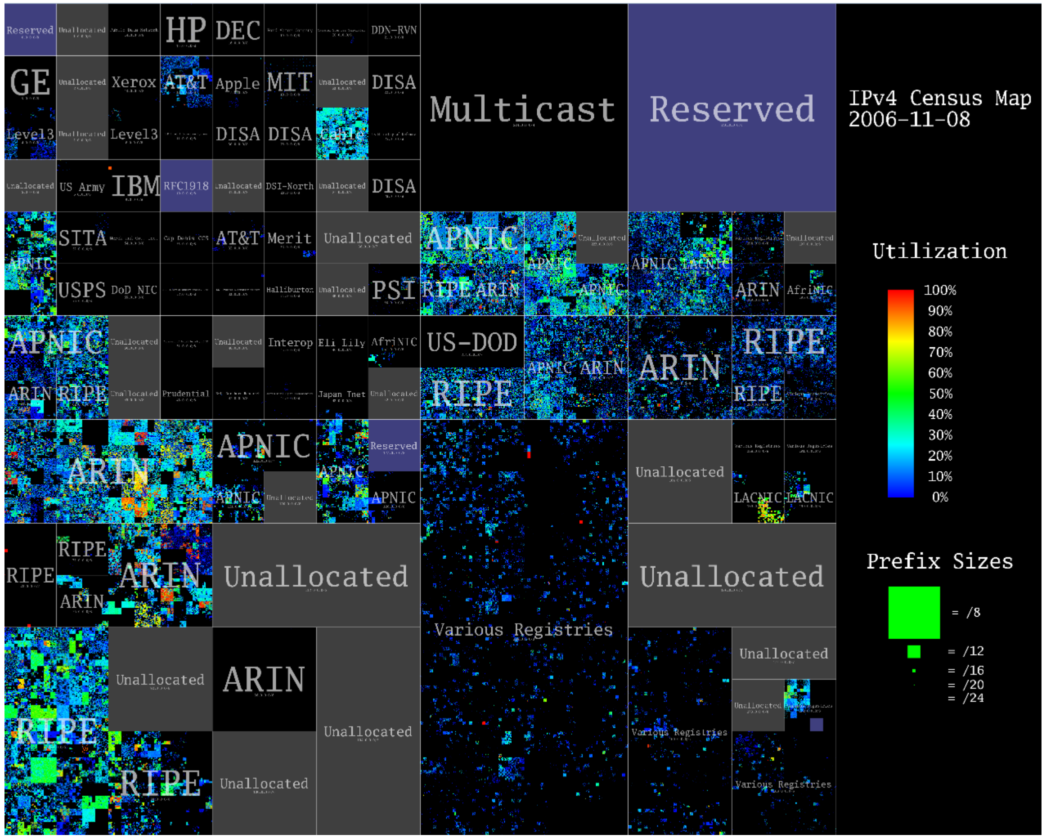
  Default gateway has host number set to '1', e.g., e.g., 192.0.1.**1**

Room-Filling Curves

- **Hilbert curve** for 2D representation of IPv4 address space



xkcd Comic



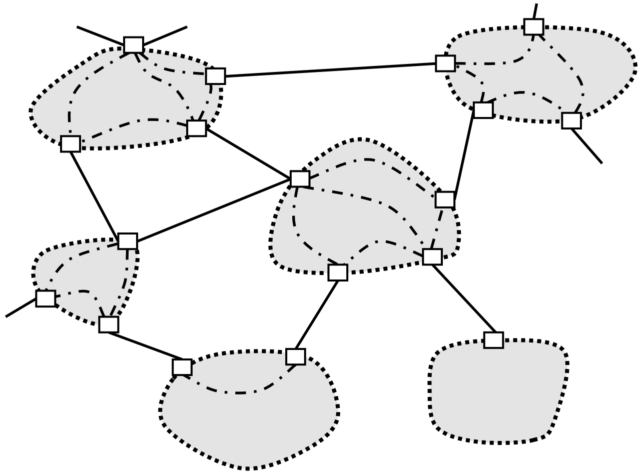
CAIDA IPv4 Census

How big is the Internet?

