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

Master Course Computer Networks IN2097

Prof. Dr.-Ing. Georg Carle
Christian Grothoff, Ph.D.

Chair for Network Architectures and Services
Institut für Informatik
Technische Universität München
http://www.net.in.tum.de

Technische Universität München

Outline

- Project announcements
- Recapitulation on last lectures
- Internet development
- Node property fundamentals: delay, loss, throughput

Project announcements

- Currently 30 teams
- If you did not register so far, write Email to guenther@in.tum.de
- SVN accounts: planned available by Monday evening, Nov 7th
- Submission 1 - Project plan - due by Tuesday evening, Nov 8th
- Submission 2 - IPv6 today - due by Tuesday evening, Nov 15th
- Submission 3 - Your own Site - due by Thursday Dec 15th

Recapitulation on last lectures

- DNS
- Tunneling
- IPv4
- IPv6

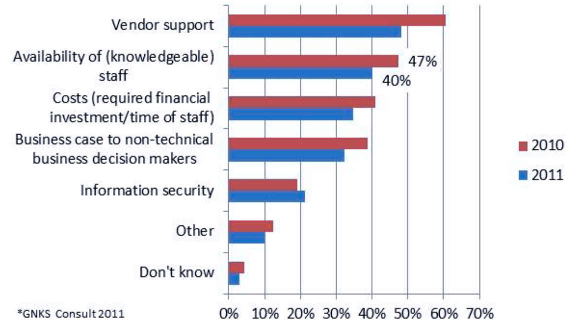
Chair for Network Architectures and Services – Prof. Carle
Department for Computer Science
TU München

IPv6 Deployment

Standardisation

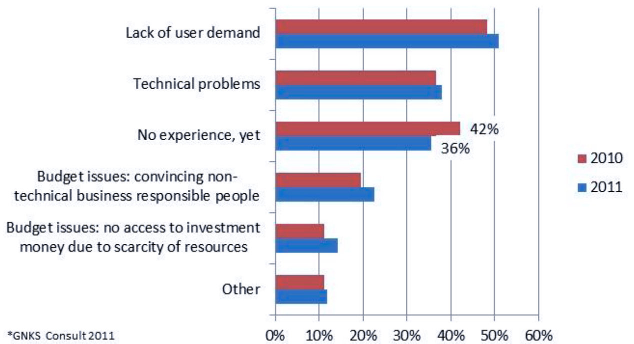
Technische Universität München

Biggest hurdles when deploying IPv6



- Maarten Botterman, GNKS Consult: Results of the 2011 Global IPv6 Deployment Monitoring Survey - Presentation at RIPE-63

Biggest problems with IPv6 in practice



RFC 2460: IPv6 Specification

- The routing header is used by an IPv6 source to list one or more intermediate nodes to be "visited" on the way to packet's destination.
- Each extension header should occur at most once, except for the destination options header which should occur at most twice.
- IPv6 nodes must accept and attempt to process extension headers in any order and occurring any number of times in the same packet.

- c.f. Merike Kaeo, merike@doubleshotsecurity.com Presentation „IPv6 Routing Header Security " - RIPE54 Meeting, Tallin, Estonia, May 2007

Router Configurations

- Cisco
  - "no ipv6 source-route„
- Linux
  - # Filter all packets that have RT0 headers
  - ip6tables -A INPUT -m rt--rt-type 0 -j DROP
  - ip6tables -A FORWARD -m rt--rt-type 0 -j DROP
  - ip6tables -A OUTPUT -m rt--rt-type 0 -j DROP
  - (of course before accepting anything else ;)
- FreeBSD
  - Upgrade the kernel with at least the following patch in place: http://www.freebsd.org/cgi/cvsweb.cgi/src/sys/netinet6/route6.c.diff?r1=1.12&r2=1.13

Routing Header Processing

- Disabling IPv6 type 0 routing header processing still allows other nodes to be used for attack
- Dropping is required for ISP's
- RFC 5095 - deprecate [„ablehnen"/"missbilligen"]

```
Network Working Group                                           J. Abley
Request for Comments: 5095                                       Afilias
Updates: 2460, 4294                                            P. Savola
Category: Standards Track                                      CSC/FUNET
                                                          G. Neville-Neil
                                                  Neville-Neil Consulting
                                                            December 2007

Deprecation of Type 0 Routing Headers in IPv6
Abstract
   The functionality provided by IPv6's Type 0 Routing Header can be
   exploited in order to achieve traffic amplification over a remote
   path for the purposes of generating denial-of-service traffic.  This
   document updates the IPv6 specification to deprecate the use of IPv6
   Type 0 Routing Headers, in light of this security concern.
```


