An Introduction to Routing the Internet

Geoff Huston

APNIC


松江路
SongJiang Rd.
N 9th Ave. S
Sung Chiang Rd.
9th Ave. S
Singsheng N. Rd.
Sec. 3
S 8th Ave. N
民族東路
Min Tsu E. Rd.
濱江街
Pin Chiang St.
濱江街 1
Binjiang St.
基隆 23
Keelung
新生北路三段
XinSheng N. Rd.
Sec. 3
N 8th Ave. S
民權西路
民權東路一段

```
 2  85.132.40.1 (85.132.40.1)  1.202 ms  1.272 ms  1.425 ms
 3  94.20.50.141 (94.20.50.141)  2.394 ms  2.649 ms  2.044 ms
 4  r02-greenxchange-r04.az-ix.net (85.132.60.61)  2.634 ms  2.700 ms  2.292 ms
 5  so-7-0-0-ycr1.skt.cw.net (166.63.220.21)  71.707 ms  77.259 ms  78.499 ms
 6  xe-0-3-1-xcr2.amd.cw.net (195.2.9.217)  102.346 ms
    xe-0-2-1-xcr2.amd.cw.net (195.2.9.233)  103.459 ms
    xe-0-3-1-xcr2.amd.cw.net (195.2.9.217)  100.372 ms
 7  ae3-xcr1.amd.cw.net (195.2.30.105)  98.816 ms  99.660 ms  97.998 ms
 8  xe-4-1-0-xcr1.lsw.cw.net (195.2.25.93)  97.892 ms
    xe-7-2-0-xcr2.lsw.cw.net (195.2.25.138)  97.840 ms
    xe-4-1-0-xcr1.lsw.cw.net (195.2.25.93)  132.251 ms
 9  xe-3-0-1-xcr1.lns.cw.net (195.2.25.250)  101.963 ms  97.870 ms  97.707 ms
10  195.66.236.166 (195.66.236.166)  102.975 ms  101.481 ms  107.319 ms
11  i-2-0-0.ulco-core01.bi.telstraglobal.net (202.40.148.218)  104.572 ms  98.277 ms  102.415 ms
12  i-6-0-0.eig-core01.bx.telstraglobal.net (202.84.249.21)  241.369 ms  238.537 ms  240.648 ms
13  i-0-4-1-0.eqnx-core01.bi.telstraglobal.net (202.84.249.34)  239.569 ms  239.580 ms  239.513 ms
14  i-0-4-0-0.sydo-core02.bx.telstraglobal.net (202.84.144.186)  386.366 ms  386.084 ms  398.815 ms
15  tengige0-2-0-0.oxf-gw1.sydney.telstra.net (203.50.13.29)  388.800 ms  399.029 ms  397.750 ms
16  bundle-ether1.ken-core4.sydney.telstra.net (203.50.6.5)  394.591 ms  406.403 ms  396.072 ms
17  bundle-ether5.cha-core4.brisbane.telstra.net (203.50.11.73)  410.061 ms  411.584 ms  413.307 ms
18  tengigabitethernet7-1.cha30.brisbane.telstra.net (203.50.51.40)  432.973 ms  401.504 ms  400.854 ms
19  4608resolvers.brisbane.telstra.net (139.130.130.194)  405.478 ms  407.010 ms  408.246 ms
20  wattle.rand.apnic.net (203.133.248.2)  406.757 ms  410.261 ms  410.676 ms
```

The Objective of Routing

- To ensure that every packet gets delivered
  - To ensure that every packet gets delivered in the shortest possible time
- This implies that every router needs to understand the relative direction of all possible destinations
- So every router needs to know the current relative location of all possible destinations
- This is the task of the routing system