- Many measures:
  - networks (routed entities)
  - domains, host names (but: several names per host!)
  - directly (continuously) attached hosts ("ping' able")
  - IP-connected hosts (including dialin, e.g. PPP)
  - firewalled hosts
  - e-mail reachable

- What is the German Internet?
  - Entities within Germany
  - Entities operated by Germans / German organisations
  - Entities used by Germans / German organisations

Counting

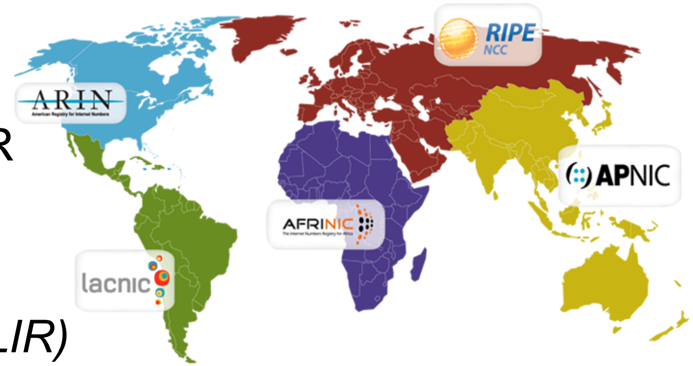


- Worldwide
  - > 700.000.000 hosts
  - > 37.000 Autonomous Systems
  - > 3.000.000.000 assigned IP addresses
  - > 2.180.000.000 reachable IP addresses
- Europe
  - > 126.600.000 hosts
  - > 19.000 Autonomous Systems
  - > 420.000.000 reachable IP addresses
  - > 500.000.000 assigned IP addresses
- Germany
  - > 13.300.000 hosts
  - > 1.200 Autonomous Systems
  - > 70.700.000 assigned IP addresses (5.500 prefixes)
  - > 62.700.000 reachable IP addresses

Snapshot 2011

Associations for Internet Names and Numbers

- **ICANN**
  - *„Internet Corporation for Assigned Names and Numbers“*
  - Private endowment (non-profit)
  - Administration of DNS Top Level Domains
  - Close collaboration with IETF and other related Intenet bodies (e.g., ISOC)
- **IANA**
  - *„Internet Assigned Numbers Authority“* (non-profit)
  - Operations via ICANN
  - Assignment of Internet numbers and Internet names
  - Administration of DNS root name servers
  - Administration of reverse DNS infrastructure (.arpa)
  - Assignment of protocol names and protocol numbers
- **NRO**
  - *„Number Resource Organization“*(non-profit)
  - Association of the 5 *„Regional Internet Registrars“ (RIR)*
  - Represents interest of RIRs towards ICANN/IANA
- **Regional Registrars** (non-profit)
  - ARIN, RIPE, APNIC, LACNIC, AfriNIC
  - *RIPE: „Réseaux IP Européens“* European RIR
  - Assigns IP addresses and AS numbers
  - Delegation of reverse DNS
  - Operation of Registrar database
  - Administration of *„Local Internet Registries“ (LIR)*



