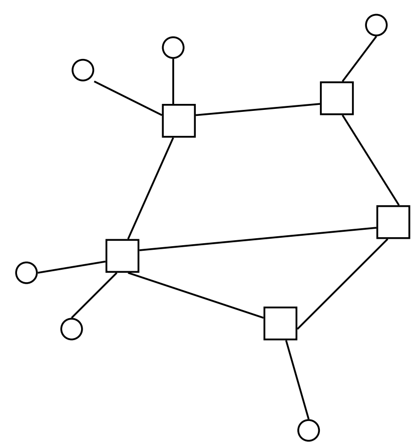
2. Modelling of telecommunication systems (part 1)

Contents

- Telecommunication networks
- Network level: switching and routing
- Link level: multiplexing and concentration
- Shared media: multiple access

Telecommunication network

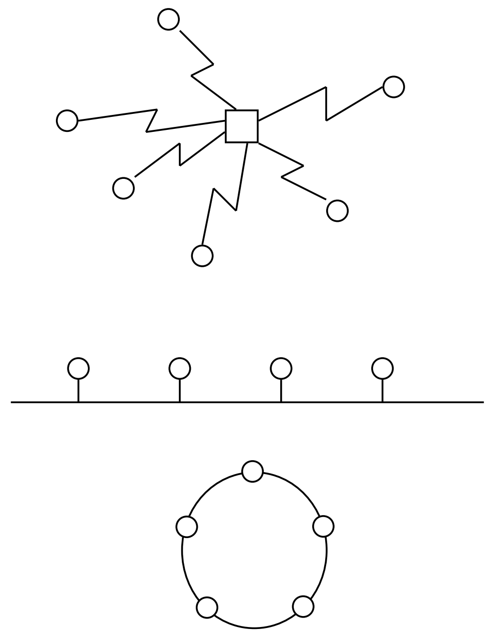
- A simple model of a telecommunication network consists of
  - **nodes**
    - terminals
    - network nodes
  - **links** between nodes
- **Access network**
  - connects the terminals to the network nodes
- **Trunk network**
  - connects the network nodes to each other



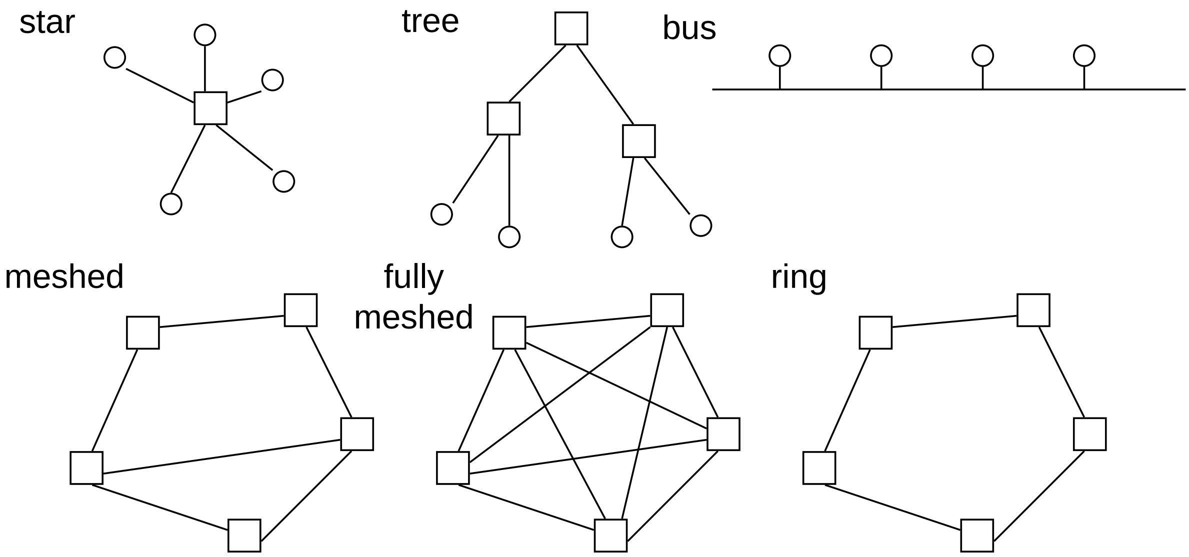
Shared medium as an access network

- In the previous model,
  - connections between terminals and network nodes are **point-to-point** type (=> no resource sharing within the access netw.)
- In some cases, such as
  - mobile telephone network
  - local area network (LAN) connecting computers

  the access network consists of **shared medium**:
  - users have to **compete** for the resources of this shared medium
  - **multiple access** (MA) techniques are needed



Network topologies



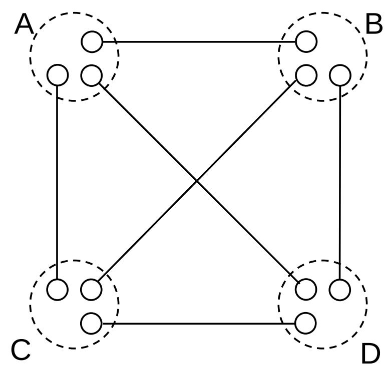
Network hierarchy

- Networks typically constructed level-by-level
- Relations to network topology
  - flat topologies (topology within one level)
  - hierarchial topologies
- One natural hierarchy:
  - access vs. trunk network
- Traditionally:
  - many hierarchial levels (5 in AT&T)
- Current tendency:
  - to reduce the number of levels in hierarchy
  - “We see future large national networks with only three levels.”