IETF Structure and Internet Standards Process

Scott Bradner

Harvard University
http://www.sobco.com/sob/sob.html

77th IETF - March 2010
Anaheim, California, USA

The IETF - Internet Engineering Task Force

- Formed in 1986
  - evolved out of US government activities
  - ARPA's Internet Configuration Control Board (ICCB) (1979) and Internet Activities Board (1983)
- Was not considered important for a long time - good!!
- Not government approved - great!!
  - but funding support from U.S. Government until 1997
- Specifications always available without charge (vs. ITU-T, IEEE)
- People not companies

"We reject kings, presidents and voting.

We believe in rough consensus and running code"

Dave Clark (1992)

IETF Organisation

- 1K to 2K people at 3/year meetings (many more on mail lists)
- >100 working groups with working group chairs
- 8 areas with Area Directors (ADs):
  GEN, APS, INT, O&M, RAI, RTG, SEC, TSV:
  - IETF Chair & AD for General Area (gen) - 0 WGs
  - Applications (app) - 15 WGs
  - Internet (int) - 28 WGs
  - Operations & Management (ops) - 15 WGs
  - Real-time Applications and Infrastructure (rai) - 19 WGs
  - Routing (rtg) - 16 WGs
  - Security (sec) - 17 WGs
  - Transport Services (tsv) - 14 WGs
- Internet Enginnering Steering Group (IESG): ADs + IETF Chair
- Internet Architecture Board (IAB): architectural guidance, liaisons
- IETF produces standards and other documents

Working Groups

- no defined membership
  - just participants
- "***Rough consensus** and running code...*"
  - no formal voting - can not define constituency
    - can do show of hands or hum - but **no** count
  - does **not** require unanimity
  - chair determines if there is consensus
  - disputes resolved by discussion
  - mailing list and face-to-face meetings
  - final decisions must be verified on mailing list
    - to ensure those not present are included
      - but taking into account face-to-face discussion
- sessions are being streamed & recorded

IETF Standardisation Procedure

- Proposals published as Internet Drafts (ID)
- Worked on in a Working Group (WG)
- WG sends to IESG request to publish an ID 'when ready'
- proposal reviewed by AD
  - can be sent back to working group for more work
- IETF Last-Call
- IESG review
  - last call comments + own technical review
  - can be sent back to Working Group for more work
- publication as RFC

RFC Repository Contains:

- standards track
  - OSPF, IPv6, IPsec ...
- obsolete Standards
  - RIPv1
- requirements
  - Host Requirements
- policies
  - Classless Inter-Domain Routing
- april fool's day jokes
  - IP on Avian Carriers ...
  - ... updated for QoS
- poetry
  - 'Twas the night before startup
- white papers
  - On packet switches with infinite storage
- corporate documentation
  - Ascend multilink protocol (mp+)
- experimental history
  - Netblt
- process documents
  - IETF Standards Process

Standards Track RFCs

- Best Current Practices (BCP)
  - policies or procedures (best way we know how)
- 3-stage standards track (not all that well followed)
  - Proposed Standard (PS)
    - good idea, no known problems
  - Draft Standard (DS)
    - PS + stable
    - multiple interoperable implementations
    - note: interoperability not conformance
  - Internet Standard (STD)
    - DS + wide use
- *"The Internet runs on proposed standards"* – perhaps first said by Fred Baker, Cisco Fellow, IETF Chair 1996-2001

Challenge Interoperability

Example:
IPFIX Interoperability Test Event, 63rd IETF

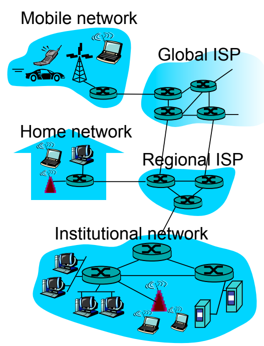
- Participants
  - CISCO
  - IBM Research Zürich
  - NEC Laboratories Heidelberg
  - Fraunhofer FOKUS, Berlin
  - University team of Prof. Carle
    - c.f. RFC 3333, 5477, 5815
- Lession learned:
  Organisation of interoperability activities is useful. We do not necessarily need to organize joint meetings, but should make more of a habit of organizing joint testing, e.g. combined with chat sessions.