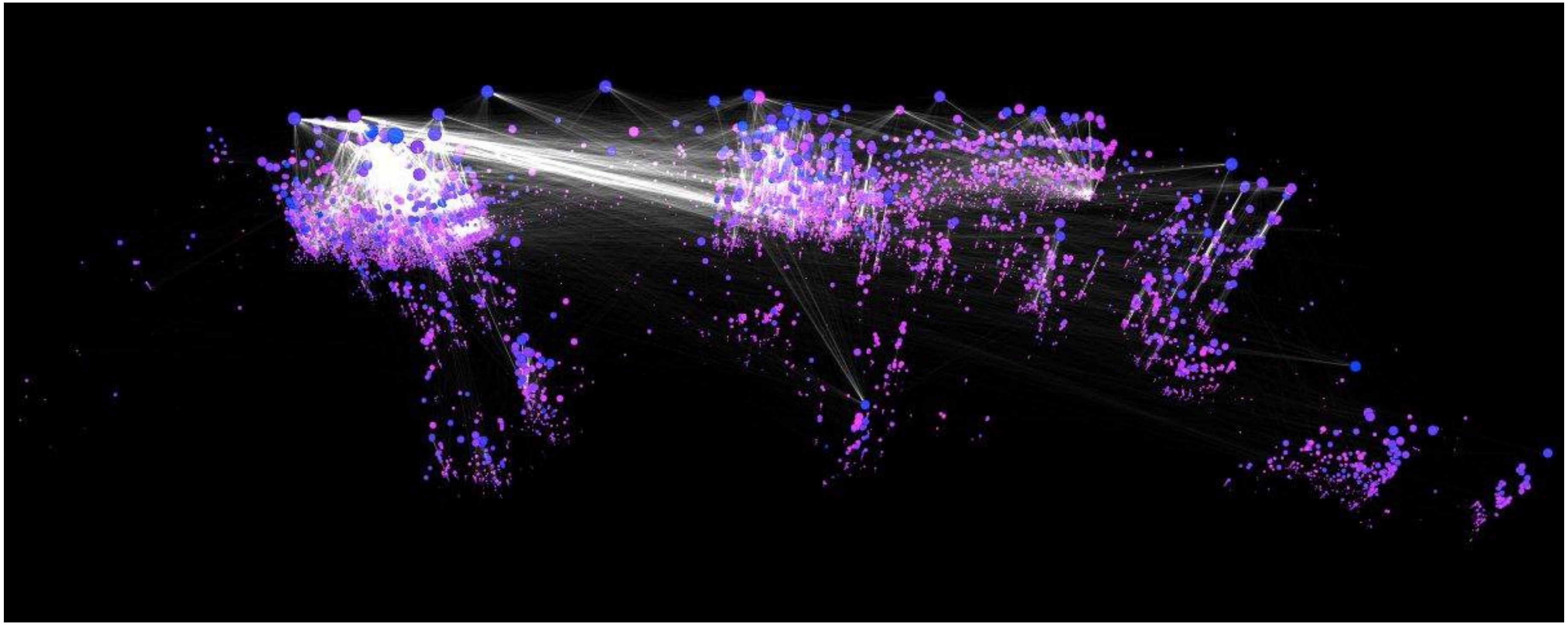
RIPE Database

- RIPE: registration and administration of Internet resources
  - AS information
  - Prefix information
  - Routing information
  - administrative work
- Online Whois service, and offline data bases
  - provides non-personal meta data
  - example:

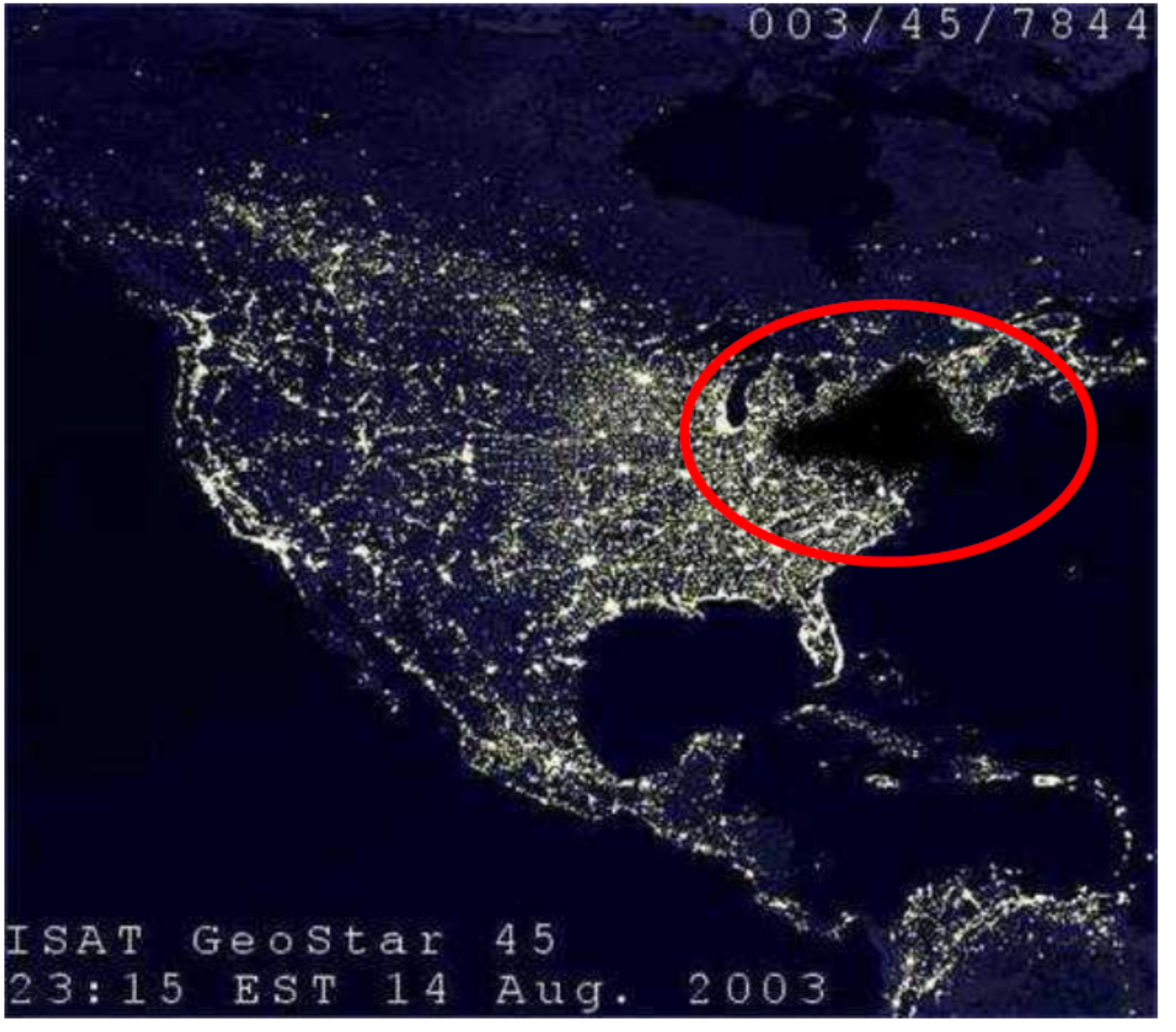
```
% Information related to 'AS56357'

aut-num:    AS56357
as-name:    TUM-I8-AS
descr:      Technische Universitaet Muenchen
descr:      Chair for Network Architectures and Services
import:     from AS680 accept ANY
import:     from AS33926 accept ANY
import:     from AS48918 accept ANY
export:     to AS680 announce AS56357
export:     to AS33926 announce AS56357
export:     to AS48918 announce AS56357
```