Example: Why networks? (1)

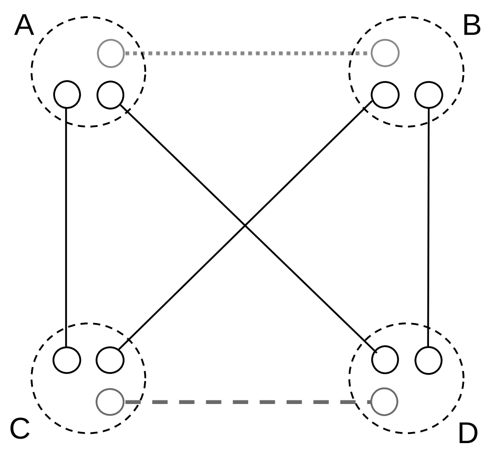
- Assume that
  - there are N = 100 persons who want to be connected with each other
- Solution 1: **Separate networks**
  - point-to-point links with separate terminals for each user pair
  - no switches
  - no resource sharing
  - very low utilization

Resource	Number	Max. Util.
○	N(N-1)	1%
—	N(N-1)/2	1%
□	0	--



Example: Why networks? (2)

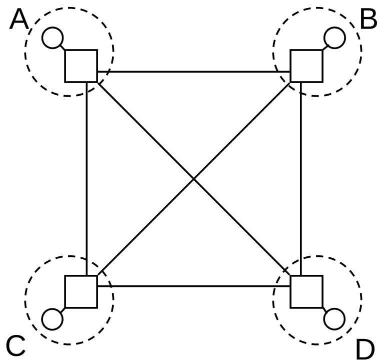
- Simultaneous connections A-B and C-D in solution 1



Example: Why networks? (3)

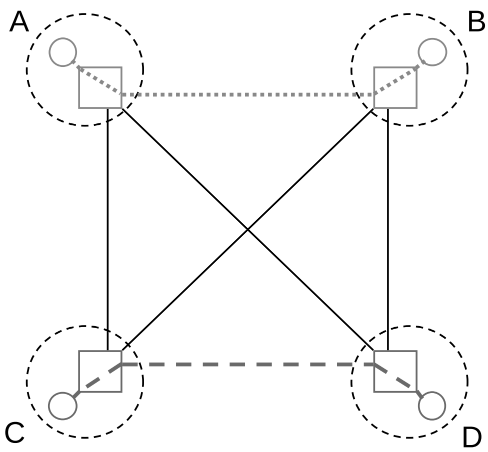
- Still assume that
  - there are N = 100 persons who want to be connected with each other
- Solution 2:
  Fully meshed network
  - one terminal per user
  - one 1x(N-1) switch per user
  - point-to-point links for each switch pair
  - partial resource sharing
  - better utilization

Resource	Number	Max. Util.
○	N	100%
———	N(N-1)/2	1%
□	N	100%



Example: Why networks? (4)

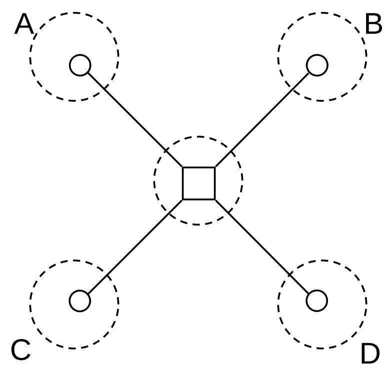
- Simultaneous connections A-B and C-D in solution 2



Example: Why networks? (5)

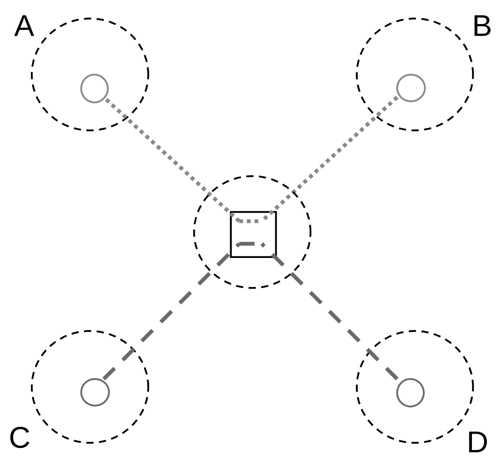
- Still assume that
  - there are N = 100 persons who want to be connected with each other
- Solution 3: **Star network**
  - one terminal per user
  - just one NxN switch for the whole network
  - NxN switch can be implemented with N/2 parallel Nx1/1xN switch pairs
  - complete resource sharing
  - best utilization