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

Delay, loss and throughput

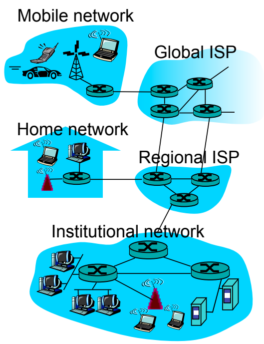
What's the Internet: "nuts and bolts" view

- millions of connected computing devices: *hosts = end systems*
  - running *network apps*
- *communication links*
  - fiber, copper, radio, satellite
  - transmission rate = *bandwidth*
- *routers:* forward packets (chunks of data)



What's the Internet: "nuts and bolts" view

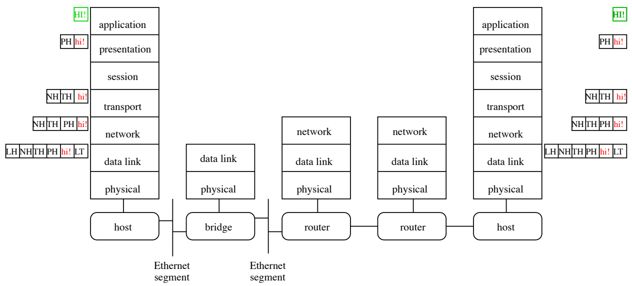
- *protocols* control sending, receiving of messages
  - e.g., TCP, IP, HTTP, Skype, Ethernet
- *Internet:* "network of networks"
  - loosely hierarchical
  - public Internet versus private intranet
- Internet standards
  - RFC: Request for comments
  - IETF: Internet Engineering Task Force
- communication *infrastructure* enables distributed applications:
  - Web, VoIP, email, games, e-commerce, file sharing
- communication services provided to applications:
  - reliable data delivery from source to destination
  - "best effort" (unreliable) data delivery



Protocol Mechanisms

- All or some of the following:
  - addressing/naming: manage identifiers
  - fragmentation: divide large message into smaller chunks to fit lower layer
  - resequencing: reorder out-of-sequence messages
  - error control: detection and correction of errors and losses
    - retransmission; forward error correction
  - flow control: avoid flooding/overwhelming of slower receiver
  - congestion control: avoid flooding of slower network nodes/links
  - resource allocation: administer bandwidth, buffers among contenders
  - multiplexing: combine several higher-layer sessions into one "channel"
  - compression: reduce data rate by encoding
  - privacy, authentication: security policy (others are listening)

Protocol Layering



- **send side** layer N takes protocol data (PDU) from layer N +1, adds header, and passed to N-1
- **receive side** layer N takes PDU from N –, strips N headers, processes and passes rest to N + 1

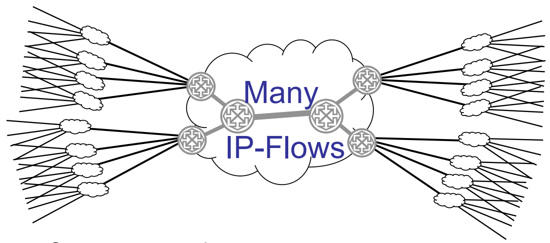
Layering Considered Harmful?

- Benefits of layering
  - need layers to manage complexity
    - don't want to reinvent Ethernet-specific protocol for each application
  - common functionality
    - "ideal" network
- but:
  - layer N may duplicate lower layer functionality (error recovery)
  - different layers may need same information
  - layer N may need to peek into layer N+x

Routers: Forwarding and Routing

- Forwarding: data plane
  - Directing a data packet to an outgoing link
  - Individual router using a forwarding table
- Routing: control plane
  - Computing the paths the packets will follow
  - Routers talking amongst themselves
  - Individual router creating a forwarding table

Goal: Scalability



- Core router requirements
  - Large number of IP flows
  - High packet rate
  - No 'per-Flow' state

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

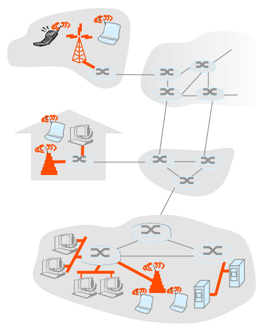
Access networks and physical media

Q: How to connect end systems to edge router?