Internet Structure



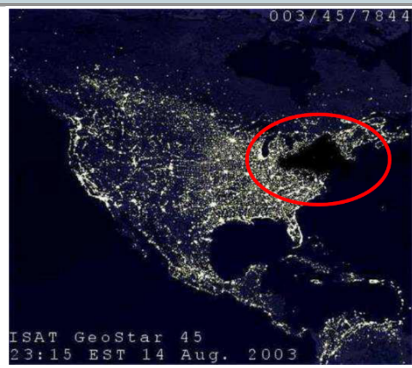
Incident



Blackout in North East USA, August 2003

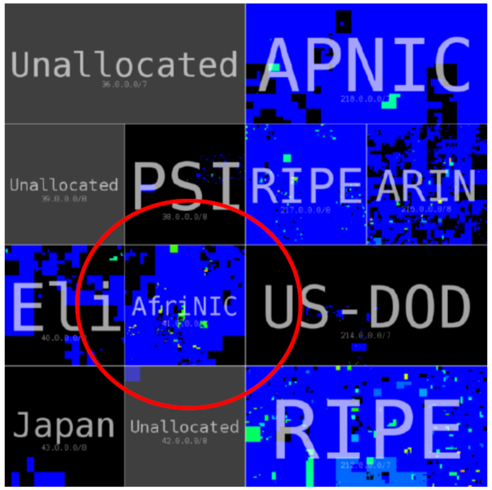
Analysis of BGP Anomalies

- **Goal:** find correlation between real-world incidents and BGP anomalies
  - Process large streams of BGP updates
  - Visualize changes over time
  - Analyze cause and effect