Routing Architecture

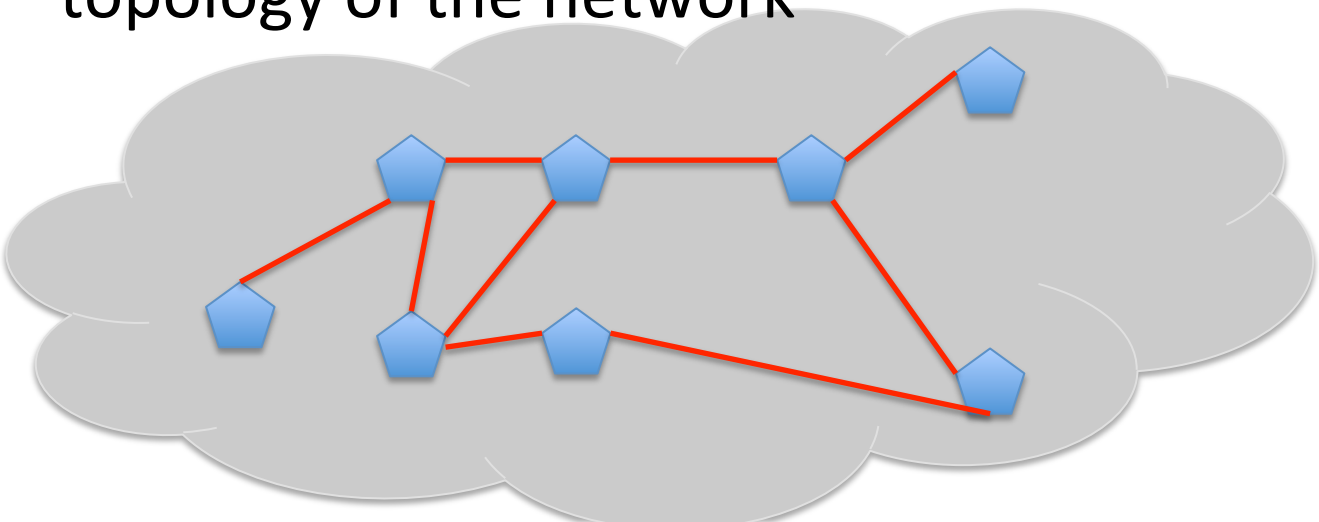
- The Internet uses a *decoupled* routing architecture
  - A routing protocol is designed to pass forwarding information between routers. It does not perform the forwarding function itself
  - Different networks may (and do) use different routing protocols
  - Routing protocols change over time
    - But need not change all at once

Routing Protocols

- Routing protocols are “self discovery” distributed protocols
  - Two flavours of algorithm:
    - Distance vector (rumours)
      - I tell you everything I know
      - You tell me everything you know
      - Repeat until nothing more to day!
    - Link State (map drawing)
      - I tell everyone about my links (link state messages)
      - I listen to all link state messages
      - I compute a optimal path map of the network