Resource	Number	Max. Util.
○	N	100%
——	N	100%
□	1	100%



Example: Why networks? (6)

- Simultaneous connections A-B and C-D in solution 3



Contents

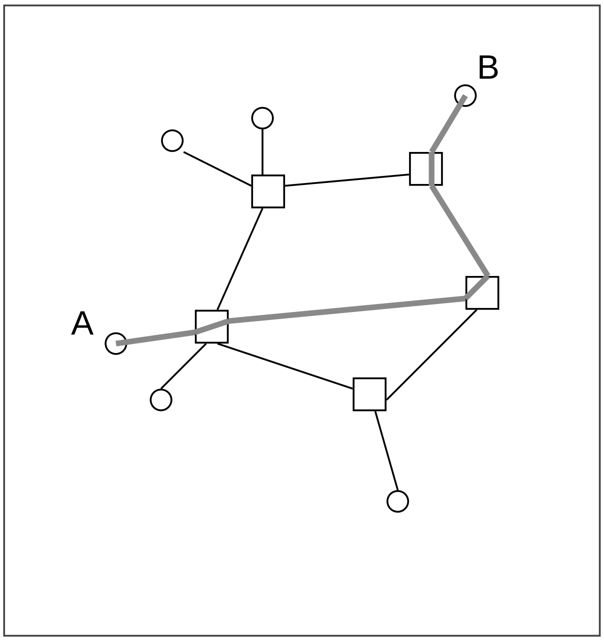
- Telecommunication networks
- Network level: switching and routing
- Link level: multiplexing and concentration
- Shared media: multiple access

Switching modes

- **Circuit switching**
  - telephone networks
  - mobile telephone networks
  - even applied to data networks
- **Packet switching**
  - data networks
  - two possibilities
    - **connection oriented**: e.g. X.25, Frame Relay
    - **connectionless**: e.g. Internet (IP), SS7 (MTP)
- **Cell switching**
  - fast packet switching with fixed length packets (cells): ATM
  - integration of different traffic types (voice, data, video)
    $\Rightarrow$ multiservice networks

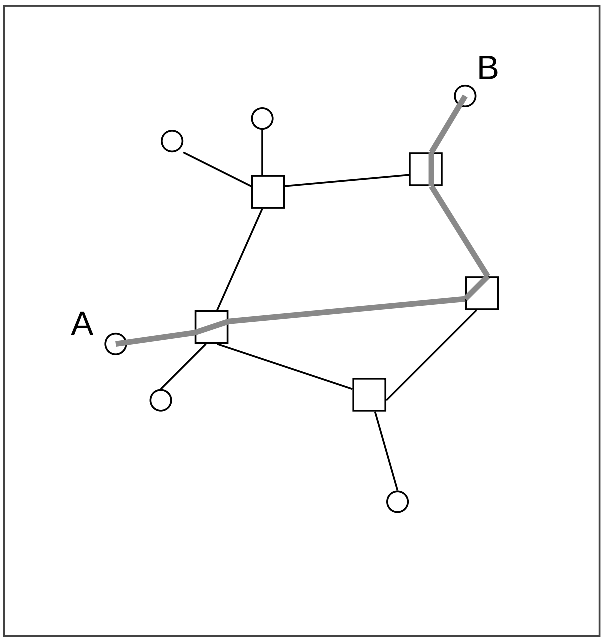
Circuit switching (1)

- **Connection oriented**:
  - connections **set up** end-to-end before information transfer
  - resources **reserved** for the whole duration of connection
- Information transfer as **continuous stream**



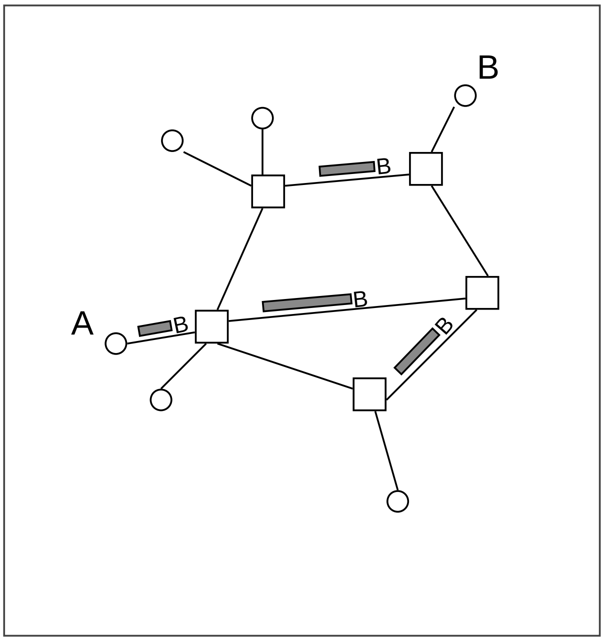
Circuit switching (2)