- residential access networks
- institutional access networks (school, company)
- mobile access networks

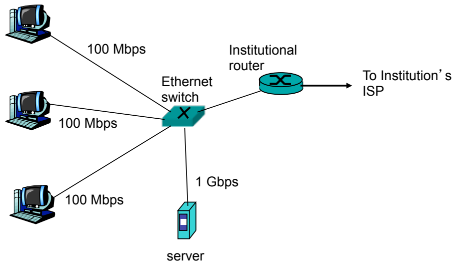
Relevant:

- bandwidth (bits per second) of access network?
- shared or dedicated?



Ethernet Internet access

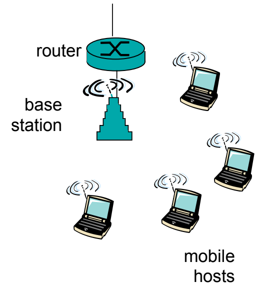
- Typically used in companies, universities, etc
  - 10 Mbs, 100Mbps, 1Gbps, 10Gbps Ethernet
  - Today, end systems typically connect into Ethernet switch



⇨ why?

Wireless access networks

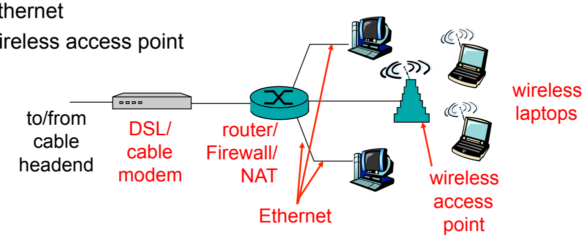
- shared *wireless* access network connects end system to router
  - via base station - "access point"
- wireless LANs:
  - 802.11b/g (WiFi): 11 or 54 Mbps
- wide-area wireless access
  - provided by telco operator
  - ~1Mbps over cellular system (HSDPA)
  - next cellular network technology: LTE (10's Mbps) over wide area



Home networks

Typical home network components:

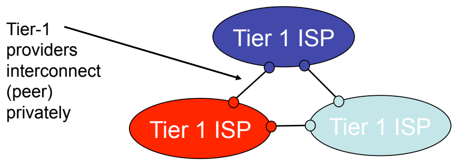
- DSL or cable modem
- router/firewall/NAT
- Ethernet
- wireless access point



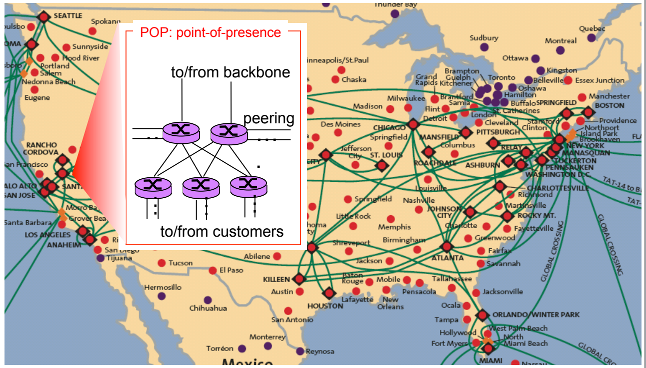
⇨ Research at chair I8: Autonomic Home Networks

