Cloud and Datacenter Networking

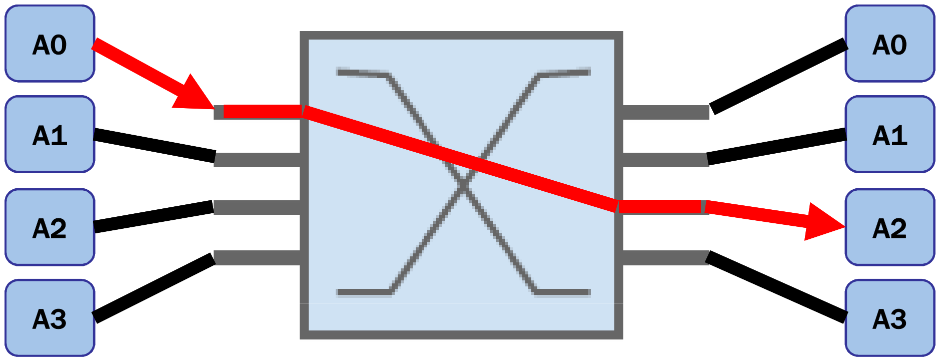
Università degli Studi di Napoli Federico II
Dipartimento di Ingegneria Elettrica e delle Tecnologie dell'Informazione DIETI
Laurea Magistrale in Ingegneria Informatica

Prof. Roberto Canonico

Switching theory basic concepts

Switching system

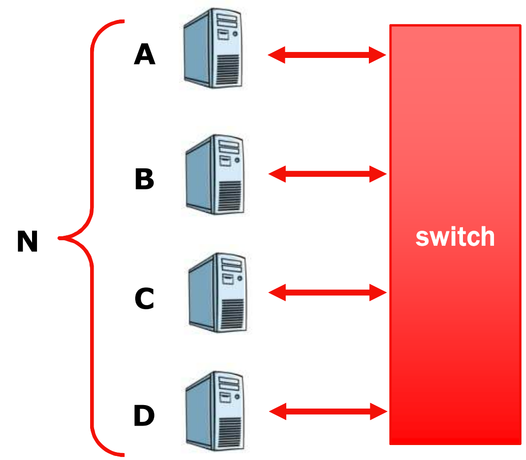
- A switching system connects N terminals (A_0, A_1, ..., A_{n-1}) so that a (unidirectional) transmission is possible from the transmitting side of any given terminal A_i to the receiving side of any other terminal A_j ($i \neq j$)

