Applications and TCP

Markus Peuhkuri

2001-09-18

Lecture Topics

- What network requirements
- Requirements by application
- Requirements for elastic flows (or look-alike)

Definition of properties

Throughput [1] is the number of binary digits that a system accepts and delivers per unit of time. Traditionally throughput at network level is expressed as bits per second and ”kilo” equals 1,000. At application level also bytes or octets per second is used and ”kilo” equals *usually* 1024. To avoid vagueness, one should use new SI prefixes of kibi, mebi and gibi when multiplicative prefixes of 2^{10*3^N} are used.

Delay [2]b (transit-) is the time elapsing between the emission of the first bit of data block by transmitting system and its reception by the receiving system. Other delays include:

access delay is time to wait when data is ready before transmission may start

transmission delay or *serialisation* delay is time needed to transmit all of the data block

Delay variation expresses how much delay of successive data units vary. Short timescale variation is called jitter[3].

Error rate is count of errors per transmitted data or per time unit. (*PER*: packet error rate) Possible errors are:

data alteration usually referred as bit errors. Rare in optical networks but may take place inside networking equipment because of faulty hardware or software. More frequent in wireless networks. (*BER*: bit error rate)

data loss is the most common error in packet networks when a packet is discarded because of congestion. It is possible that intermediate system discards data unit because of data alteration. (*PLR*: packet loss rate)

data duplication occurs mostly because of transitory malfunction of network equipment

Network building blocks

- End systems (hosts)
  - servers
  - PCs, workstations
  - mobile devices
- Routers

[1] Läpäisy
[2] viive
[3] Värinä

  - access routers: a large number of slow-speed interfaces and few high-speed interfaces. Extensive authentication, accounting and filtering facilities.
  - core routers: possibly a large number of high-speed interfaces. Optimised for high-speed routing with large routing tables.
  - application gateways / NAT-boxes basically divide network into two different domains by limiting network visibility end to end.

- Links interconnecting routers and end systems
  - point-to-point: dedicated connection; serial lines, PCM/SDH pipes
  - shared media: fixed total capacity for all systems; ”old” Ethernet and wireless LANs
  - overlay network: virtual network hiding physical topology; ATM, MPLS

End system properties

Throughput is determined by speed of

- CPU
- internal busses: peripheral busses (PCI in PCs) and memory bus
- storage devices
- network interface

Delay same factors as for throughput, data processing . For example, many audio compression methods need some length of audio data before they can compress.

Delay variation increases by system load

- interrupt latency
- processor sharing (scheduling)

Error rate is usually small

data alteration memory or DMA (Direct Memory Access) errors

data loss lack of buffer space or too slow processing of data

Router properties

Throughput Modern routers can deliver data at full line speed. Using some advanced features such as constrain based routing or accounting may drop performance by a factor.

Delay is dependent on traffic load and available buffer space. Small buffers result in small delay but high data loss rate.

Delay variation is caused by queue length variations.

Error rate is usually small

data alteration may happen if there is hardware error

data loss is the most common error as queues grow full

data duplication happens in routers because of transitory state in packet forwarding. For example a routing table may change just when a packet is being forwarded and the packet is transmitted to both old and new route.

Link properties

Throughput is set by selected technology and depends on

- distance, specially in wireless networks
- number of hosts, if media is multi-access

Delay

transit delay is set by distance and signal propagation speed. For optical cables one can use speed of 200,000 km/s (5 ms/1000 km).

transmission delay depends on link speed

Delay variation is caused by variation in access delay In Ethernet-style networks there is no upper bound; in token passing networks upper bound may be set.

Circuit-based networks do not have much of jitter (typically few nanoseconds or microseconds).

If link is overlay network, the delay variation may be much larger because of buffering in intermediate nodes (as in routers).

Error rate on links is caused by bit errors

optical fibres 10^{-9}–10^{-12}

satellite links 10^{-6}–10^{-8}

Overlay networks may also loss data if network nodes gets congested.