- Before information transfer
  - delay (to set up the connection)
- During information transfer
  - no overhead
  - no extra delays



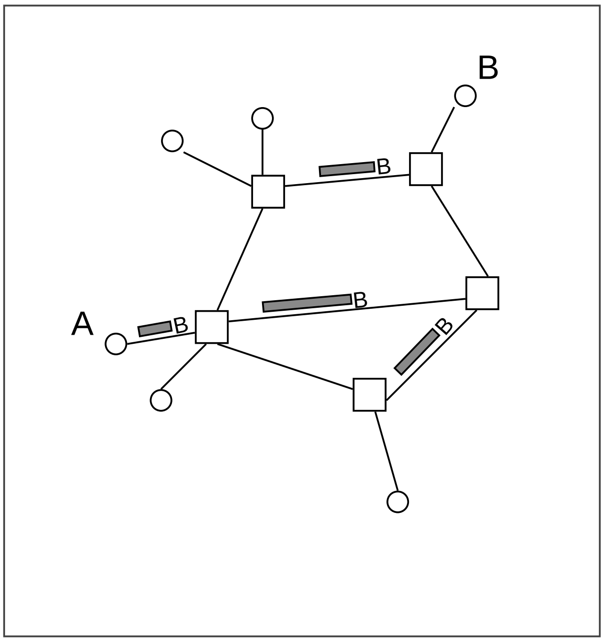
Connectionless packet switching (1)

- **Connectionless**:
  - no connection set-up
  - no resource reservation
- Information transfer as **discrete packets**
  - varying length
  - global address (of the destination)



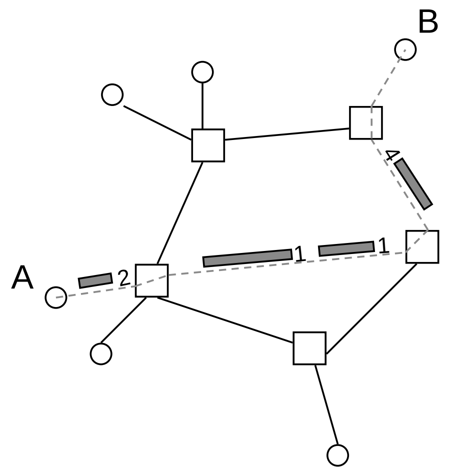
Connectionless packet switching (2)

- Before information transfer
  - no delays
- During information transfer
  - overhead (header bytes)
  - packet processing delays
  - queueing delays (since packets compete for joint resources)



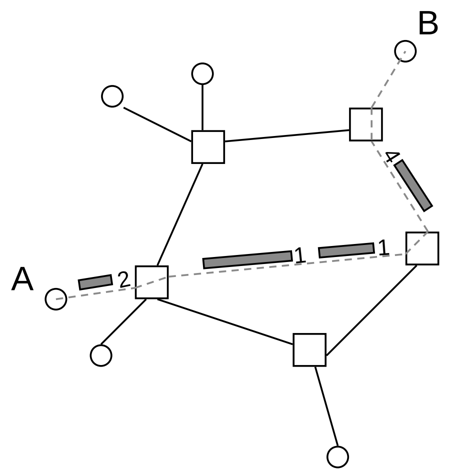
Connection oriented packet switching (1)

- **Connection oriented**:
  - virtual connections **set up** end-to-end before information transfer
  - no resource reservation
- Information transfer as **discrete packets**
  - varying length
  - local address (logical channel index)



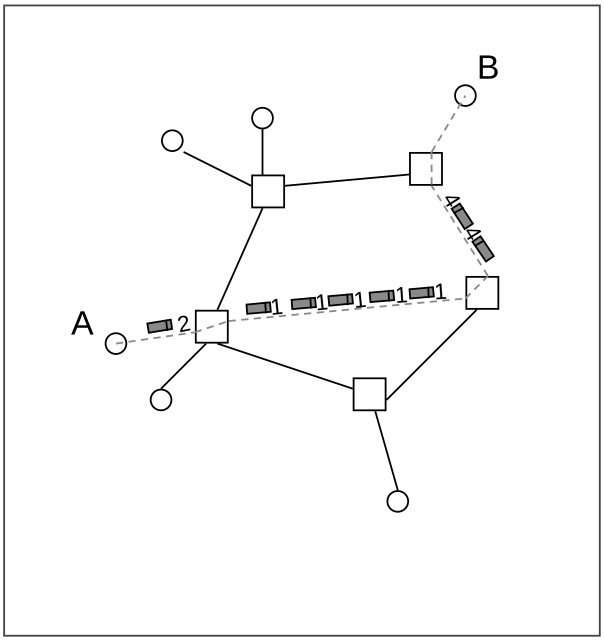
Connection oriented packet switching (2)