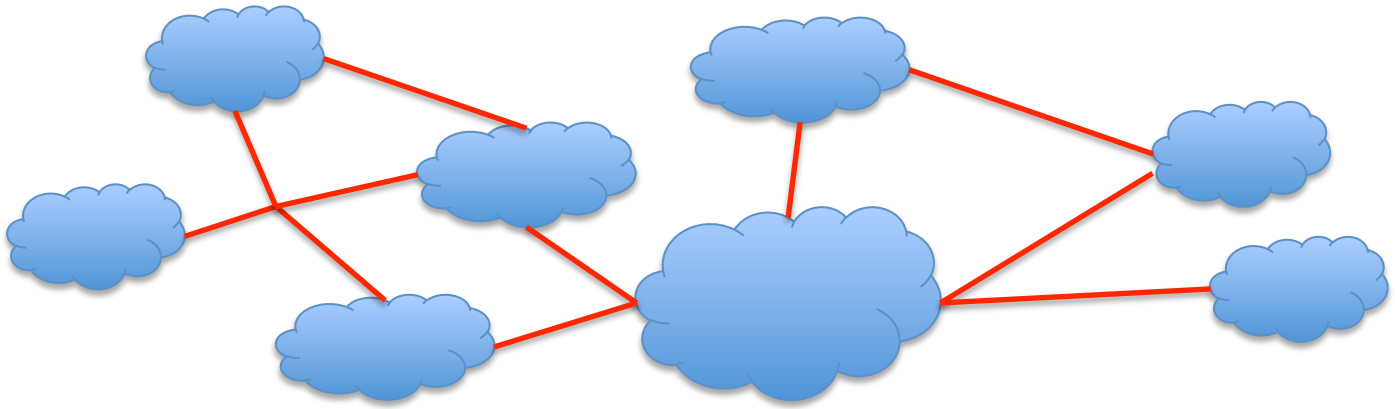
Routing Architecture

- The Internet uses a *two level* routing hierarchy:
  - Interior Routing Protocols, used by each network to determine how to reach all destinations that line within the network
  - Interior Routing protocols maintain the current topology of the network



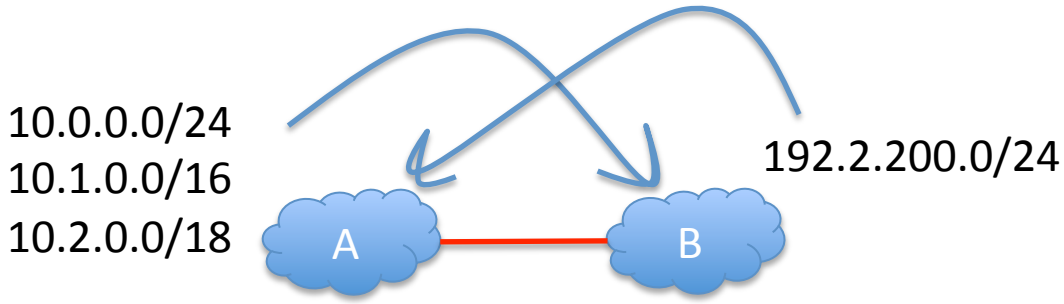
Routing Architecture

- The Internet uses a *two level* routing hierarchy:
  - Exterior Routing Protocol, used to link each component network together into a single whole
  - Exterior protocols assume that each network is fully interconnected internally



Exterior Routing: BGP

- BGP is a large set of bilateral (1:1) routing sessions
  - A tells B all the destinations (prefixes) that A is capable of reaching
  - B tells A all the destinations that B is capable of reaching



BGP decisions

- BGP is an “offer and accept” protocol
  - A network “offers” to a neighboring network prefixes that it is prepared to act as a forwarder
  - The network need not offer all its prefixes to a routing neighbour
    - It may have internal routing policies to moderate what it is prepared to offer
  - The neighboring network is not obliged to accept the offer
    - It may have better offers from other routing peers that show a preferred path
    - It may have policies about redistribution of prefixes that determine what is offered and what is accepted

A view of a Routing Table

```
show ip bgp
BGP table version is 0, local router ID is 203.133.248.2
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
              r RIB-failure, S Stale, R Removed
Origin codes: i - IGP, e - EGP, ? - incomplete

   Network          Next Hop            Metric LocPrf Weight Path
*  1.0.0.0/24       202.12.28.1                            0 4777 15169 i
*>                  203.119.76.3                           0 4608 15169 i
*  1.0.4.0/22       202.12.28.1                            0 4777 2516 6453 7545 7545 7545 56203 i
*>                  203.119.76.3                           0 4608 1221 2764 24130 7545 7545 56203 i
*> 1.0.16.0/23      202.12.28.1                            0 4777 2519 i
*                   203.119.76.3                           0 4608 1221 4637 3356 2516 2519 i
*> 1.0.18.0/23      202.12.28.1                            0 4777 2519 i
*                   203.119.76.3                           0 4608 1221 4637 3356 2516 2519 i
*> 1.0.20.0/23      202.12.28.1                            0 4777 2519 i
*                   203.119.76.3                           0 4608 1221 4637 3356 2516 2519 i
*> 1.0.22.0/23      202.12.28.1                            0 4777 2519 i
*                   203.119.76.3                           0 4608 1221 4637 3356 2516 2519 i
*> 1.0.24.0/23      202.12.28.1                            0 4777 2519 i
*                   203.119.76.3                           0 4608 1221 4637 3356 2516 2519 i
*> 1.0.26.0/23      202.12.28.1                            0 4777 2519 i
```