Other network components

- Store-and-forward application relay
  - email servers: Message/mail transfer agents

  Store-and-forward application relay differs from a router in sense that s-a-f takes responsibility to store the message until it is delivered to next s-a-f or recipient.
- Multicasting component
  - multicast router
  - document replicator
  - mail distributor: for example mailing list software
- Application translator
  - protocol translators: email, wap gateways
  - media transcoders
- Mirror and cache
  - server replicators
    * loose: Usenet news
    * tight: server mirrors (for geographical distribution) For example Akamai `http://www.akamai.com/` provides service to replicate media content (pictures, video, audio) to multiple servers. The user is directed to closest server (think to be).
  - document or media caches: for example web cache

  The difference between mirror and cache is that mirror is replicated at time intervals (daily or so) while document is stored info cache when it is requested.
- Network services
  - name service: DNS in Internet
  - directory service such as LDAP

Application traffic characteristics

Peak rate is highest rate source will transmit data

Burst size is maximum amount of data source transmits at peak rate

Average rate calculated over time interval. Time interval may be order of few seconds or whole lifetime of connection

Connection hold time is time application is active

Application traffic requirements

Minimum bandwidth needed to maintain application fidelity

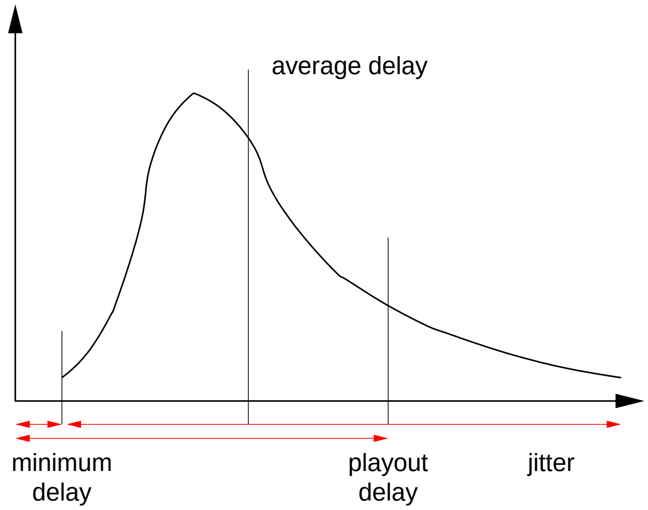
Delay application may tolerate. Transmission delay and queueing delay are part of bandwidth requirement.

Delay variation is difference between maximum and minimum delay

Loss rate allowed without fidelity degradation

Delay distribution

- Packets that arrive before playout delay ar ok
- Late packets are “lost”



Different sources and receivers

- Source

  interacting: maximum delay may not exceed one second, preferable round-trip delay less than 0.5 s

  live: greater delay allowed but delay variation should not exceed playout buffer

  stored: available bandwidth and delay may vary much because data can be sent well in advance. One must take account possible controls and handle those locally.

- Receiver

  human: sensitive for delay, not much sensitive for transitory errors
  - eye integrates information
  - less tolerance for audio errors

  recorder: not sensitive for delay, but prefers error-free transmission

Audio streams

Telephone quality 300–3400 Hz

- uncompressed 64 kbit/s (BER $< 10^{-2}$)
- compressed 4–16 kbit/s

CD quality 0–20 kHz

- uncompressed 1.4 Mbit/s
- compressed 128–384 kbit/s (BER $< 10^{-4}$)

- Echo cancelling needed if one-way delay over 24 ms
- Delay for interactive should be less than 100 ms to provide delay offset of 500 ms
- Higher requirements for virtual reality
- Intermedia synchronisation better than 100 ms

Video streams

HDTV (1920x1080x60)

- uncompressed 2 Gbit/s
- compressed 20–34 Mbit/s

Studio/Broadcast (625x864x25)

- uncompressed 166 Mbit/s
- compressed 3–6 Mbit/s

VCR (625x864x25)

- compressed 1,2 Mbit/s

Video conference (352x288x5)

- compressed 112 kbit/s

- Delay requirements set by audio
- BER 10^{-4}–10^{-6}
- Hierarchical coding for multicast delivery
- Compression adds burstiness up to ratio 10:1[4]

Other application

- ”Non-real-time” *do* have requirements
- Web surfing: a page should load in less than 1 second and it must not take more than 10 seconds, otherwise user may get impatient and aborts transfer. [2, p. 135]
- Data transfer applications require minimum bandwidth

[4] This depends on MPEG coding how group of pictures (GOP) defined i.e. how many P and B frames there are for a I frame.

