

Internet Architecture

Web Architecture and Information Management [./] Spring 2009 — INFO 190-02 (CCN 42509)

Erik Wilde, UC Berkeley School of

Information

2009-02-18



<http://creativecommons.org/licenses/by/3.0/>

This work is licensed under a [CC Attribution 3.0 Unported License](http://creativecommons.org/licenses/by/3.0/) [http://creativecommons.org/licenses/by/3.0/]

Contents

• Abstract	2
• 1 Computer Networks	
◦ Network History	4
◦ Networks vs. Internetworks	5
◦ Internet	6
◦ Internet Principles	7
◦ Internet Protocols	8
◦ Network Convergence	9
• 2 Internet Protocol (IP)	
◦ IP Features	11
◦ IP Address	12
• 3 Transmission Control Protocol (TCP)	
◦ TCP Features	14
◦ Reliable Connections	15
◦ TCP Window	16
• 4 Domain Name System (DNS)	
◦ Naming vs. Addressing	18
◦ DNS Properties	19
◦ Names Matter	20
◦ Domain Name Space	21
◦ DNS Namespace Organization	22
◦ Using DNS	23
• Conclusions	24

Abstract

(2)

The Internet is the technical infrastructure on top of which the Web is built. Some of the services provided by the Internet are essential for the Web, most importantly the naming service and the data transfer service. The *Domain Name System (DNS)* provides the human-readable names for computers, which can then be used in the addresses of Web servers and ultimately Web pages. The *Transmission Control Protocol (TCP)* provides the reliable data transfer service between Web Servers and Web Browsers, building on the very robust *Internet Protocol (IP)*.

Computer Networks

Network History

(4)

- First regarded as a convenient workaround for floppy disks
 - “real computer scientists write compilers”
 - the value of computer networks depends on their size
- Early networking solutions were vendor-specific islands
 - DECnet for *Digital Equipment Corporation (DEC)* customers
 - XNS for *Xerox* customers
 - SNA for *IBM* customers
 - transmitting data between these networks was very cumbersome
- Bridging networks transparently became increasingly important
 - more computers and networks increase the benefit of interconnections
 - layering being used for internetworks, not only for networks

Networks vs. Internetworks (5)

- Specific networks use specific abstractions
 - how to address nodes (computers, phones, PDAs, RFID tags)
 - how to address applications on these nodes
 - how to transmit data to these applications
- Internetworks provide a network-independent abstraction
 - nodes are addressed uniformly (IP addresses)
 - applications are identified uniformly (ports)
 - data transmission uses one set of protocols (TCP/UDP)

Internet (6)

- Very early start and a lot of experience
 - pragmatic and evolutionary approach
 - "if it's not broken, don't fix it"
- Standardization by independent technical experts
 - avoids the "designed by committee" effect of consortiums
 - conservative and concentrating on stability
 - implementations are required to prove technical feasibility
 - simplicity whenever possible

Internet Principles (7)

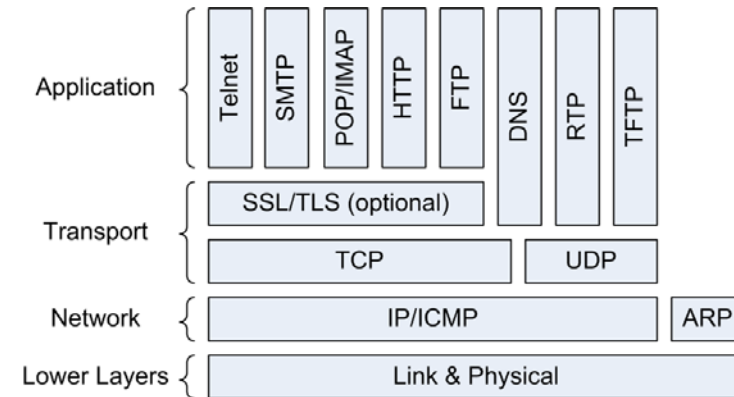
Be liberal in what you accept, and conservative in what you send.

[Jon Postel](http://www.postel.org/postel.html) [http://www.postel.org/postel.html], [RFC 1122](http://dret.net/rfc-index/reference/RFC1122) [http://dret.net/rfc-index/reference/RFC1122]

Whenever possible, communications protocol operations should be defined to occur at the end-points of a communications system, or as close as possible to the resource being controlled.

