CS615 - Aspects of System Administration

Networking I

Department of Computer Science
Stevens Institute of Technology
Jan Schaumann
jschauma@stevens.edu
http://www.cs.stevens.edu/~jschauma/615A/

IPv4 Basics

10011011111101100101100101100100

IPv4 addresses are 32-bit numbers.

IPv4 Basics

```
10011011   11110110   01011001   01100100
```

Each IPv4 address consists of four octets.

IPv4 Basics

```
10011011  11110110  01011001  01100100

 155   .   246   .    89   .   100
```

Each IPv4 address consists of four octets.

IPv4 Basics

```
10011011 11110110    01011001 01100100
```

IPv4 addresses are divided into a *network part* and a *host part*.

Hosts on the same network (*broadcast domain*) can talk to each other without the help of a router.

IPv4 Basics

```
10011011 11110110    01011001 01100100
```

There are three different *classes* of IPv4 networks.

IPv4 Basics

```
10011011 11110110    01011001 01100100
```

There are three different *classes* of IPv4 networks.
Well, five, really.

IPv4 Basics

Class	Leading bits	Size of *network number* bit field	Size of *rest* bit field	Number of networks	Addresses per network	Start address	End address
Class A	0	8	24	128 (2^7)	16,777,216 (2^{24})	0.0.0.0	127.255.255.255
Class B	10	16	16	16,384 (2^{14})	65,536 (2^{16})	128.0.0.0	191.255.255.255
Class C	110	24	8	2,097,152 (2^{21})	256 (2^8)	192.0.0.0	223.255.255.255
Class D (multicast)	1110	not defined	not defined	not defined	not defined	224.0.0.0	239.255.255.255
Class E (reserved)	1111	not defined	not defined	not defined	not defined	240.0.0.0	255.255.255.255

Subnets

```
10011011   11110110   01011001   01100100

11111111   11111111   00000000   00000000
```

A *netmask* splits the IPv4 address into *network* and *host* parts.

Subnets

```
10011011   11110110   01011001   01100100

11111111   11111111   11111111   00000000
```

A *netmask* splits the IPv4 address into *network* and *host* parts.

Subnets

```
$ ipcalc -n 155.246.89.100/16
Address:   155.246.89.100       10011011.11110110. 01011001.01100100
Netmask:   255.255.0.0 = 16     11111111.11111111. 00000000.00000000
Wildcard:  0.0.255.255          00000000.00000000. 11111111.11111111
=>
Network:   155.246.0.0/16       10011011.11110110. 00000000.00000000
HostMin:   155.246.0.1          10011011.11110110. 00000000.00000001
HostMax:   155.246.255.254      10011011.11110110. 11111111.11111110
Broadcast: 155.246.255.255      10011011.11110110. 11111111.11111111
Hosts/Net: 65534                 Class B
```

Try also: `sipcalc -a 155.246.89.100/16`

Subnets

```
$ ipcalc -n 155.246.89.100/24
Address:   155.246.89.100       10011011.11110110.01011001. 01100100
Netmask:   255.255.255.0 = 24   11111111.11111111.11111111. 00000000
Wildcard:  0.0.0.255            00000000.00000000.00000000. 11111111
=>
Network:   155.246.89.0/24      10011011.11110110.01011001. 00000000
HostMin:   155.246.89.1         10011011.11110110.01011001. 00000001
HostMax:   155.246.89.254       10011011.11110110.01011001. 11111110
Broadcast: 155.246.89.255       10011011.11110110.01011001. 11111111
Hosts/Net: 254                   Class B
```

Try also: `sipcalc -a 155.246.89.100/24`

CIDR cheat sheet

A.B.C.D/N

- N = bits describing network portion of address
- $M = 32 - N$ = bits in host portion of address
- 2^M = number of addresses on this subnet
- $2^M - 2$ = number of possible hosts
  - first address on subnet = network address
  - last address on subnet = broadcast address
- subnet division need not occur on dotted boundary only
  - for example, you can divide 155.246.89.0/24 into four /26 networks
  - networks starting at .0, .64, .128, .192

Which of the following is not a valid netmask?

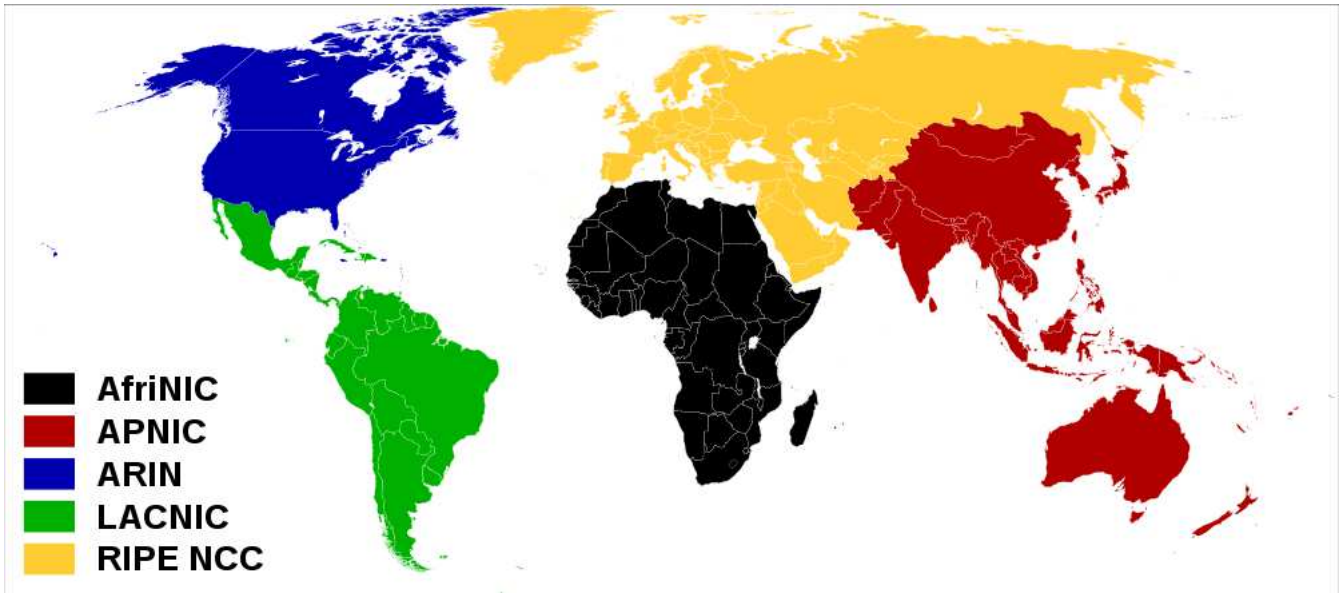
`255.255.253.0, 255.255.250.0, 255.255.240.0`

Mommy, where do IP addresses come from?