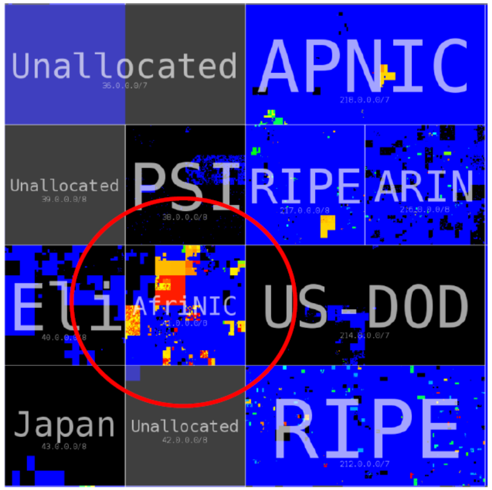


Power blackout in USA, Aug. 2003

- **Example:** Egyptian Internet shutdown, January 2011



Regular BGP activity

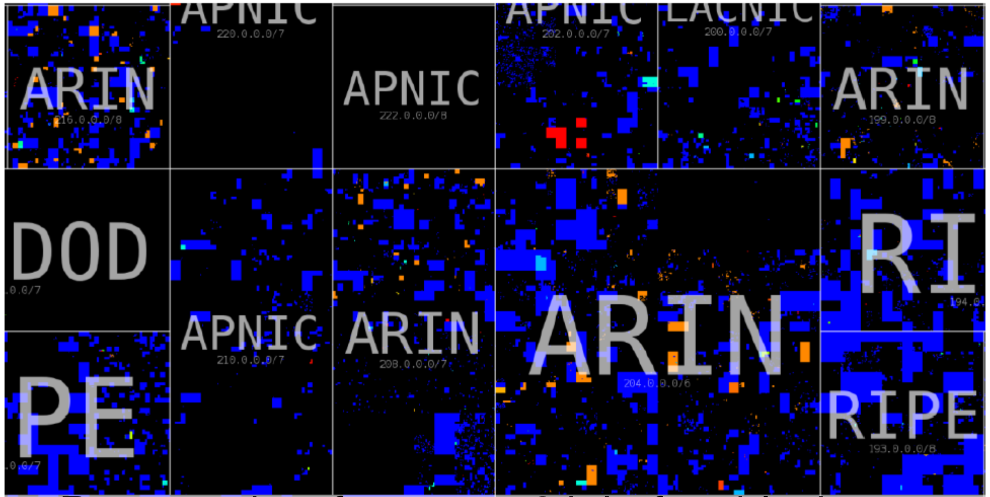


BGP activity at shutdown time

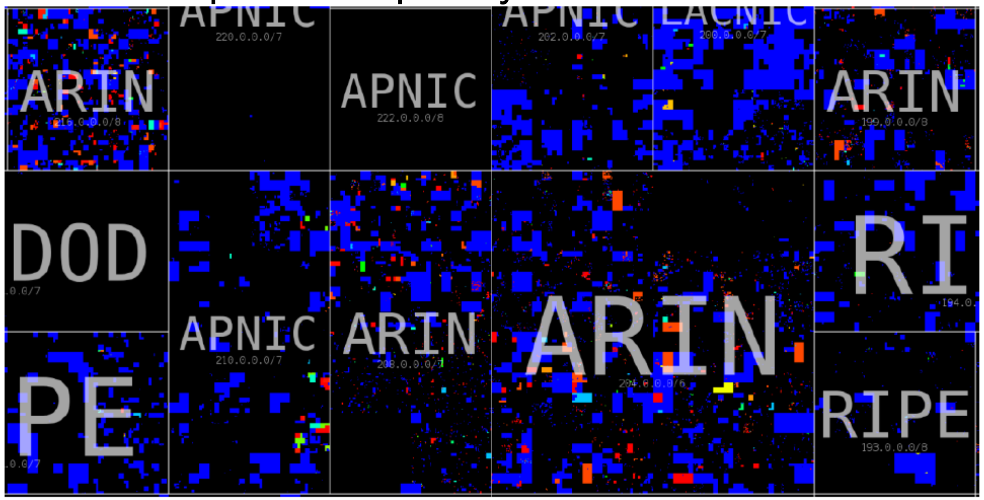
c.f. Sebastian Rampfl. **Analysis and visualization of BGP routing dynamics**. *Bachelor thesis supervised by Johann Schlamp*, TUM, March 2011.

Blackout USA, 2003

- > 100 power stations affected
- > 3.000 networks in > 1.700 organisations affected