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

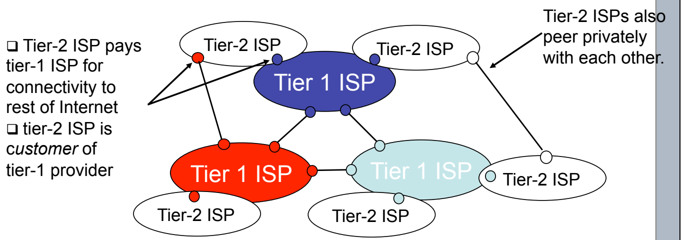


Tier-1 ISP: e.g., Sprint



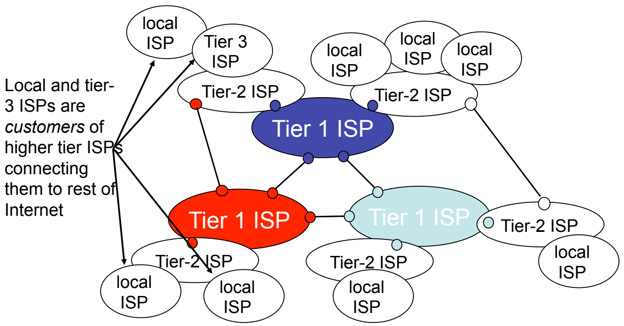
Internet structure: network of networks

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



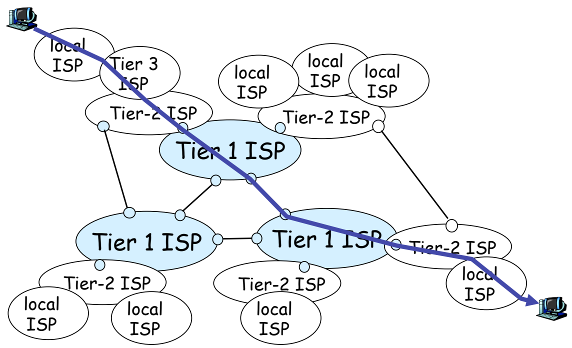
Internet structure: network of networks

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

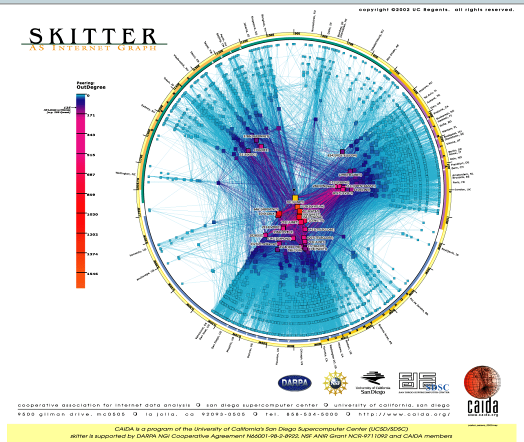


Internet structure: network of networks

- a packet passes through many networks!



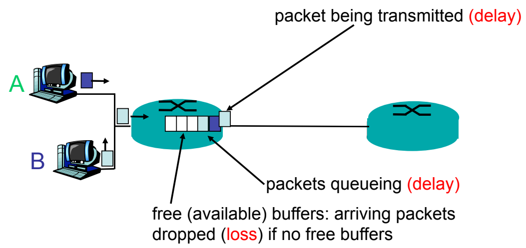
ISP Peering Relations



Reasons for delay and loss

packets *queue* in router buffers

- packet arrival rate to link exceeds output link capacity
- packets queue, wait for turn



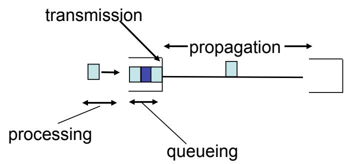
Background: Sources of packet delay

1. Processing delay:
   - Sending: prepare data for being transmitted
   - Receiving: interrupt handling

2. Queueing delay
   - time waiting at output link for transmission

3. Transmission delay:
   - L=packet length (bits)
   - R=link bandwidth (bps)
   - time to send bits into link = L/R

4. Propagation delay:
   - d = length of physical link
   - s = propagation speed in medium (~2×10^8 m/sec)
   - propagation delay = d/s



Nodal delay

- d_{proc} = processing delay
  - typically a few microseconds (µs) or less
- d_{queue} = queuing delay
  - depends on congestion - may be large
- d_{trans} = transmission delay
  - = L/R, significant for low-speed links
- d_{prop} = propagation delay
  - a few microseconds to hundreds of msecs

$$d_{nodal} = d_{proc} + d_{queue} + d_{trans} + d_{prop}$$

Impact Analysis: Advances in Network Technology

Data rate	Delay (1bit)	Length (1bit)	Delay (1kbyte)	Length (1kbyte)
1 Mbit/s	1 us	200 m	8 ms	1600 km
10 Mbit/s	100 ns	20 m	0,8 ms	160 km
100 Mbit/s	10 ns	2 m	80 us	16 km
1 Gbit/s	1 ns	0,2 m	8 us	1600 m
10 Gbit/s	100 ps	0,02 m	0,8 us	160 m
100 Gbit/s	10 ps	0,002 m	80 ns	16 m

- Assessment
  - Transmission delay becomes less important
    ⇨ over time; in the core
  - Distance becomes more important
    ⇨matters for communication beyond data center
  - Network adapter latency less important
    ⇨ Latency of communication software becomes important

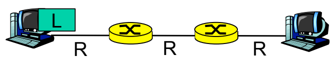
Propagation Delay

- Propagation speed: $2x10^8$ m/sec
- Transmission of 625 byte (= 5000 bit): t= L/R=5000 / 1Gbit/s = 5 us