The Internet Assigned Numbers Authority (IANA) oversees global IP address/AS number allocation, root zone management etc.

`https://www.iana.org/`

Mommy, where do IP addresses come from?



Regional Internet Registries (RIR) manage the allocation and registration of Internet number resources within a region of the world.

Mommy, where do IP addresses come from?

RIRs assign blocks of IP addresses to the Local Internet Registries (LIR).

LIRs are either ISPs, enterprises using a lot of addresses, or academic institutions.

IPv4 Subnets: Common CIDRs

```
10011011  11110110  01011001  01100100
       |         |  ||||   |  |||||||| /32   Host route
       |         |  ||||   |  ||||||   /30   "Glue network" (Point-to-point)
       |         |  ||||   |  |||||    /29   Smallest multi-host network
       |         |  ||||   |  ||||     /28   Small LAN
       |         |  ||||   |  |||      /27   Small LAN
       |         |  ||||   |  ||       /26   Small LAN
       |         |  ||||   |  |        /25   Large LAN
       |         |  ||||   |           /24   Large LAN
       |         |  ||||               /20   Small ISP / Large business
       |         |  |||                /19   LIR / ISP / Large business
       |         |  ||                 /18   LIR / ISP / Large business
       |         |  |                  /17   LIR / ISP / Large business
       |         |                     /16   LIR / ISP / Large business
       |                               /8    RIR
```

IPv4 Exhaustion

IPv4 address space depletion:

- private IP space (RFC1918): `10.0.0.0/8`, `172.16.0.0/12`, `192.168.0.0/16`
- class D (`224.0.0.0/4`) and E (`240.0.0.0/4`)
- class As (16M addresses each!) initially handed out liberally (ATT, Apple, MIT, Stanford, Xerox, ...)
- subnetting often inefficient
- more and more devices added

IPv4 Exhaustion

IPv4 address space depletion:

Total theoretically available IP addresses: 2^{32}

RFC1918: `10.0.0.0/8`, `172.16.0.0/12`, `192.168.0.0/16`

RFC5735 etc.: `0.0.0.0/8`, `100.64.0.0/10`, `127.0.0.0/8`, `169.254.0.0/16`, `192.0.0.0/24`, `192.0.2.0/24`, `192.88.99.0/24`, `198.18.0.0/15`, `198.51.100.0/24`, `203.0.113.0/24`

Class D/E: `224.0.0.0/4`, `240.0.0.0/4`

”Limited broadcast”: `255.255.255.255/32`

What is the percent/number of actually available IP addresses?

IPv4 Exhaustion

Past and predicted:

IANA Address Pool Exhaustion:	2011-02-03
APNIC reached final /8:	2011-04-19
RIPENCC reached final /8:	2012-09-14
LACNIC reached final /8:	2014-06-10
ARIN reached final /8:	2015-09-15
AFRINIC(predicted):	2018-06-06

`http://www.potaroo.net/tools/ipv4/`
`http://www.iana.org/assignments/ipv4-address-space/`

IPv6 Basics

10011011111101100101100101100100

IPv4 addresses are 32-bit numbers.

IPv6 Basics

0010000000000001
0000010011111000
0000000000000100
0000000000000111
0000001011100000
1000000111111111
1111111001010010
1001101001101011

IPv6 addresses are 128 bits.

IPv6 Basics

IPv4: 32 bits $=>$ 2^{32} addresses

IPv6: 128 bits $=>$ 2^{128} addresses

IPv6 Basics

IPv4: 32 bits $=>$ $4,294,967,296$ addresses

IPv6: 128 bits $=>$ 2^{128} addresses

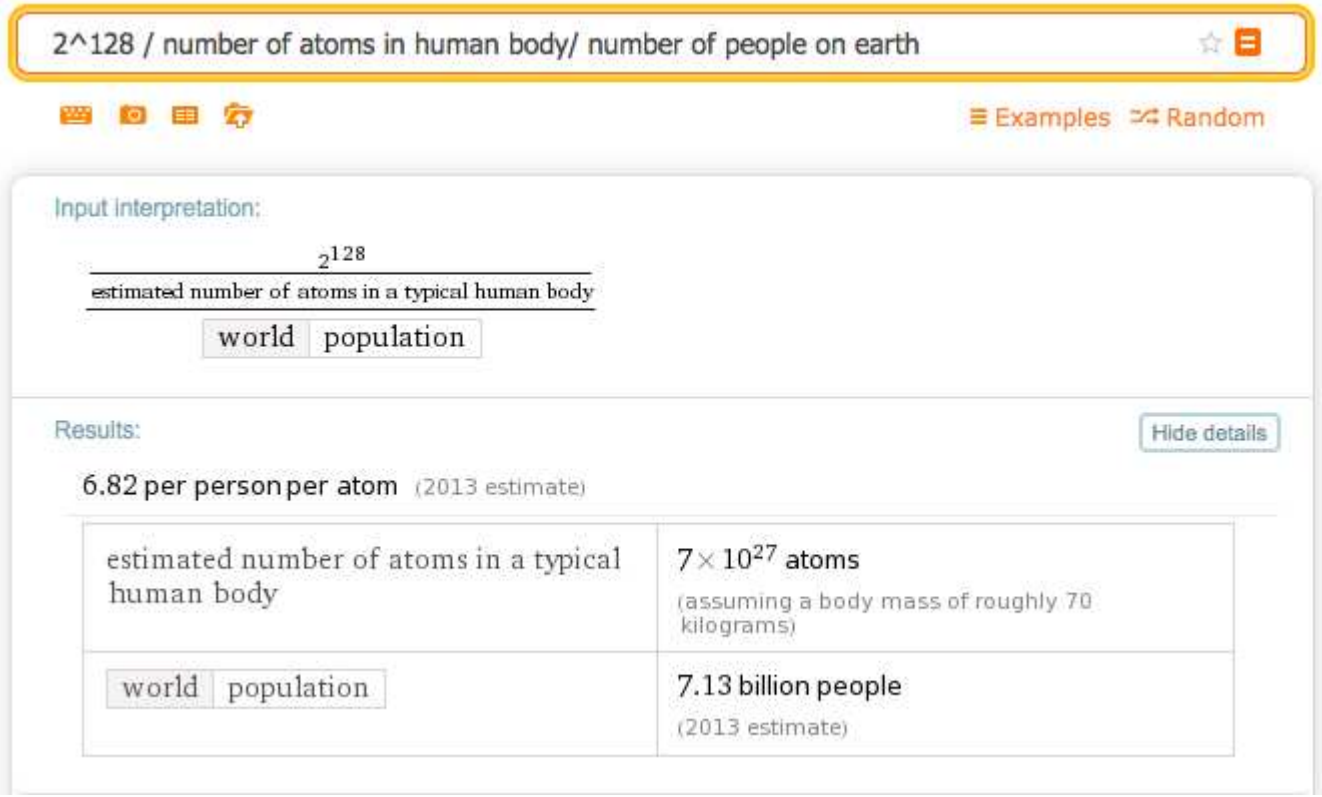
IPv6 Basics

IPv4: 32 bits $=>$ $4,294,967,296$ addresses

IPv6: 128 bits $=>$

$340,282,366,920,938,463,463,374,607,431,768,211,456$ addresses

IPv6 Basics



http://is.gd/94ve91

IPv6 Basics

- 8x16 bit fields (words) in case insensitive colon hexadecimal representation

```
2031:0000:0000:030F:0000:0000:0000:130B
```