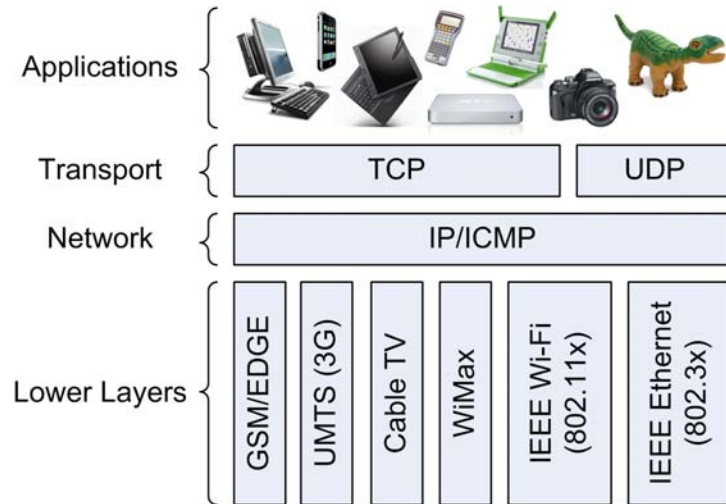
[J. Saltzer, D. Reed, D. Clark, "End-to-end Arguments in System Design"](http://dret.net/biblio/reference/sal84) [http://dret.net/biblio/reference/sal84]

Internet Protocols (8)



Network Convergence

(9)



Internet Protocol (IP)

IP Features

(11)

- End-to-end data transfer (IP addresses)
- Hiding lower-level heterogeneity
- Connection-less (each packet routed individually)
- Unreliable (packets may be lost or duplicated)

IP Address

(12)

- IP identifies nodes by an IP address
- IP addresses are globally unique ([and can be geocoded](http://api.hostip.info/get_html.php?position=true))
- IP uses 4 bytes for addresses (e.g., 128.32.226.29)
 - maximum number of addresses: $2^{32} = 4$ billion
 - IPv6 extends the address format to 16 bytes (2^{128} addresses)
- IP address shortage led to the some trickery using IP addresses
 - [Dynamic Host Configuration Protocol \(DHCP\)](http://en.wikipedia.org/wiki/Dhcp) is used to assign addresses on-demand
 - [Network Address Translation \(NAT\)](http://en.wikipedia.org/wiki/Network_address_translation) uses one IP address for more than one device
- IP addresses are well-organized
 - important for routing (i.e., sending packets to the target host)
 - not ideally suited for mobile or ad-hoc networks

Transmission Control Protocol (TCP)

TCP Features

(14)

- Flow-controlled (avoiding congestion)
- Reliable (no data lost or duplicated)
- Connection-oriented
- Application addressing

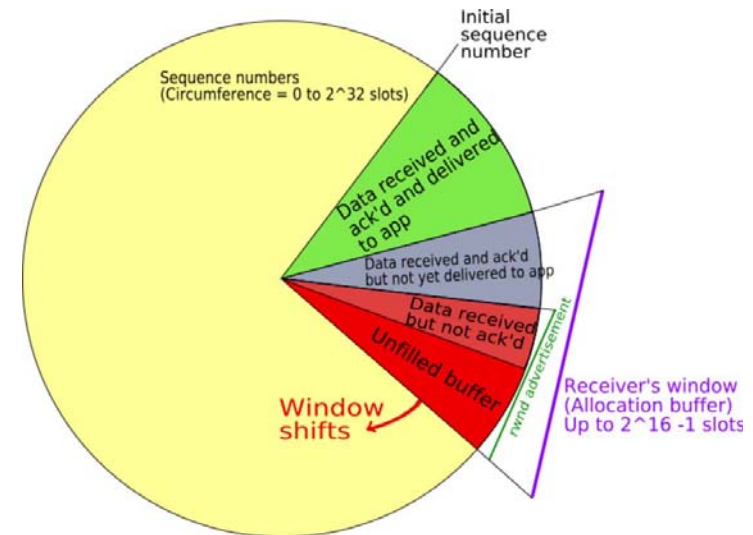
Reliable Connections

(15)

- IP may drop or duplicate packets
 - TCP adds serial numbers in data packets
 - if problems are detected, TCP recovers automatically
- TCP avoids network congestion and system overload
 - *slow start* avoid flooding receivers with data they cannot process
 - *fast retransmit* for avoiding timeouts when losing data
 - a *sliding window* for controlling the amount of outstanding packets

TCP Window

(16)



Domain Name System (DNS)

Naming vs. Addressing (18)

- IP addresses depend on network topology and organization
 - reorganizing a network may change all IP addresses
 - identifying important hosts should not be address-based
- Names are supposed to be more stable than addresses
 - a name is an abstract identification of something
 - names can be used to obtain more information
- Network services should use names instead of addresses
 - before using the service, a mapping has to be performed
 - the *Domain Name System (DNS)* is providing this service