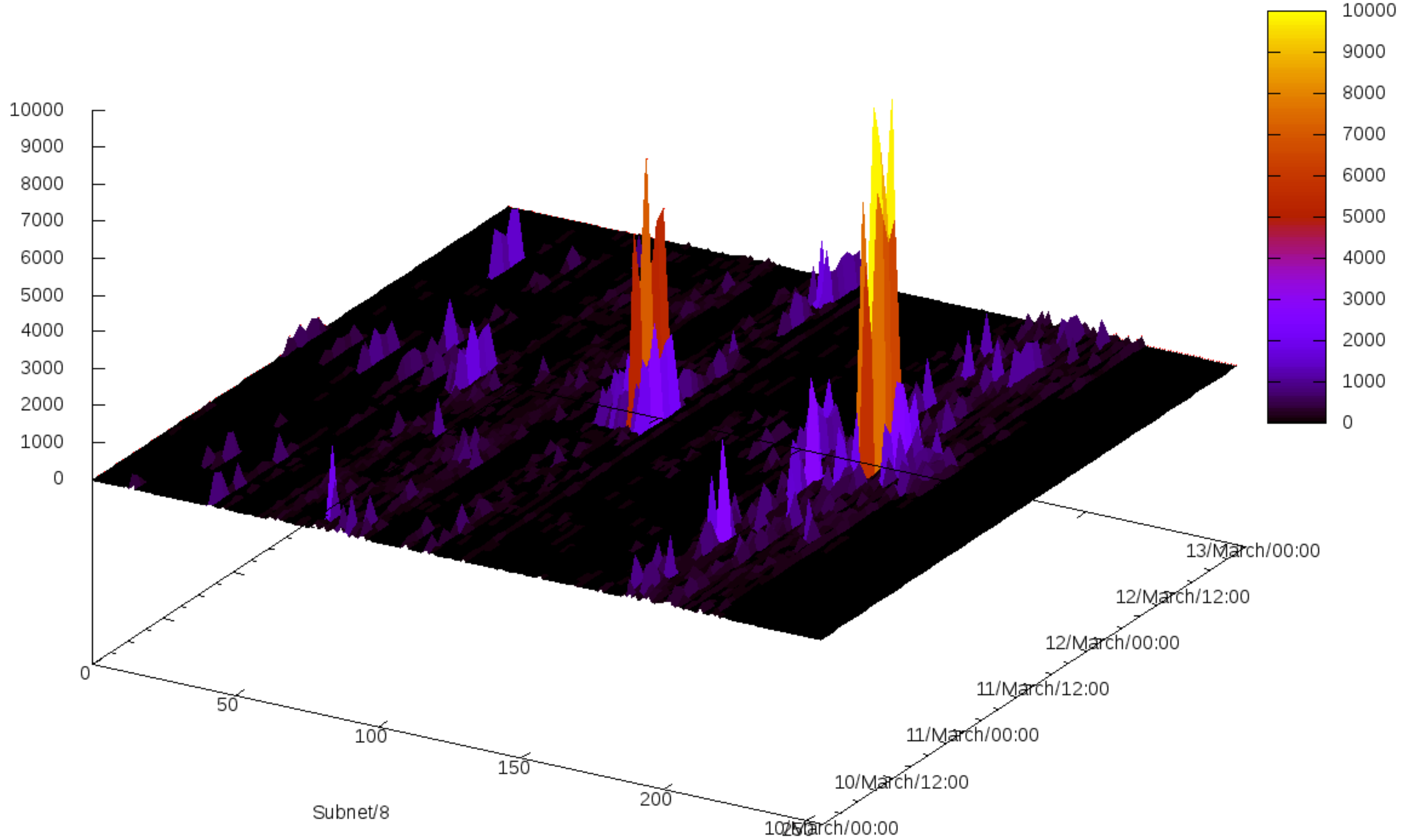
Route update frequency 2 h before blackout