IPv6 Basics

- 8x16 bit fields (words) in case insensitive colon hexadecimal representation

  `2031:0000:0000:030F:0000:0000:0000:130B`

- Leading zeros in a field are optional:

  `2031:0:0:30F:0:0:0:130B`

IPv6 Basics

- 8x16 bit fields (words) in case insensitive colon hexadecimal representation

```
2031:0000:0000:030F:0000:0000:0000:130B
```

- Leading zeros in a field are optional:

```
2031:0:0:30F:0:0:0:130B
```

- Successive fields of 0 represented as ::, but only once in an address:

```
2031::30F:0:0:0:130B          ok
2031:0:0:30F::130B            ok
2031::30F::130B               not ok
```

IPv6 Basics

- 8x16 bit fields (words) in case insensitive colon hexadecimal representation

```
2031:0000:0000:030F:0000:0000:0000:130B
```

- Leading zeros in a field are optional:

```
2031:0:0:30F:0:0:0:130B
```

- Successive fields of 0 represented as ::, but only once in an address:

```
2031::30F:0:0:0:130B          ok
2031:0:0:30F::130B            ok
2031::30F::130B               not ok
```

- ```
  0000:0000:0000:0000:0000:0000:0000:00001 =>
                              0:0:0:0:0:0:0:1 => ::1
  ```
  ```

IPv6 Basics - Address Oddities

- Address may include a link name:

```
2001:470:1f07:3d1::1%eth0
```

IPv6 Basics - Address Oddities

- Address may include a link name:

```
2001:470:1f07:3d1::1%eth0
```

- IPv4-mapped addresses

```
0:0:0:0:0:ffff:66.163.162.9
::ffff:66.163.162.9
```

IPv6 Basics - Address Oddities

- Address may include a link name:

```
2001:470:1f07:3d1::1%eth0
```

- IPv4-mapped addresses

```
0:0:0:0:0:ffff:66.163.162.9
::ffff:66.163.162.9
```

- You need brackets to distinguish a port from an address:
  - IPv4: `66.163.162.9:22`
  - IPv6: `[2001:470:1f07:3d1::1]:22`

IPv6 Configuration Types

- Static Configuration
- Stateful Autoconfiguration (DHCPv6)
- Stateless Address Autoconfiguration (SLAC)
  - RFC2462
  - use of autonomously configured link-local address using its EUI-64 address

    `fe80::213:d3ff:fe9c:1840%eth0`
  - at boot time, send Router Solicitation (RS) to request Router Advertisements (RAs)

IPv6 Subnets

```
$ sipcalc 2001:470:30:84:e276:63ff:fe72:3900/64
-[ipv6 : 2001:470:30:84:e276:63ff:fe72:3900/64] - 0

[IPV6 INFO]
Expanded Address        - 2001:0470:0030:0084:e276:63ff:fe72:3900
Compressed address      - 2001:470:30:84:e276:63ff:fe72:3900
Subnet prefix (masked)  - 2001:470:30:84:0:0:0:0/64
Address ID (masked)     - 0:0:0:0:e276:63ff:fe72:3900/64
Prefix address          - ffff:ffff:ffff:ffff:0:0:0:0
Prefix length           - 64
Address type            - Aggregatable Global Unicast Addresses
Network range           - 2001:0470:0030:0084:0000:0000:0000:0000 -
                          2001:0470:0030:0084:ffff:ffff:ffff:ffff
```

IPv6 Subnets: Common CIDRs