DNS Properties (19)

- DNS has a bootstrap problem
 - DNS provides a service and should thus be identified by a name
 - for resolving names into addresses, the DNS service is required
- DNS configuration is part of basic Internet configuration
 - *Dynamic Host Configuration Protocol (DHCP)* provides [IP Address](#) [IP Address (1)], netmask, gateway, and DNS server address
- DNS names are hierarchically structured
 - `ischoo1.berkeley.edu`, `edu` is the *Top-Level Domain (TLD)*
 - TLDs are either *generic (gTLD)* or *country code (ccTLD)*
 - subdomains are federated (e.g., `edu`, `us`, `uk`, `tv`)

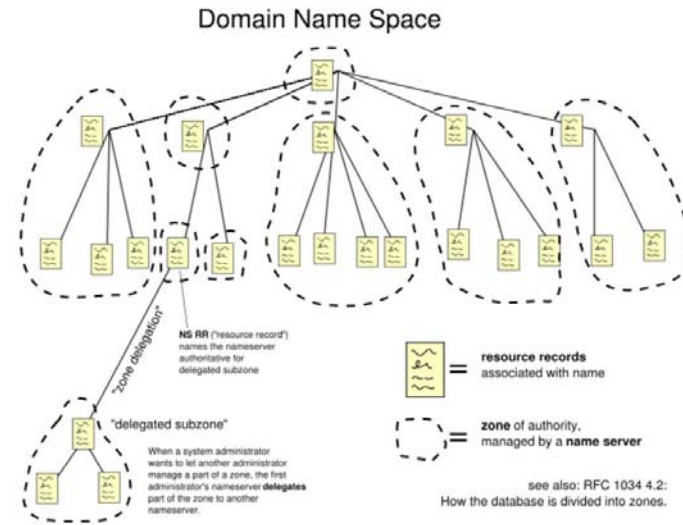
Names Matter

(20)

- Names are not unique and namespaces are finite
 - name disputes arise which were irrelevant before the Web
 - "cybersquatting" as a popular way to make money
- Names can be worth a lot of money
 - business.com was sold for \$7.5 million
- Name inflation can be used to generate money
 - aero, biz, coop, info, jobs, mobi, museum, name, pro, travel
 - starting 2009, [user-defined top-level domains will go on sale](http://dret.typepad.com/dretblog/2008/06/dret.html) [http://dret.typepad.com/dretblog/2008/06/dret.html]
- Names can have political significance
 - ccTLDs are assigned based on the UNO's idea of what a country is
- Names can have symbolic significance
 - Catalonia managed to get a domain of its own (cat)

Domain Name Space

(21)



DNS Namespace Organization (22)

- Domain owners can organize the assignment of subdomains
 - [berkeley.edu](http://www.berkeley.edu/) [http://www.berkeley.edu/] is an U.S. educational institution
 - [ethz.ch](http://www.ethz.ch/) [http://www.ethz.ch/] is a Swiss university
 - [imperial.ac.uk](http://www.imperial.ac.uk/) [http://www.imperial.ac.uk/] is a British university
 - [uts.edu.au](http://www.uts.edu.au/) [http://www.uts.edu.au/] is an Australian university
- Special rules may apply (Germany does not assign car license plate names)
- Organizations may be countries or companies
 - countries have national organizations for assigning names
 - companies may create an internal multi-level namespace (www.ischool.berkeley.edu [http://www.ischool.berkeley.edu/])

Using DNS (23)

- DNS is used by virtually all Internet applications
 - names are more stable than addresses
- E-mail has some dedicated features built into DNS
 - special entries (MX records) identify the e-mail server for a domain
 - fallback entries help dealing with failing e-mail servers
- most URIs are based on DNS names
 - <http://ischool.berkeley.edu/> identifies the access protocol and the host
 - the browser first performs a DNS lookup
 - a TCP connection is then established to the address returned by the DNS

Conclusions

(24)

- The Internet is a network of networks
 - [Ted Stevens](http://en.wikipedia.org/wiki/Ted_Stevens) [http://en.wikipedia.org/wiki/Ted_Stevens]: “[It's not a big truck, it's a series of tubes!](http://www.youtube.com/watch?v=f99PcP0aFNE)” [http://www.youtube.com/watch?v=f99PcP0aFNE]”
- The Internet provides basic data transfer services
- IP is used to send *datagrams* between end-points
- TCP is used for reliable communications (building on IP)
- DNS provides human-readable names for Internet hosts