- Before information transfer
  - delay (to set up the virtual connection)
- During information transfer
  - overhead (however, less than in connectionless mode)
  - packet processing delays (less, due to the shorter address)
  - queueing delays (since packets compete for joint resources)



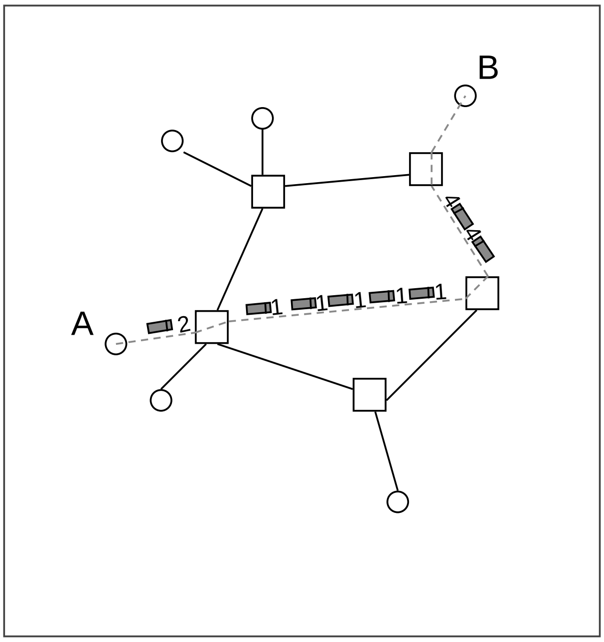
Cell switching (1)

- **Connection oriented**:
  - virtual connections **set up** end-to-end before information transfer
  - resource reservation possible but not mandatory
- Information transfer as **discrete packets** (cells)
  - fixed (small) length
  - local address



Cell switching (2)

- Before information transfer
  - delay (to set up the connection)
- During information transfer
  - overhead (per packet even more than in connectionless mode)
  - packet processing delays (less, due to the shorter address and the fixed length of cells)
  - queueing delays (if resources not reserved beforehand)



Switching modes: summary

- Circuit switching
  - suitable for real-time traffic (voice, RT-video, …)
  - inefficient for VBR traffic and data
  - transparent but inflexible
- Cell switching
  - quite flexible
  - efficient use of network resources
  - seq. integrity guaranteed
  - real-time guarantees possible
  - possible to integrate different traffic types
- Connection oriented packet switching
  - quite flexible
  - efficient use of network resources
  - seq. integrity guaranteed
  - no real-time guarantees
- Connectionless packet switching
  - flexible and fault tolerant
  - efficient use of network resources
  - seq. integrity not guaranteed
  - no real-time guarantees

Routing methods in telephone networks

- **Fixed**
  - no alternative routes
- **Hierarchic**
  - traditionally used in telephone networks
  - alternative routes searched in a fixed order
- **Dynamic (non-hierarchic)**
  - time-dependent (dynamic) routing tables
  - AT&T has shown some 15% improvement with routing performance with DNHR (dynamic non-hierarchial routing)
- **Adaptive (non-hierarchic)**
  - state-dependent (adaptive) call-by-call routing decisions
  - AT&T’s next generation: RTNR (real-time network routing)
  - Another adaptive method: DAR (dynamic alternative routing) by BT

Contents

- Telecommunication networks
- Network level: switching and routing
- Link level: multiplexing and concentration
- Shared media: multiple access

Analogue vs. digital systems (1)

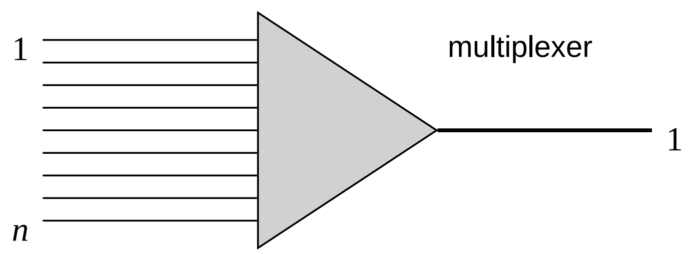
- Originally telecommunication networks (i.e. telephone networks) were purely analogical
  - First: digital trunks between exchanges
  - Then: digital exchanges
  - In the current telephone network, the telephone itself and the access line are still based on the analogue technology
  - ISDN and GSM are the first completely digital telephone networks (including the terminals and the access part)
- Packet switched networks have always been completely digital
  - e.g. LANs
- Cell switched networks (ATM) are also completely digital

Analogue vs. digital systems (2)