```
2001:0db8:0123:4567:89ab:cdef:1234:5678
|||| |||| |||| |||| |||| |||| |||| |||128   Single end-points and loopback
|||| |||| |||| |||| |||| |||| |||| ||124
|||| |||| |||| |||| |||| |||| |||| |120
|||| |||| |||| |||| |||| |||| |||| 116
|||| |||| |||| |||| |||| |||| |||112
|||| |||| |||| |||| |||| |||| ||108
|||| |||| |||| |||| |||| |||| |104
|||| |||| |||| |||| |||| |||| 100
|||| |||| |||| |||| |||| |||96
|||| |||| |||| |||| |||| ||92
|||| |||| |||| |||| |||| |88
|||| |||| |||| |||| |||| 84
|||| |||| |||| |||| |||80
|||| |||| |||| |||| ||76
|||| |||| |||| |||| |72
|||| |||| |||| |||| 68
|||| |||| |||| |||64                        Single End-user LAN (default prefix size for SLAAC)
|||| |||| |||| ||60
|||| |||| |||| |56                          Proposed minimal end sites assignment
|||| |||| |||| 52
|||| |||| |||48                              Default end sites assignment
|||| |||| ||44
|||| |||| |40
|||| |||| 36
|||| |||32                                  Local Internet registry minimum allocations
|||| ||28                                   Local Internet registry medium allocations
|||| |24                                    Local Internet registry large allocations
|||| 20                                     Local Internet registry extra large allocations
|||16
||12                                       Regional Internet Registry allocations from IANA
|8
```

Hooray!

5 Minute Break

Networking Buzzwords

"The network is the computer."

John Gage, Sun Microsystems

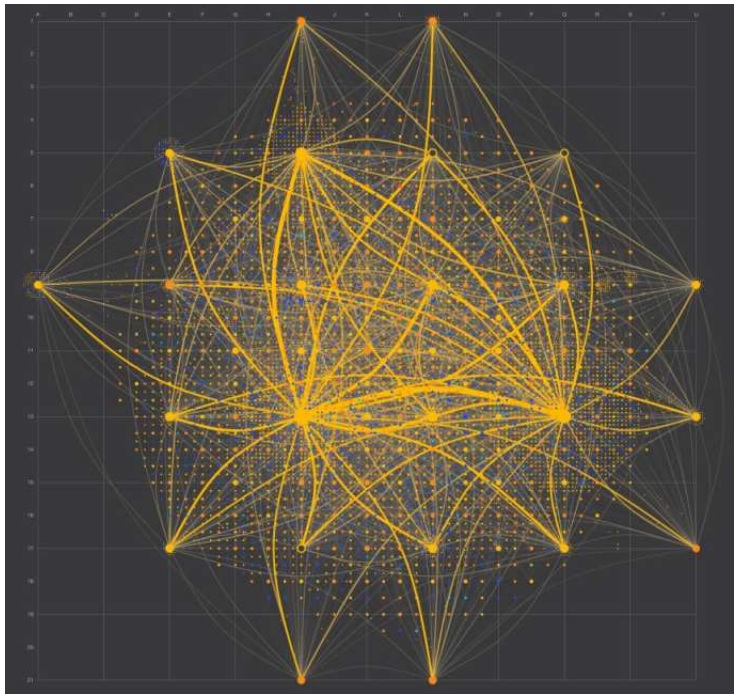
Networking Buzzwords

"The network is the
network,
the computer is the
computer -
sorry about the confusion."