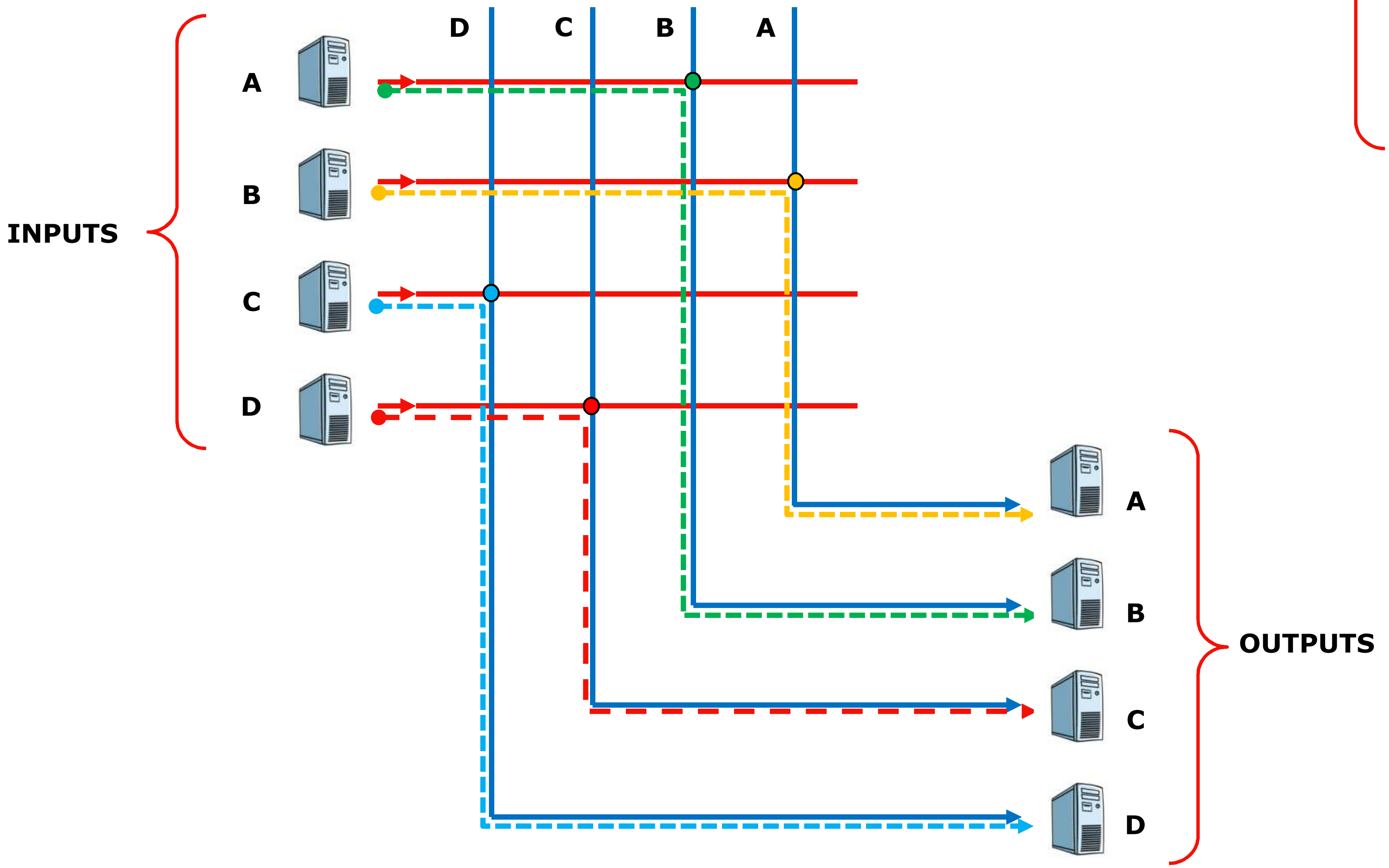


- The public telephone system is a switching system in which connections are established on a per-call basis by means of signalling (*circuit switching*)
  - In the early telephone networks, human operators closed circuits manually, while today this is done automatically by digital switches
- A switched computer LAN is also a switching system in which transmission is segmented in packets bringing in the header the address of destination

Switching system: crossbar matrix

- # Bidirectional connections: N
- # switching elements in the matrix: N^2
- Non blocking





B → A

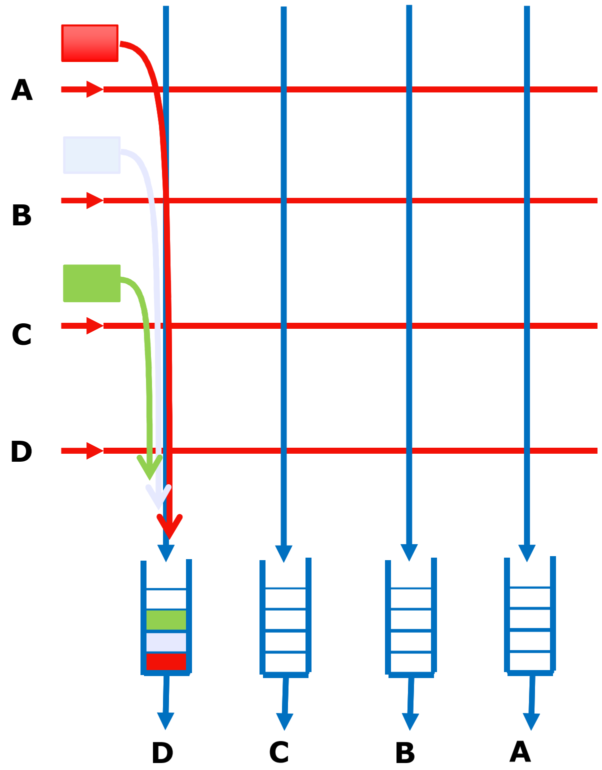
A → B

D → C

C → D

Switching matrix and queues (1)