- Analogical circuit switched system:
  - one connection occupies a single one or a multiple of channels
  - link capacity expressed in number of channels
- Digital circuit switched system:
  - one connection occupies a single one or a multiple of channels
  - channel capacity expressed in bits per second (bps, kbps, Mbps, ...)
    - typically: 64 kbps
  - link capacity expressed either in number of channels or in bits per second (being then a multiple of the channel capacity)
- Digital packet/cell switched system:
  - link capacity occupied dynamically on-demand
  - capacity demand (of a connection) expressed in bits per second
  - link capacity expressed in bits per second

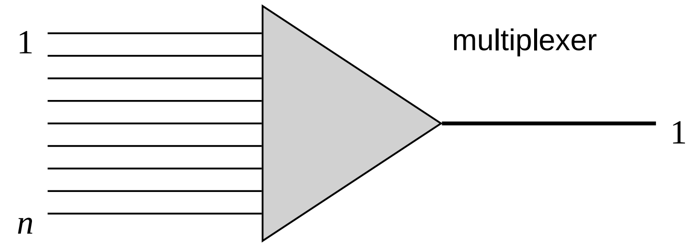
Transmission: multiplexing (1)

- Originally,
  - each connection in a telephone network required its own physical link
- By multiplexing,
  - the capacity of a single physical link is divided into multiple channels
  - each connection typically occupies one channel
  - thus, multiple connections can be conveyed by a single link
- The device implementing this is called a **multiplexer**



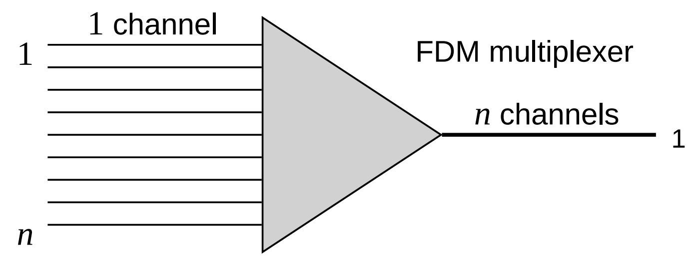
Transmission: multiplexing (2)

- In circuit switched networks, there are two different multiplexing techniques:
  - frequency division multiplexing (FDM)
  - time division multiplexing (TDM)
- In packet switched networks, there is just
  - statistical multiplexing



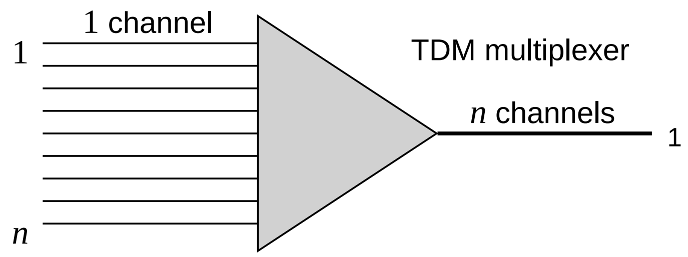
Frequency division multiplexing (FDM)

- FDM
  - oldest multiplexing technique
  - used in analogue circuit switched systems
  - fixed portion (frequency band) of the link bandwidth reserved for each channel
- FDM multiplexer is lossless
  - input: n 1-channel physical connections
  - output: 1 n-channel physical connection



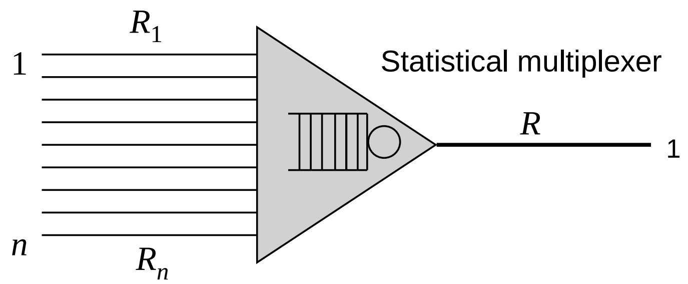
Time division multiplexing (TDM)

- TDM
  - used in digital circuit switched systems
  - information conveyed on a link transferred in **frames** of fixed length
  - fixed portion (time slot) of each frame reserved for each channel
- TDM multiplexer is lossless
  - input: n 1-channel physical connections
  - output: 1 n-channel physical connection



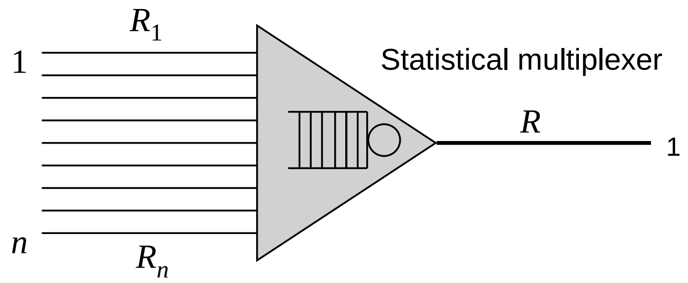
Statistical multiplexing

- Statistical multiplexing
  - used in digital packet/cell switched systems (e.g. Internet, ATM, ...)
  - information transferred in packets of varying or fixed length
  - each packet belongs to exactly one connection
    - packet header includes the connection identifier
  - link capacity reserved dynamically and asynchronously as packets arrive $\Rightarrow$ need for buffering