Route update frequency 2 h after blackout

Internet Analysis – Path Graphs

- Example: Earth quake in Japan, March 2011

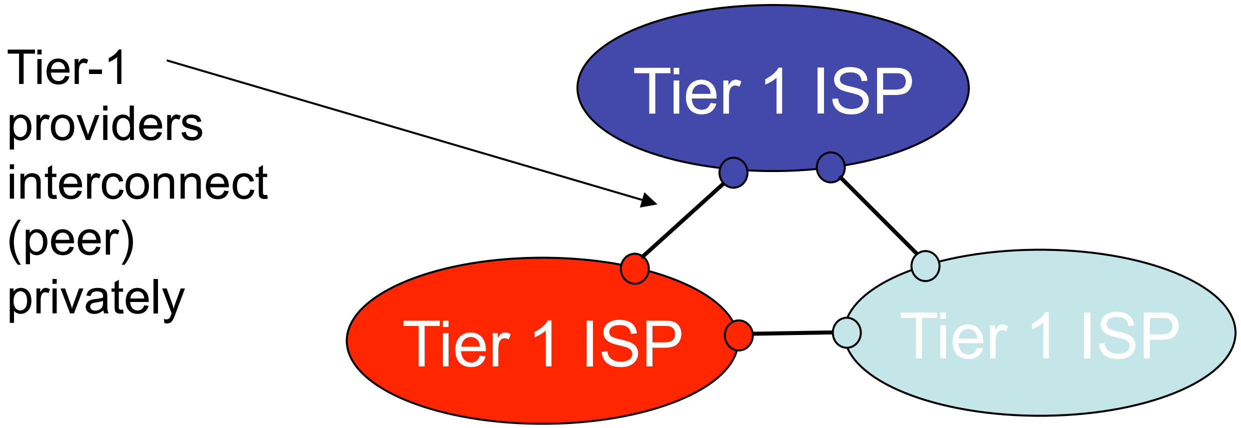


Internet Ecosystem

- >30,000 autonomous networks
- Networks with different
  - different roles and business type
    - stub networks
    - transit networks
    - content providers
  - Influenced by traffic patterns, application popularity, economics, regulation, ….
- Peering
  - bilateral contracts
  - Customer-provider, settlement-free peering, or in between
- Internet Exchange Points

Internet structure: network of networks

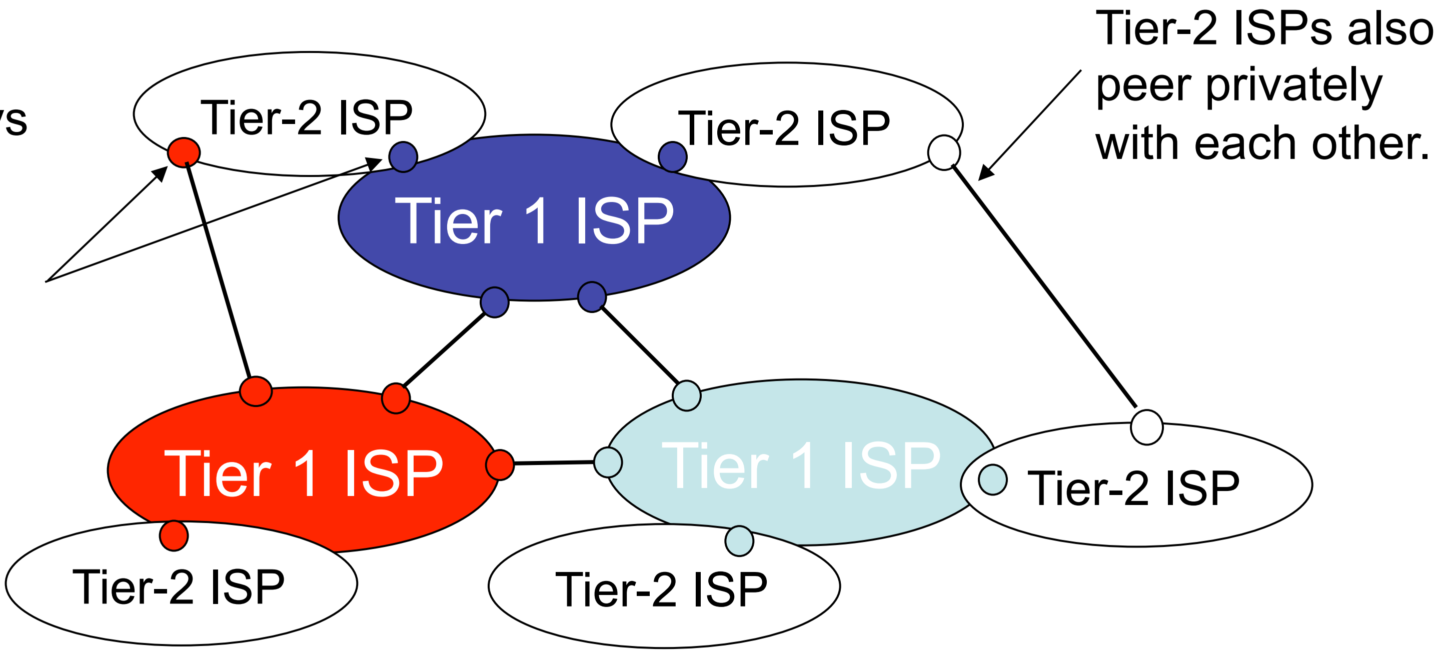
- roughly hierarchical
- at center: "tier-1" ISPs (AT&T, Global Crossing, Level 3, NTT, Qwest, Sprint, Tata, Verizon (UUNET), Savvis, TeliaSonera), national/international coverage
  - treat each other as equals
  - can reach every other network on the Internet without purchasing IP transit or paying settlements