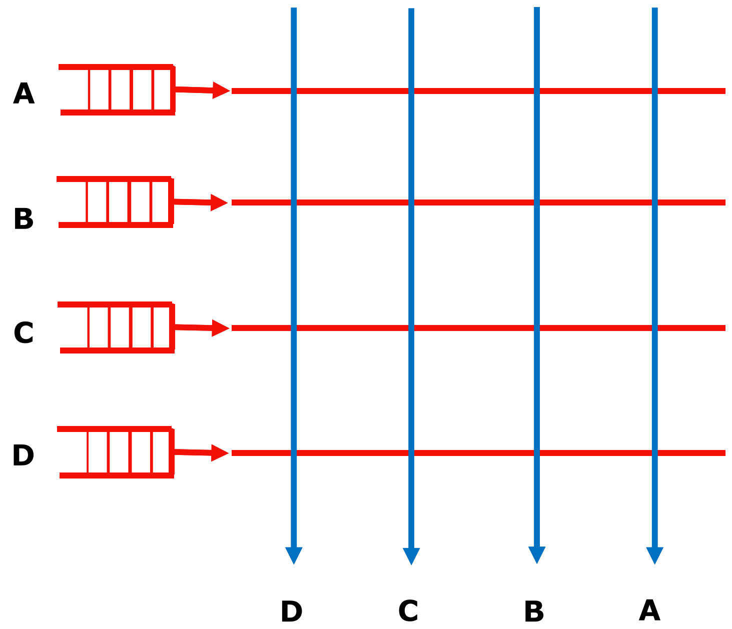
- Let's consider a crossbar switch operating synchronously:
  - Packets enter the switch through input ports at the beginning of each time slot T
  - At each time slot, one packet is transmitted through each output ports
- Queues at input and output lines are necessary to manage the circumstance in which more packets (arriving from different input ports) are to be forwarded through the same output port
- In a switch with output queues <u>only</u>, the switching matrix needs to operate at a rate <u>N times higher than the line transmission rate</u>



Crossbar switch with output queues

Switching matrix and queues (2)

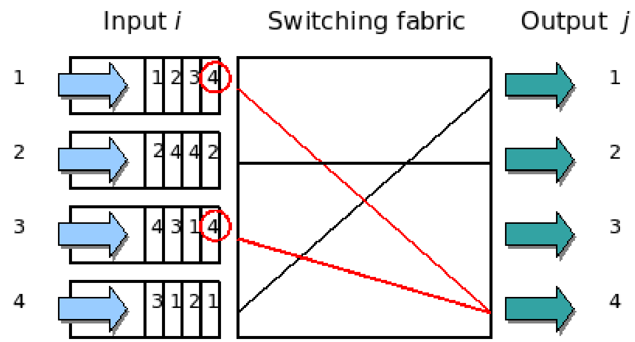
- If the switching matrix operates at the same rate than the line transmission rate
  - The matrix may accept at most one packet per output line at each time slot T
  - In such a case, an input queue is necessary to hold packets that find their output port busy



Crossbar switch with input queues

Head-of-Line blocking in input-queued switches

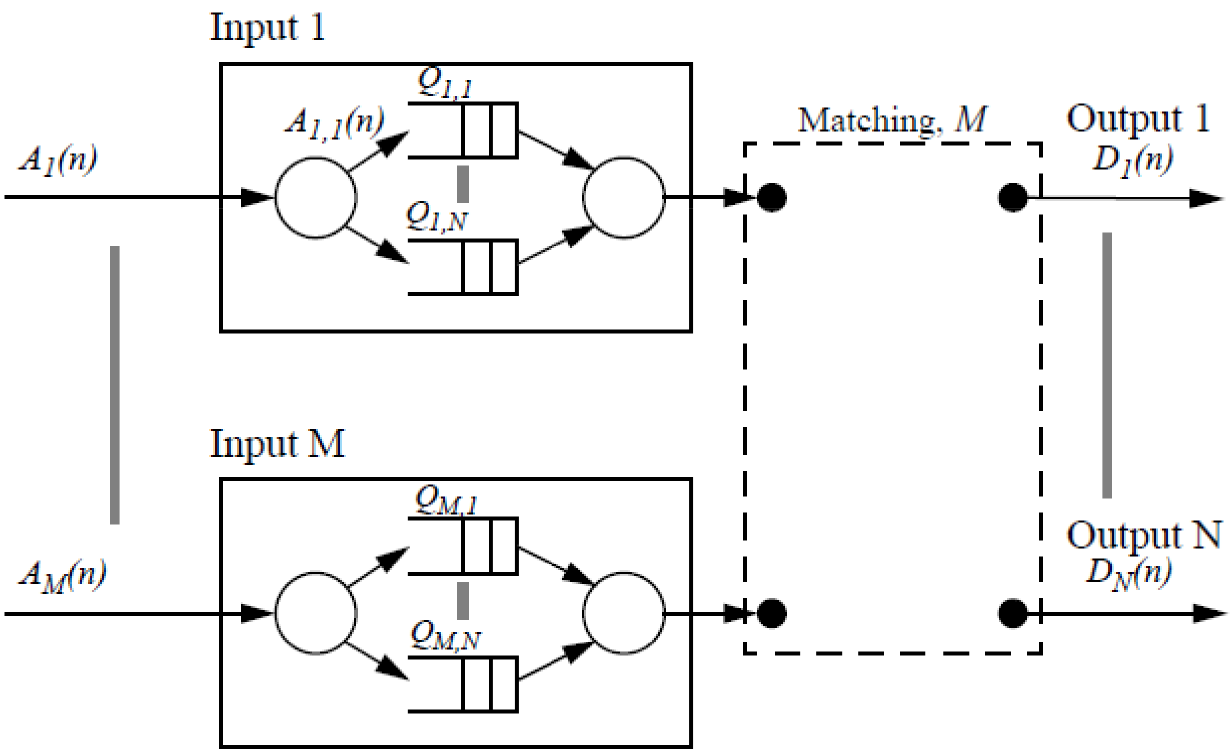
- Let’s consider an input-queued switch
- Head-of-Line blocking problem
  - Example: the 1st and 3rd input flows are competing to send packets to the same output interface
  - The switching fabric decides to transfer the packet from the 3rd input flow
  - The 1st input flow cannot be processed in the same clock cycle and blocks a packet for output interface 3, which is available for processing