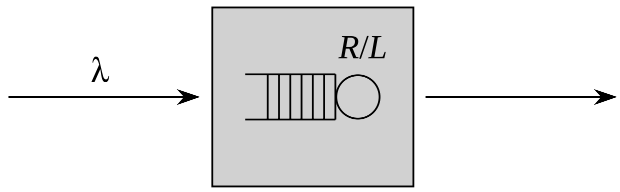
Statistical multiplexer

- Statistical multiplexer is (typically) lossy
  - input: n physical connections with link speeds R_i $(i = 1,\ldots,n)$
  - output: 1 physical connection with link speed $R \le R_1 + \ldots + R_n$
- However, the loss probability can be decreased by enlarging the buffer
  - with an "infinite" buffer, it is enough that R exceeds the average aggregated input rate



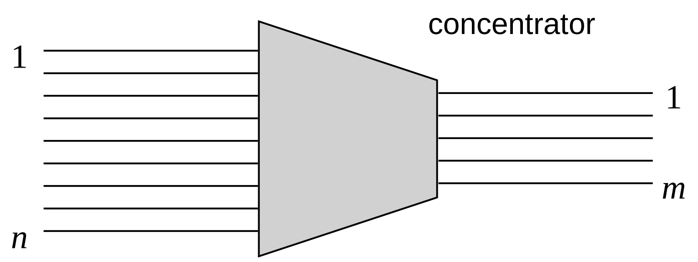
Teletraffic model for a statistical multiplexer

- Multiplexer can be modelled as
  - a pure waiting system (as below) if the buffer is large
  - a mixed system if the buffer is small
- Traffic consists of packets
  - each packet is transmitted with the full link speed R
  - let L denote the average packet length
  - packet service rate μ will be $\mu = R/L$
  - stability requirement: packet arrival rate $\lambda < \mu$



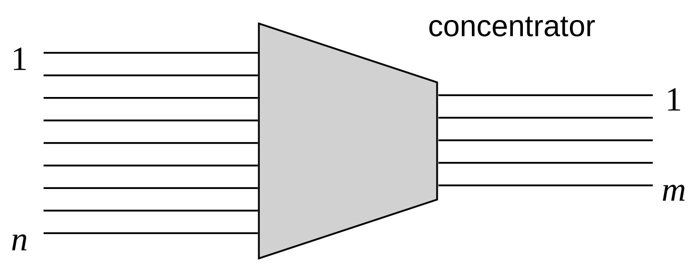
Transmission: concentration

- Concentration
  - used in circuit switched systems (analogue/digital)
  - typically in the access network part for economical reasons
  - however, switches are also (implicitly) concentrators
- In concentration,
  - traffic (= connections) from n 1-channel links is concentrated on a single m-channel link, where $m < n$
  - Idea: the probability that all n connections are active is typically very small



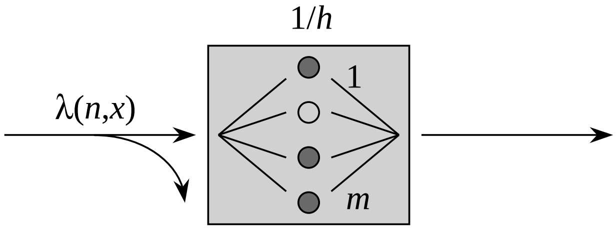
Concentrator

- Concentrator is lossy
  - input: n 1-channel physical connections
  - output: 1 m-channel physical connection with $m < n$
- Outgoing link should be dimensioned (i.e. m should be chosen) so that
  - the call blocking probability (that all m channels are occupied when a new call arrives) is small enough
  - In other words: the quality of service requirement should be fulfilled



Teletraffic model for a concentrator

- Concentrator can be modelled as
  - a pure loss system (as below) with m parallel servers
- Traffic consists of connections
  - traffic generated by a **finite** number (n) of sources
    $\Rightarrow$ arrival rate λ is not constant but depends on the total number of sources (n) and the number of active sources (x): $\lambda = \lambda(n,x)$
  - let h denote the average connection holding time
  - service rate μ will be $\mu = 1/h$



Contents

Multiple access techniques used in mobile telephone networks

- Mobile telephone networks are geographically divided into cells
  - Each cell has its own base station
- The radio frequency band reserved for the network access (within a cell) is divided into channels
  - The users (located in that cell) compete for these channels
- Dynamic channel assignment is made
  - **centralized** by the base station
- Various multiple access methods:
  - **frequency division multiple access** (FDMA)
  - **time division multiple access** (TDMA)
  - **code division multiple access** (CDMA)