Routing Properties

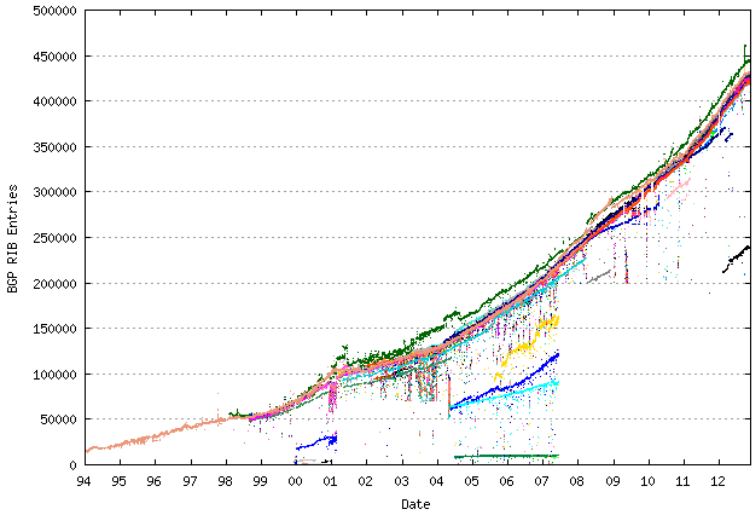
- Routing is designed to be robust
  - Within a richly connected topology the interior routing protocol is designed to “self heal” in the even of breakage
  - Within the exterior network, connectivity will also “self heal” if there are alternative paths that route around the break
- Routing is efficient
  - Interior routing can be tuned to converge in milliseconds
  - Exterior routing will converge across the span of the Internet within 70 seconds on average following a change in local connectivity

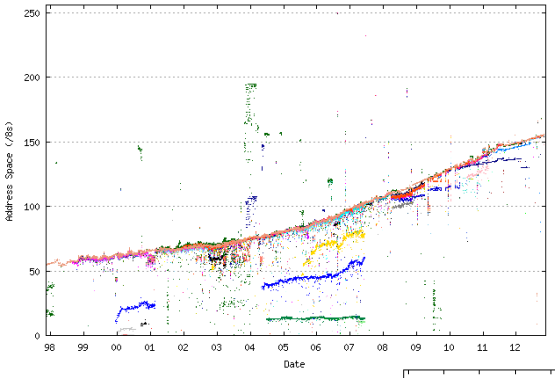
Routing Policies

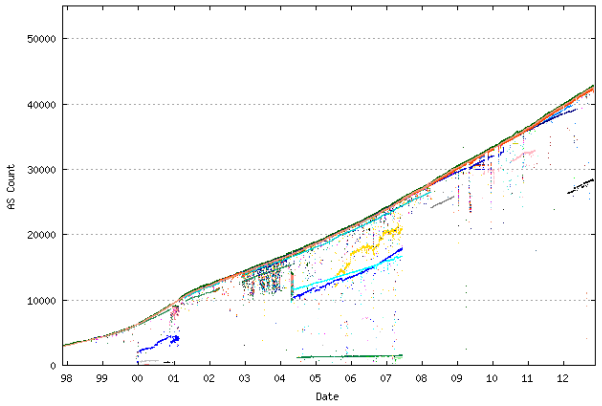
- Exterior routing systems are often used to express a number of objectives:
  - Business requirements (e.g.: customer, peer, upstream)
  - Traffic Engineering requirements (e.g.: traffic flow balancing)
  - Robustness requirements (e.g.: alternate connection arrangements and automated failover)