TCP

- Dominant network transport protocol
- Has two rate limiting mechanisms
  - flow control: not overrun receiver
  - congestion control: not overload network
- TCP congestion control reacts on packet losses
  ⇒ network must signal coming congestion by dropping packet
  ⇒ delay of RTT (round trip time) in feedback
- Round-trip time estimation essential

Estimating round trip delay

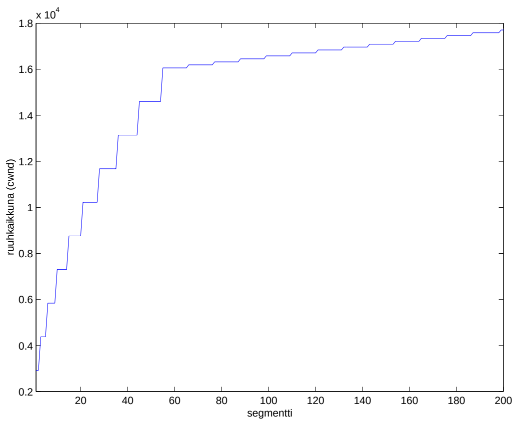
- *Too short* results superfluous retransmissions
- *Too long* reduces network utilisation and throughput
- Jacobson/Karels [1]

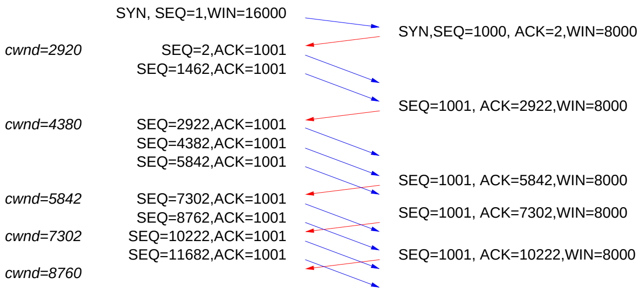
$$\begin{aligned} Diff &= RTT_{Sample} - RTT_{Est} \\ RTT_{Est} &= RTT_{Est} + \delta Diff \\ Dev &= Dev + \delta(Diff - Dev) \\ RTO &= RTT_{Est} + \phi Dev \\ & 0 \le \delta \le 1 \\ & \phi = 4 \end{aligned}$$

- Many "better" estimators proposed

Estimating network capacity

- Initially no knowledge of network status
  ⇒ $cwnd \le \min\{2SMSS, 2segment\}$
- Try to use as much network as possible
  ⇒ $cwnd = cwnd + SMSS$ by each acknowledgement
- After a point ($cwnd > ssthresh$) limit rate of increase
  ⇒ $cwnd = cwnd + SMSS^2/cwnd$ by each acknowledgement
- $ssthresh$ is threshold value between "slow start" and "congestion avoidance"





Are there other ways to signal congestion

- Packet drop is a coarce way to indicate
- ICMP Source Quench[3, p. 10]
  - adds more traffic to network
    ⇒ may create problems
- Early Congestion Notification[4]
  - experimental method to inform congestion
  - packet is not dropped but marked
  - receiver informs sender
  - sender limits its rate
- Delays increase
  - queues build up in congestion
    ⇒ queue delays are longer

Can applications adapt (or should)

- Bandwidth less than expected
  - lower media speed using for example hierarchical coding
  - drop some media
- Losses
  - lower media speed to ease congestion
  - prepare for losses
    * FEC: Forward error correction
    * redundant information send copy of data with lower fidelity in next packet
- Delay variation too large
  - increase playout delay
    ⇒ total delay increases
  - prepare to “loss”[5] some of packets

Summary

- Each network component has it properties
  - throughput
  - delay with variation
  - error rate

[5] A packet, which is late, can be handled as packet which is lost

- Componets are
  - end systems
  - routers
  - links
- *Every* application has requirements

References

[1] V. Jacobson. Congestion avoidance and control. In *Proceedings of the ACM SIGCOMM Conference*, pages 314–329, August 1988.

[2] Jakob Nielsen. *Usability Engineering*. AP Professional, Boston, 1993.

[3] J. Postel. Internet Control Message Protocol. Request for Comments RFC 792, Internet Engineering Task Force, September 1981. (Internet Standard) (Updated by RFC0950) (Obsoletes RFC0777) (Also STD0005). URL:`http://www.ietf.org/rfc/rfc792.txt`.

[4] K. Ramakrishnan and S. Floyd. A Proposal to add Explicit Congestion Notification (ECN) to IP. Request for Comments RFC 2481, Internet Engineering Task Force, January 1999. (Experimental). URL:`http://www.ietf.org/rfc/rfc2481.txt`.