Internet structure: network of networks

- "Tier-2" ISPs: smaller (often regional) ISPs
  - Connect to one or more tier-1 ISPs, possibly other tier-2 ISPs

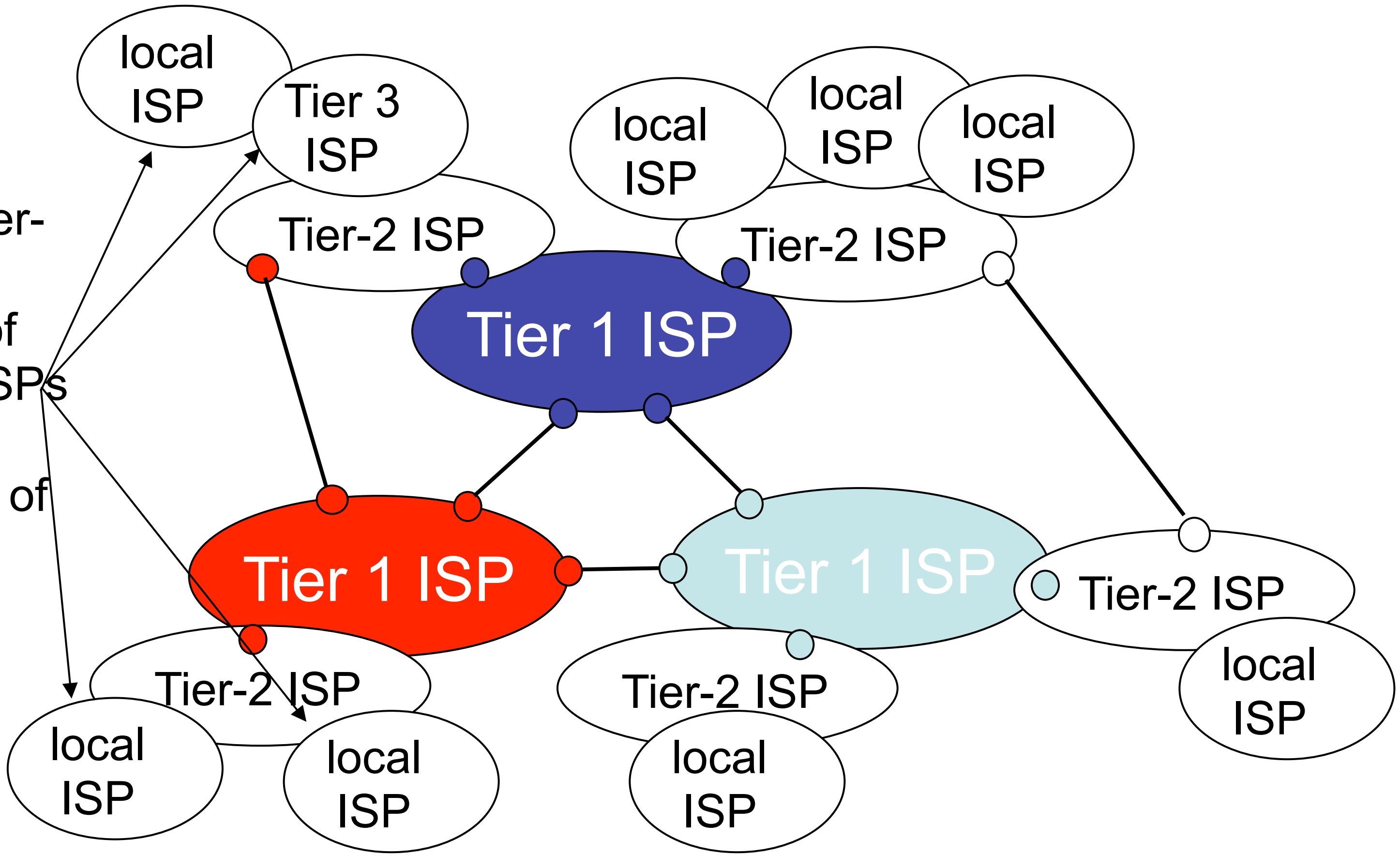
- Tier-2 ISP pays tier-1 ISP for connectivity to rest of Internet
- tier-2 ISP is *customer* of tier-1 provider



Internet structure: network of networks

- “Tier-3” ISPs and local ISPs
  - last hop (“access”) network (closest to end systems)

Local and tier-3 ISPs are *customers* of higher tier ISPs connecting them to rest of Internet



Internet structure: network of networks

- a packet passes through many networks!