FDMA and TDMA

- Frequency division multiple access (FDMA)
  - used in analogue mobile networks, e.g. NMT
  - radio frequency band reserved for the network divided into subbands (channels)
  - each connection occupies one channel
  - thus, simultaneous connections use separate frequency (sub)bands
- Time division multiple access (TDMA)
  - used in digital mobile networks, e.g. GSM
  - information transferred in frames of fixed length
  - fixed portion (time slot) of each frame reserved for each channel
  - each connection occupies one channel
  - thus, simultaneous connections use the same frequency band
  - utilization of the frequency band better than in FDMA

CDMA

- Code division multiple access (CDMA)
  - used in digital mobile networks, e.g. IS-95 (USA)
  - information coded before transfer in such a way that simultaneous transmissions do not interfere (too much)
  - each code corresponds to a channel
  - thus, simultaneous connections use the same frequency band
  - in general, the utilization of the frequency band is better than in TDMA
  - however, the notion “system capacity” in CDMA is elastic (contrary to FDMA and TDMA):
    - the more codes (channels), the greater the interference!

Teletraffic modelling of various multiple access techniques

- All multiple access techniques mentioned above (FDMA, TDMA and CDMA) can be modelled as a pure loss system
- Traffic consists of calls
  - calls either **fresh** or **handovers**
  - fresh calls may arrive according to a Poisson process, but is it so with the handovers?
  - due to handovers, call holding time is now different from coversation holding time
  - one more new feature: **mobility modelling**
- System capacity (that is, the number of parallel channels) depends on
  - the width of the frequency band reserved for the network
  - the multiple access technique used
- In CDMA, the system capacity depends additionally on
  - the allowed level of interferation (that is, on the required quality of service)

Multiple access techniques in computer LANs

- A computer LAN (local area network) transfers packets between any stations connected to this LAN
  - The stations compete for this joint transmission medium
- Dynamic channel assignment is made
  - in a fully **distributed** manner by the stations themselves
- Various multiple access methods:
  - **Random Access**
    - ALOHA, Slotted ALOHA (originally in satellite links)
  - **Carrier Sense Multiple Access / Collision Detection** (CSMA/CD)
    - Ethernet, IEEE 802.3
  - **Token Bus**
    - IEEE 802.4
  - **Token Ring**
    - IEEE 802.5

Random Access

- Stations transmit packets totally independently of each other as soon as new packets arrive
  - no prior actions to avoid collisions
  - theoretical maximum for the throughput is less than 20% of the LAN speed
- Assuming a fixed packet length,
  - a slotted system (slot = transmission time of a packet) doubles the theoretical maximum throughput

Analysis of ALOHA (1)

- (1) Assume that
  - the stations generate fixed length packets according to a Poisson process with intensity ν
  - the packets are retransmitted until they reach without any collisions the destination
- Let T denote
  - the time needed to transmit a packet (stability requirement: $\nu < 1/T$)
- Two packets **collide** with each other
  - if and only if their interarrival time is $< T$
  - Collided packets are retransmitted after a random interval
- (2) Assume further that
  - the aggregate packet stream (including all the transmitted and retransmitted packets) still obeys a Poisson process (which is certainly not true) but with a higher intensity λ such that $\lambda < 1/T$

Analysis of ALOHA (2)

- Consider a station where a new packet arrives at time 0
  - No collisions during the transmission $(0,+T)$ time if and only if no other packet arrivals (to any station) between time interval $(-T,+T)$
  - Due to assumption (2), this happens with probability $\exp(-2\lambda T)$
- Thus,
  - the throughput ν is $\nu = \lambda \cdot \exp(-2\lambda T)$
- This is maximized by $\lambda_{max} = 1/(2T)$ corresponding to
  - an offered load of $\lambda_{max} T = 1/2 = 50\%$
- The maximum throughput ν_{max} is
  - $\nu_{max} = \lambda_{max} \cdot \exp(-2\lambda_{max} T) = 1/(2eT) \approx 0.184/T \approx 20\%\ (1/T)$

Analysis of Slotted ALOHA

- (3) Assume then that
  - the packets are transmitted in slots of length T
- In this system, two packets collide with each other
  - if and only if they arrived during the same slot
  - Due to assumption (2), no collisons with probability $\exp(-\lambda T)$
- Thus,
  - the throughput ν is $\nu = \lambda \cdot \exp(-\lambda T)$
- This is maximized by $\lambda_{max} = 1/T$ corresponding to
  - an offered load of $\lambda_{max} T = 1 = 100\%$
- The maximum throughput ν_{max} is
  - $\nu_{max} = \lambda_{max} \cdot \exp(-\lambda_{max} T) = 1/(eT) \approx 0.368/T \approx 40\% \ (1/T)$
- Note that this is **double** to that of a pure ALOHA system

THE END