Joe on Computing

Networking Buzzwords

Networking



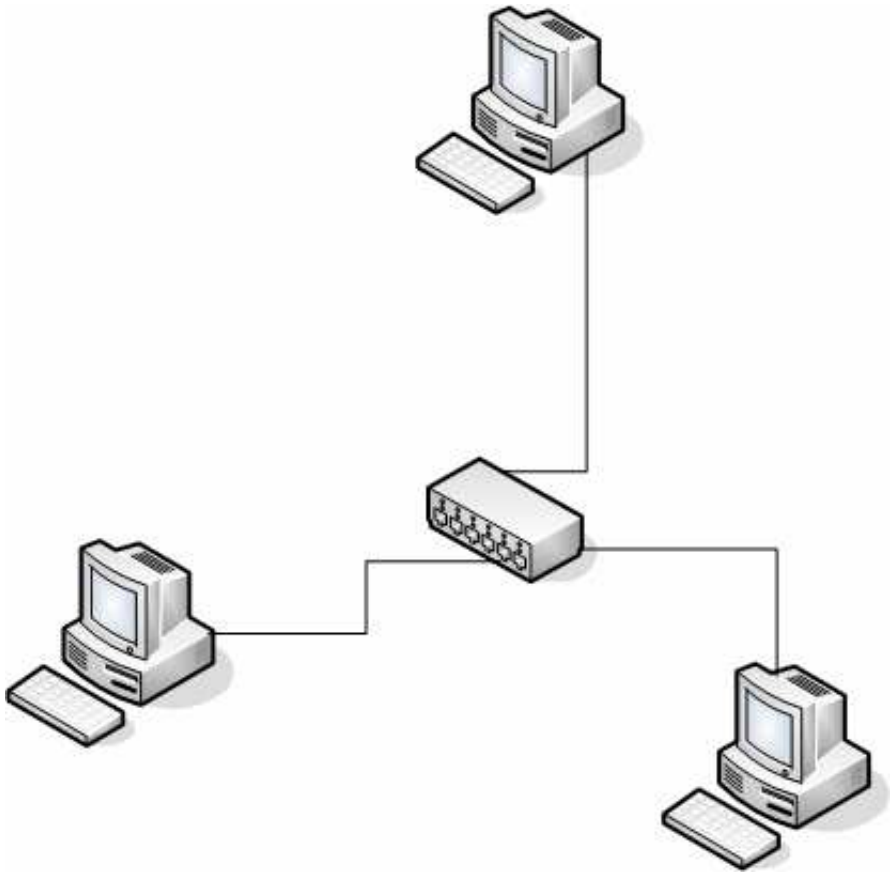
http://www.chrisharrison.net/index.php/Visualizations/InternetMap

Networking

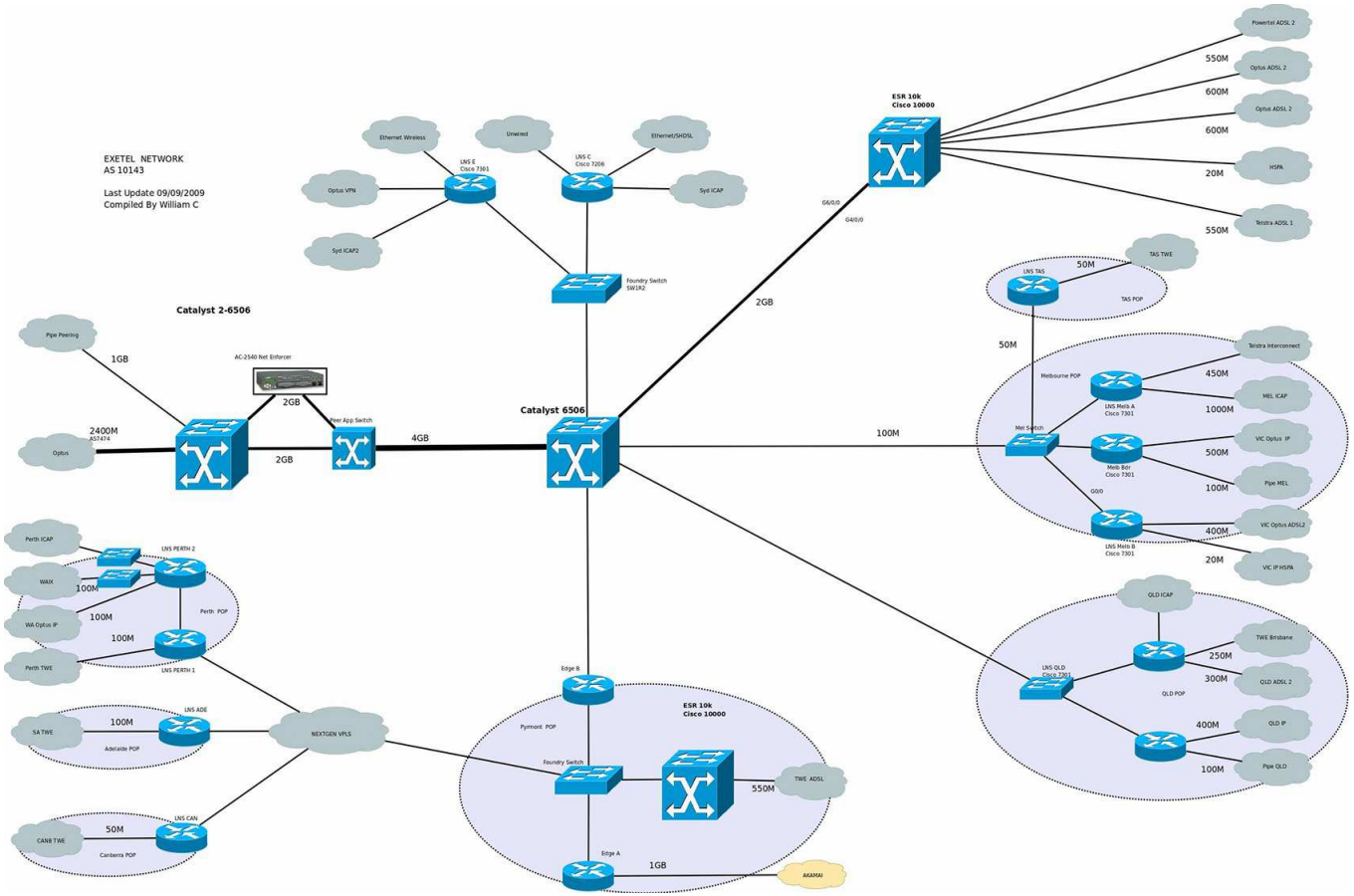
/X?

Networking

/X?



Networking


EXETEL NETWORK
AS 10143
Last Update 09/09/2009
Compiled By William C
Catalyst 2-6506
Catalyst 6506
ESR 10k
Cisco 10000
2GB
4GB
1GB
100M
2400M
Optus
Pipe Peering
AC-2540 Net Enforcer
Fiber App Switch
Ethernet Wireless
Unwired
Ethernet/SHDSL
LNS E
Cisco 7301
LNS C
Cisco 7206
Optus VPN
Syd ICAP
Syd ICAP2
Foundry Switch
SW1R2
G6/0/0
G4/0/0
Powertel ADSL 2
Optus ADSL 2
Optus ADSL 2
HSPA
Telstra ADSL 1
550M
600M
600M
20M
550M
LNS TAS
TAS TWE
TAS POP
50M
Melbourne POP
LNS Melb A
Cisco 7301
Mel Switch
Melb Bdr
Cisco 7301
LNS Melb B
Cisco 7301
Telstra Interconnect
MEL ICAP
VIC Optus IP
Pipe MEL
VIC Optus ADSL2
VIC IP HSPA
450M
1000M
500M
100M
400M
20M
QLD ICAP
LNS QLD
Cisco 7301
QLD POP
TWE Brisbane
QLD ADSL 2
QLD IP
Pipe QLD
250M
300M
400M
100M
Perth ICAP
WAIX
WA Optus IP
Perth TWE
LNS PERTH 2
LNS PERTH 1
Perth POP
100M
100M
100M
SA TWE
LNS ADE
Adelaide POP
NEXTGEN VPLS
CANB TWE
LNS CAN
Canberra POP
50M
Edge B
Edge A
Pyrmont POP
ESR 10k
Cisco 10000
Foundry Switch
TWE ADSL
550M
1GB
AKAMAI

WHOIS ASN?

The Internet Assigned Numbers Authority (IANA) oversees global IP address/AS number allocation, root zone management etc.

`https://www.iana.org/`

WHOIS ASN?

Autonomous System Numbers (ASNs) are assigned by IANA to the RIRs, see e.g. `ftp://ftp.arin.net/pub/stats/arin/`

You can query databases on the internet about e.g. IP block or ASN information via the `WHOIS` protocol:

```
$ whois 155.246.89.100 | more
NetRange:       155.246.0.0 - 155.246.255.255
CIDR:           155.246.0.0/16
NetName:        STEVENS
NetHandle:      NET-155-246-0-0-1
Parent:         NET155 (NET-155-0-0-0-0)
NetType:        Direct Assignment
Organization:   Stevens Institute of Technology (SIT)
RegDate:        1991-12-31
Updated:        2007-01-29
Ref:            https://whois.arin.net/rest/net/NET-155-246-0-0-1
```

WHOIS ASN?

Carriers connect their Autonomous Systems at *Internet Exchange Points* (IXPs) to route traffic between the different networks.

This *peering* happens amongst carriers on a tiered basis.

Examples:

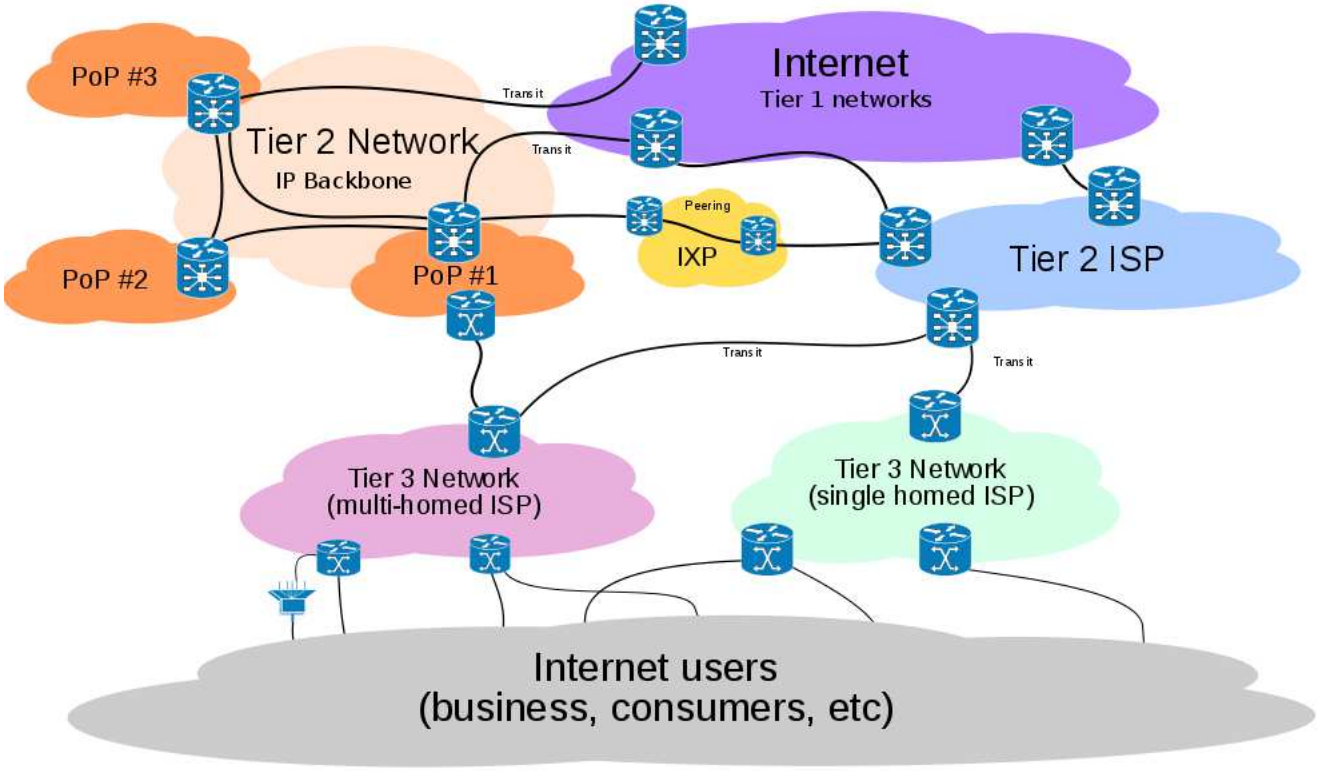
`https://peeringdb.com/net?asn=6939`
`https://peeringdb.com/net/27`
`https://peeringdb.com/net/433`
`https://peeringdb.com/net/457`

WHOIS ASN?

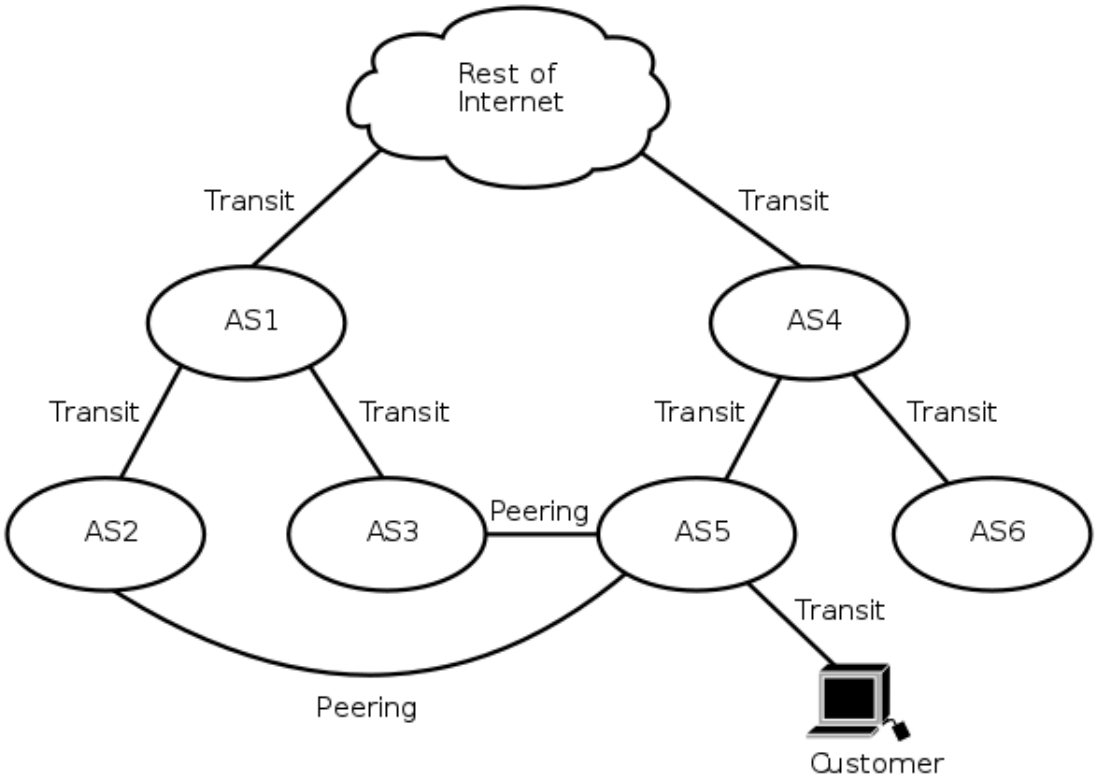
Most of these services are available via APIs or text-based interfaces:

```
$ host www.google.com
www.google.com has address 172.217.0.36
www.google.com has IPv6 address 2607:f8b0:4006:807::2004
$ whois -h whois.cymru.com 2607:f8b0:4006:807::2004
AS      | IP                                       | AS Name
15169   | 2607:f8b0:4006:807::2004                 | GOOGLE - Google Inc., US
$ curl -s https://peeringdb.com/api/net?asn=15169 | python -mjson.tool | more
{ "data": [ {
            "aka": "Google, YouTube (for Google Fiber see AS16591 record)",
            "created": "2005-02-06T06:41:04Z",
            "id": 433,
            "info_ipv6": true,
            "info_prefixes4": 15000,
            "info_prefixes6": 750,
            "info_ratio": "Mostly Outbound",
```

Networking


PoP #3
Transit
Internet
Tier 1 networks
Tier 2 Network
IP Backbone
Transit
Peering
IXP
PoP #2
PoP #1
Tier 2 ISP
Transit
Transit
Tier 3 Network
(multi-homed ISP)
Tier 3 Network
(single homed ISP)
Internet users
(business, consumers, etc)

Networking


Rest of Internet
Transit
Transit
AS1
AS4
Transit
Transit
Transit
Transit
AS2
AS3
Peering
AS5
AS6
Transit
Peering
Customer

Networking

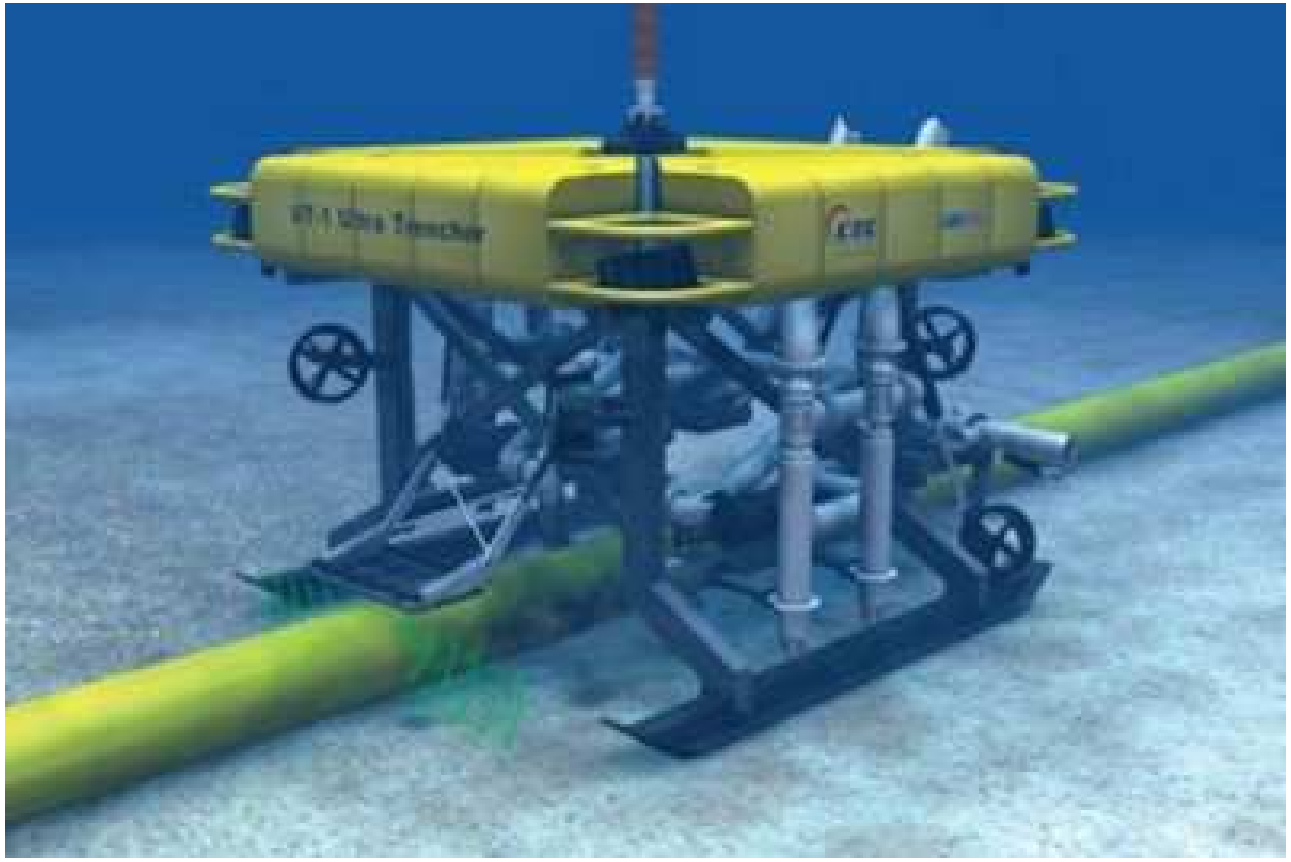
To find the path your packets might take, give `traceroute(1)` a go:

```
$ traceroute search.yahoo.com
traceroute to search.yahoo.com (63.250.200.63), 30 hops max, 60 byte packets
 1  155.246.89.2 (155.246.89.2)  0.342 ms postal0.cs.stevens-tech.edu (155.246.89.3)
 2  155.246.89.2 (155.246.89.2)  0.311 ms  0.300 ms gwa.cc.stevens.edu (155.246.151.3
 3  454a0465.cst.lightpath.net (69.74.4.101)  3.984 ms 3.761 ms  3.735 ms
 4  18267502.cst.lightpath.net (24.38.117.2)  32.559 ms  32.591 ms  32.577 ms
 5  hunt183-154.optonline.net (167.206.183.154)  4.473 ms  4.634 ms 18267502.cst.ligh
 6  451be0a9.cst.lightpath.net (65.19.113.169)  5.170 ms  5.278 ms hunt183-154.optonl
 7  nyiix.bas1-m.nyc.yahoo.com (198.32.160.121)  6.928 ms 451be0a9.cst.lightpath.net
 8  ae-1.pat2.bfw.yahoo.com (216.115.111.26)  26.422 ms ae-1.pat1.bfw.yahoo.com (216.
 9  et-18-1-0.msr1.bf2.yahoo.com (74.6.227.37)  17.812 ms et-18-1-0.msr2.bf1.yahoo.co
10  et-0-1-1.clr1-a-gdc.bf1.yahoo.com (74.6.122.15) 18.817 ms et-0-1-1.clr2-a-gdc.bf1
```

Networking

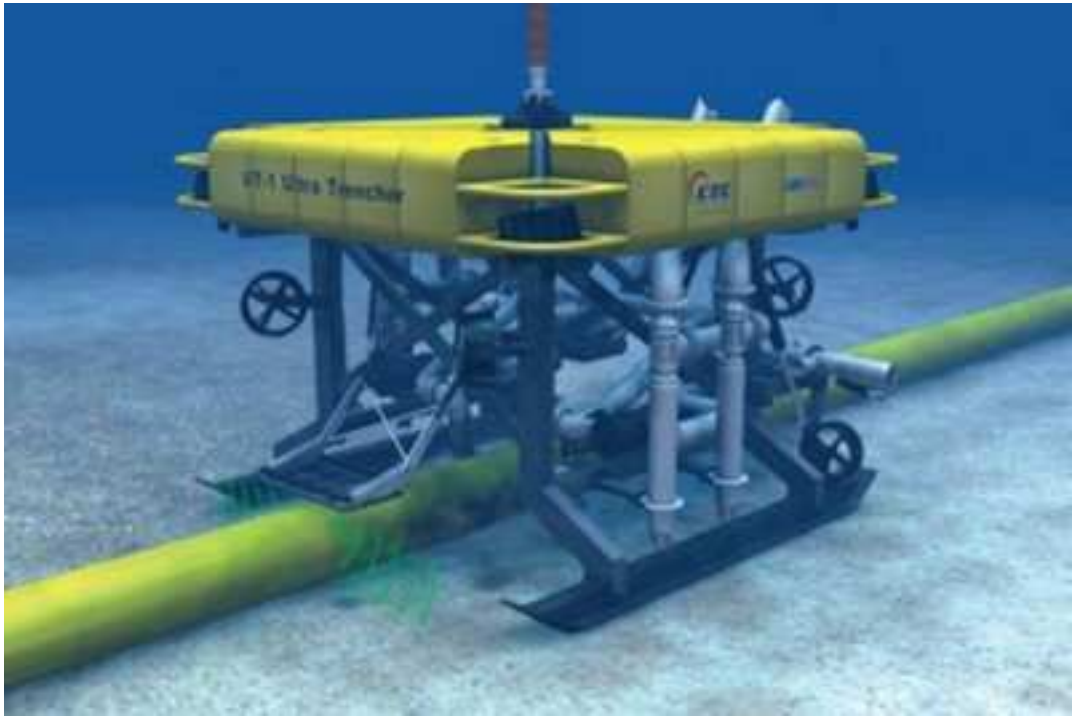
Networking

Networking



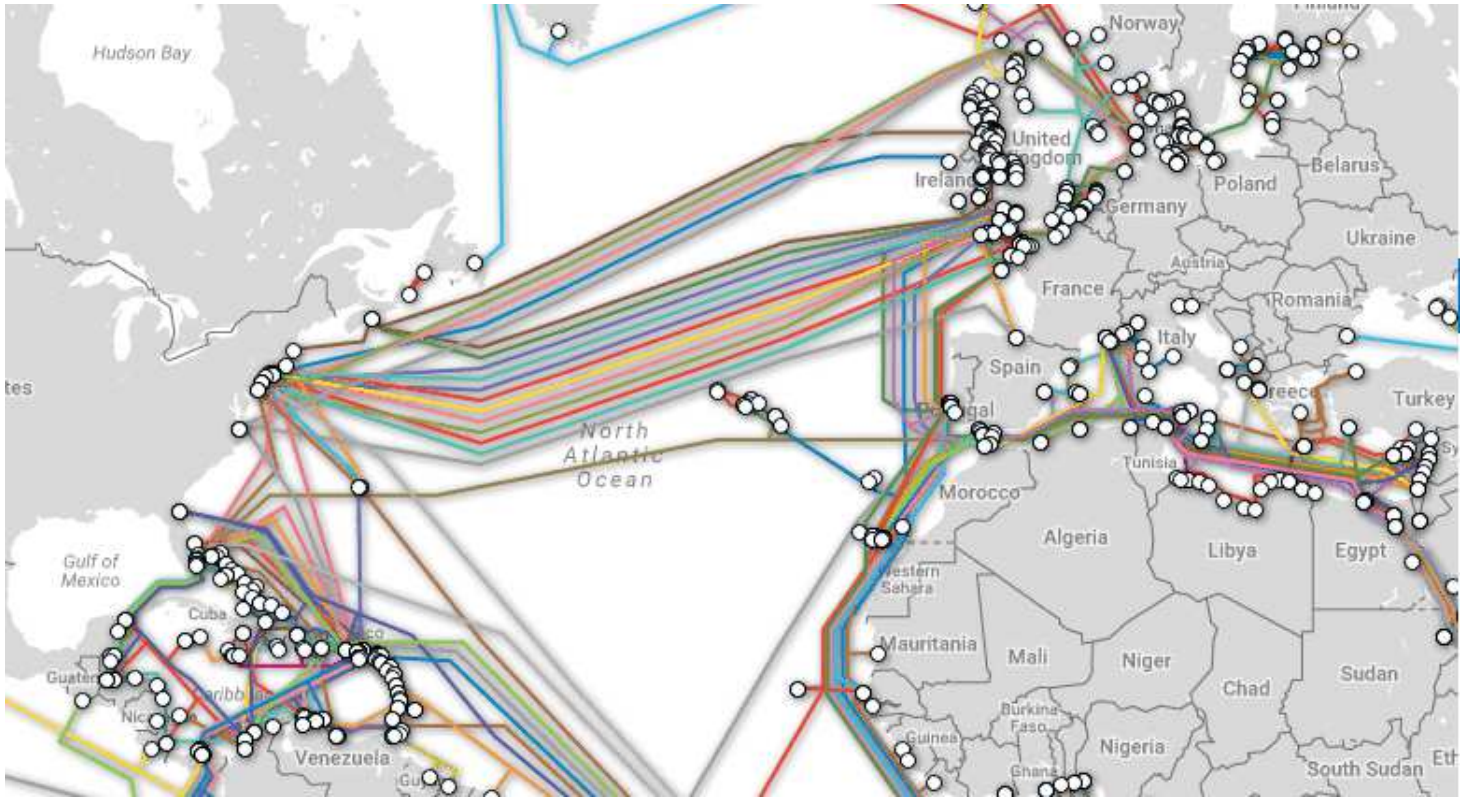
Networking

Stringing cables across the oceans' floors since 1866!



http://www.submarinecablemap.com/
http://is.gd/CjanOu

Networking

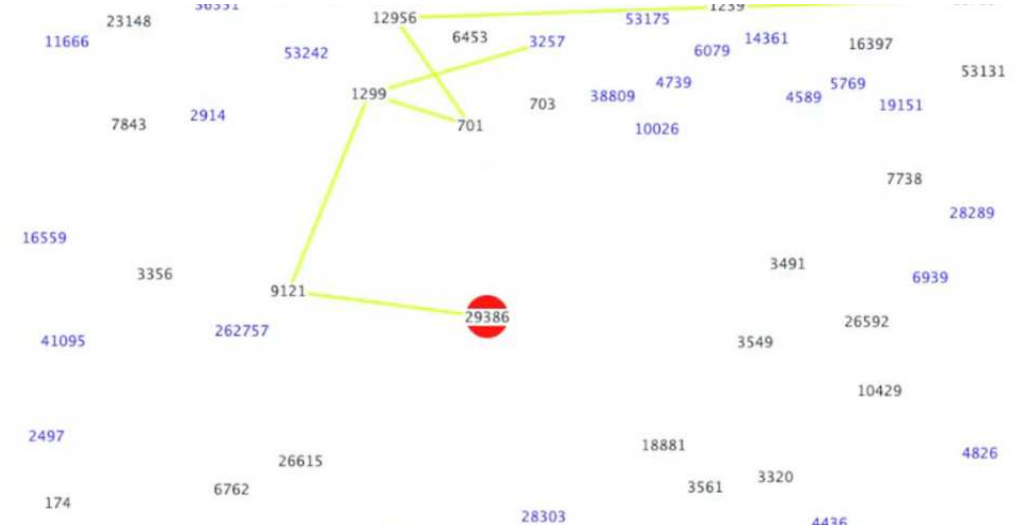


http://www.submarinecablemap.com/

Networking

"The Net interprets censorship as damage and routes around it."

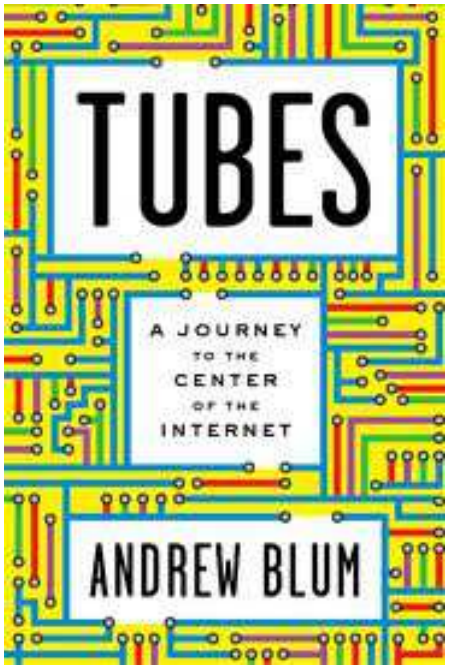
...except when it can't.



http://blog.cloudflare.com/how-syria-turned-off-the-internet

http://player.vimeo.com/video/54630037

Networking



http://amzn.com/0061994952

http://cromwell-intl.com/travel/usa/new-york-internet/

Networking

The internet is a physical place.

https://en.wikipedia.org/wiki/Room_641A

Internet Maps and Architecture

- http://www.peer1.com/map-of-the-internet
- http://is.gd/VxsE7S
- http://www.submarinecablemap.com/
- http://en.wikipedia.org/wiki/Peering
- http://is.gd/tpPNE5
- http://is.gd/B0d3kh
- http://amzn.com/0061994936
- http://bgp.he.net/

IPv6

- http://www.potaroo.net/papers/isoc/2005-07/ipv6size.html
- http://bgp.he.net/ipv6-progress-report.cgi
- https://ipv6.he.net/statistics/
- http://tunnelbroker.net/

Reading

- https://is.gd/qXVo2j
- http://www.wired.com/2015/08/shark_cable/

Commands:

- tcpdump(8)
- ktrace(1) / strace(1)
- tcp(4)/ip(4)
- netstat(1)
- nslookup(1)