Distance	Propagation Delay	equivalent Transmission Delay (625 byte)	CPU cycles per packet (1 GHz)	CPU cycles per byte (1 GHz)
100 m	500 ns	10 Gbit/s	500	<1
1 km	5 us	1 Gbit/s	5.000	8
10 km	50 us	100 Mbit/s	50.000	80
100 km	500 us	10 Mbit/s		800
1.000 km	5 ms	1 Mbit/s		8.000
10.000 km	50 ms	100 Kbit/s		80.000

- Suggestion for homework exercise: plot graphs

Store-and-Forward vs. Circuit Switching



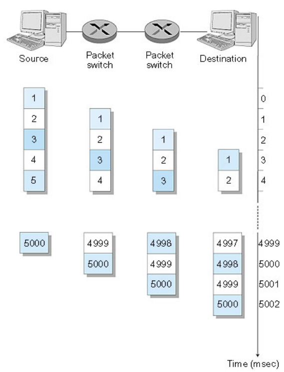
- Takes L/R seconds to transmit (push out) packet of L bits on to link or R bps
- Entire packet must arrive at router before it can be transmitted on next link: store and forward
- delay = 3L/R

Example: Large Message L

Circuit Switching:
- L = 7.5 Mbit
- R = 1.5 Mbit/s
- Transmission delay = 5 s

Store-and-Forward:
- L = 7.5 Mbit
- R = 1.5 Mbit/s
- Transmission delay = 15 s

Packet Switching: Message Segmenting



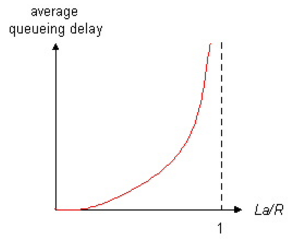
Now break up the message into 5000 packets
- Each packet 1,500 bits
- 1 msec to transmit packet on one link
- *pipelining:* each link works in parallel
- Delay reduced from 15 sec to 5.002 sec (as good as circuit switched)
- Advantages over circuit switching?
- Drawbacks (of packet vs. Message)

Queueing delay (revisited)

- R=link bandwidth (bit/s)
- L=packet length (bit)
- a=average packet arrival rate

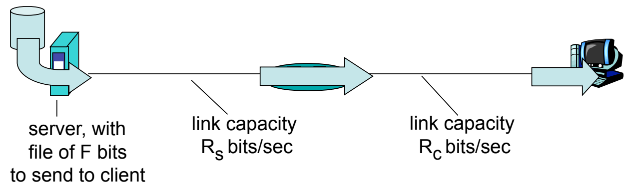
traffic intensity = a·L/R

- a·L/R ~ 0: average queuing delay small
- a·L/R → 1: delays become large
- a·L/R > 1: more "work" arriving than can be serviced, average delay infinite!



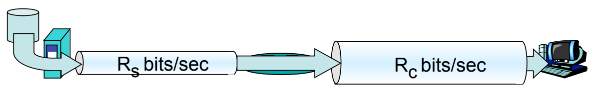
Throughput

- *throughput:* rate (bits/time unit) at which bits transferred between sender/receiver
  - *instantaneous:* rate at given point in time
  - *average:* rate over longer period of time

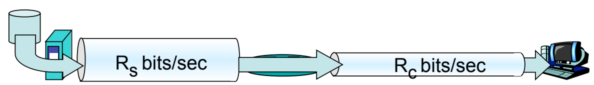


Throughput (more)

- $R_s < R_c$



- $R_s > R_c$



bottleneck link

link on end-end path that constrains end-end throughput

⇨ measurement challenge for networks with many nodes: identify bottleneck interfaces, e.g. with packet-pair measurements

Discussion

- What is the role of header lengths?
- What is the role of header compression?
- What is the cost of tunneling?
- What are the benefits of overprovisioning?
- Can you „imagine" a visualisation of packets being transmitted over different types of links?

Questions

- Why is circuit switching expensive?
- Why is packet switching cheap?
- Is best effort packet switching able to carry voice communication?
- What happens if we introduce "better than best effort" service?
- How can we charge fairly for Internet services: by time, by volume, or flat?
- Who owns the Internet?
- You've invented a new protocol. What do you do?
- How does the Internet grow? Exponentially? What is the growth perspective?

Thank you

for your attention!

Your Questions?