Routing Insights

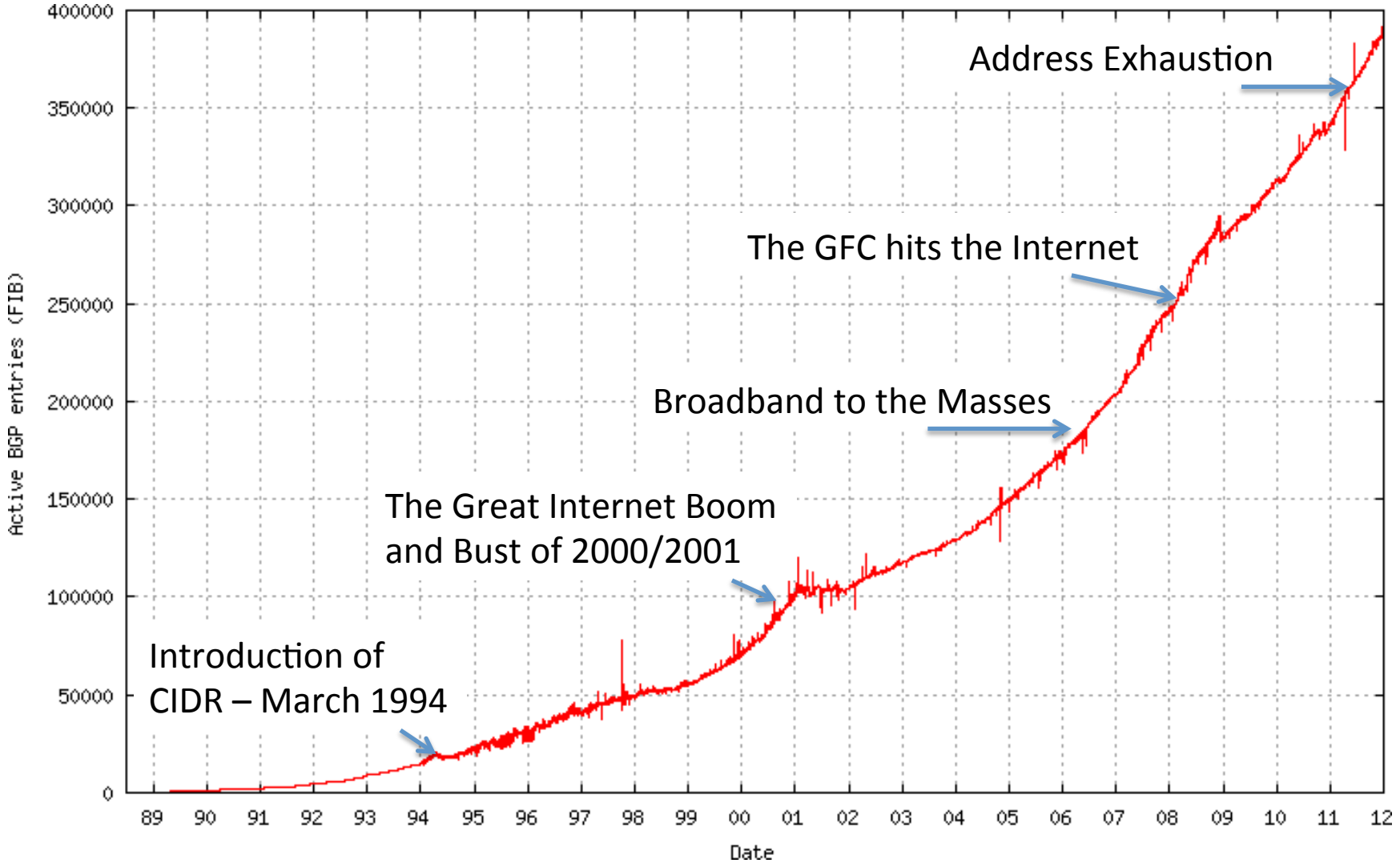
- Routing is essentially a view of “all of the network” at once, and as such it provides a number of useful insights into the connected network

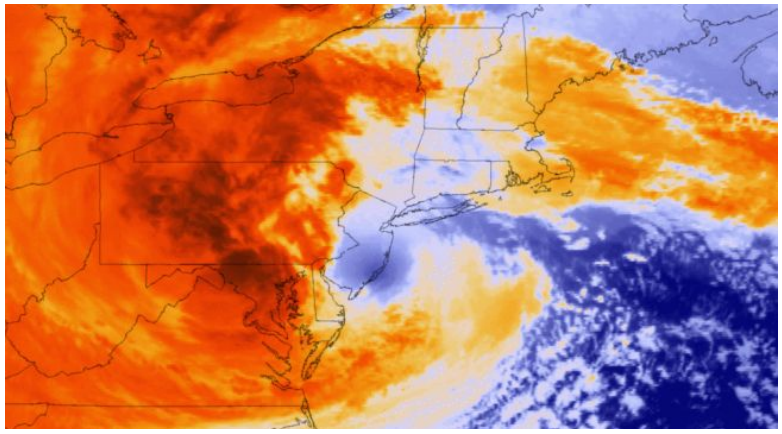




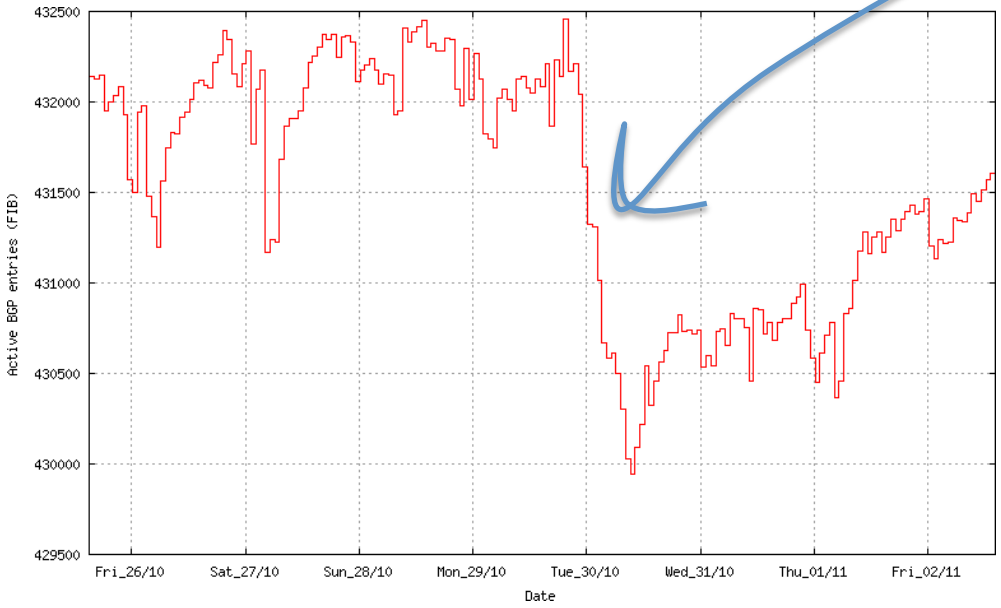


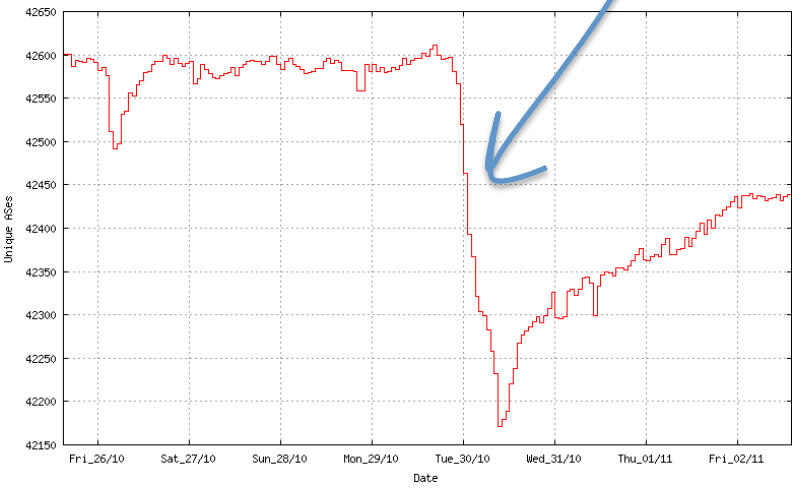
The Big Picture of the v4 Routing Table



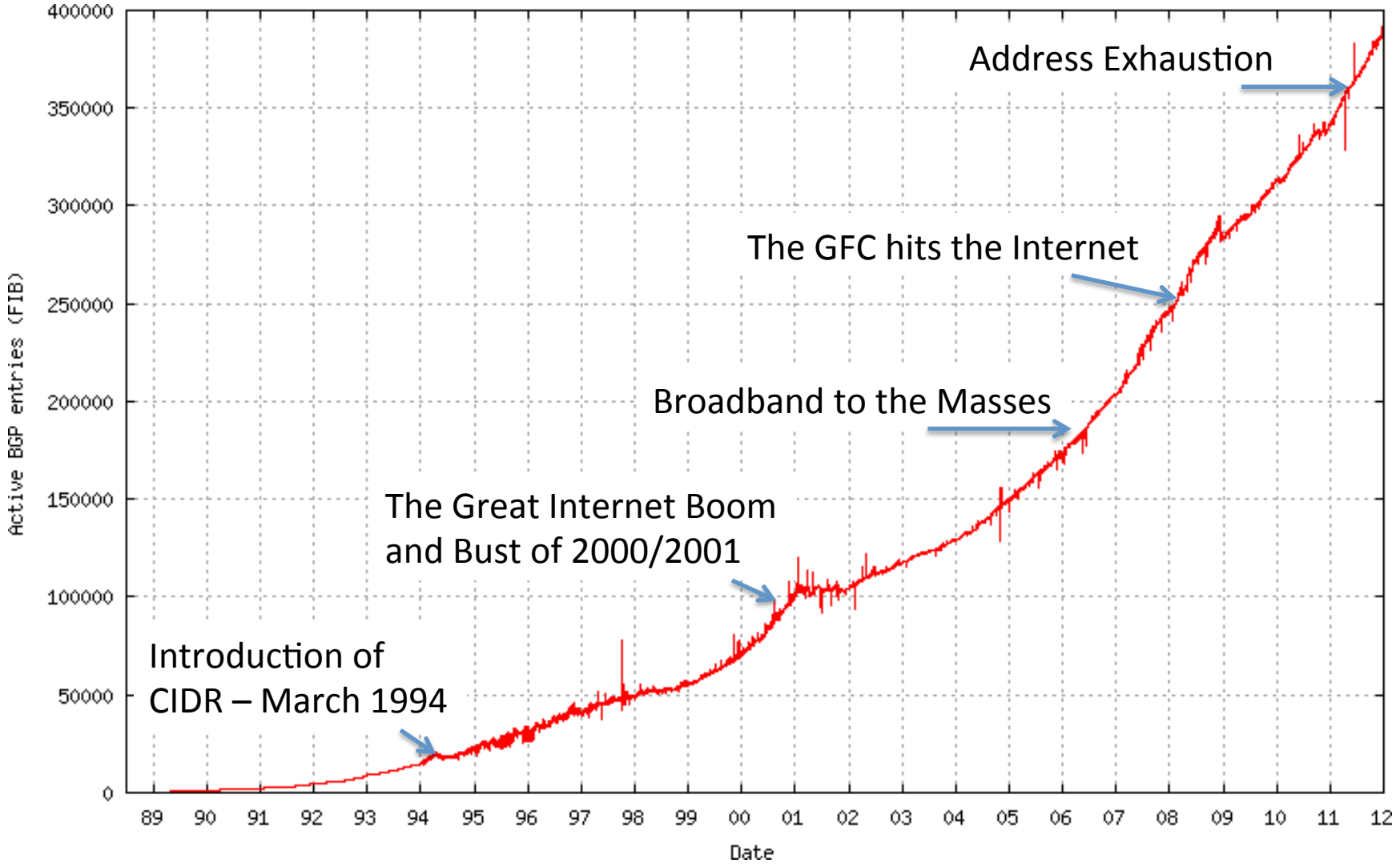


Routing Issues





The Big Picture of the v4 Routing Table

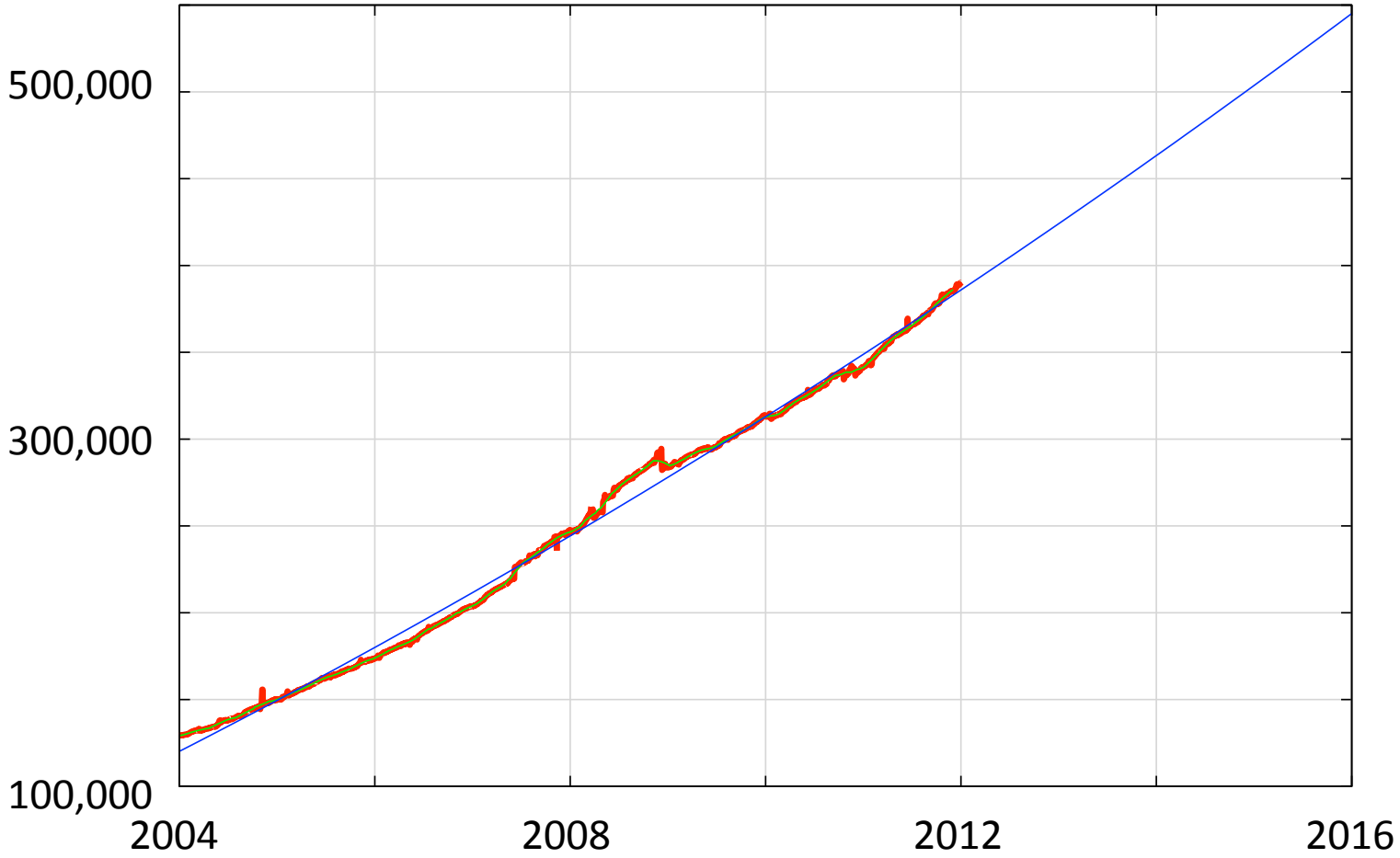


Hot topics in Routing

Security

Hot Topics in Routing

- Scaling



Open Routing Standards

- Interior Routing Protocols
  - Open Shortest Path First (OSPF)
  - Intermediate System – Intermediate System (IS-IS)
- Exterior Routing Protocols
  - Border Gateway Protocol (BGP)