- HOL blocking may cause severe performance degradation (58,6% throughput)
- Only switches with input buffering can suffer HOL blocking
- With sufficient internal bandwidth, input buffering is unnecessary; all buffering is handled at outputs and HOL blocking is avoided
  - No-input-buffering is common in small to medium-sized Ethernet switches

Ethernet switch: Virtual output queueing

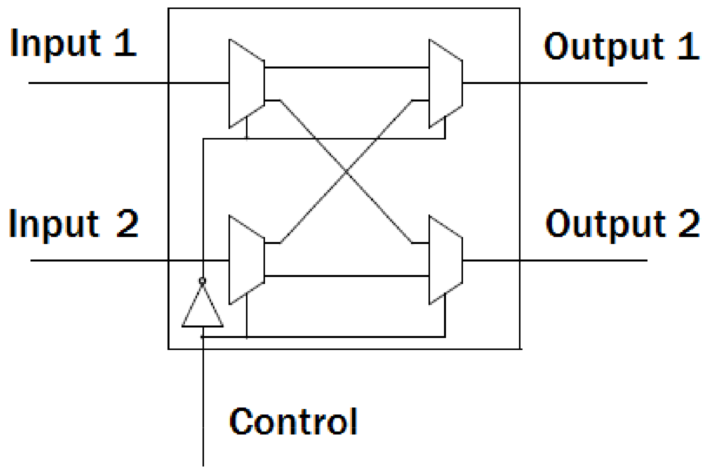
- To prevent the Head-of-Line blocking problem in input-queued switches
- Rather than keeping all traffic in a single queue per input port, separate queues are maintained for each possible output port
    - N^2 queues
- With proper scheduling, 100% throughput may be achieved



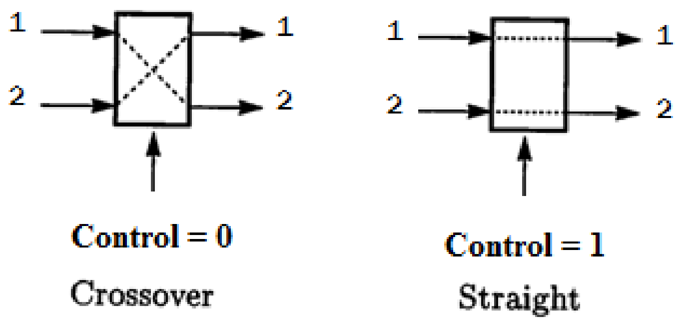
N. McKeown, A. Mekkittikul, V. Anantharam, J. Walrand. "Achieving 100% throughput in an input-queued switch". *IEEE Transactions on Communications*, vol.47, no.8, pp.1260-1267, Aug 1999

Elementary 2x2 switching matrix

- 2 inputs, 2 outputs

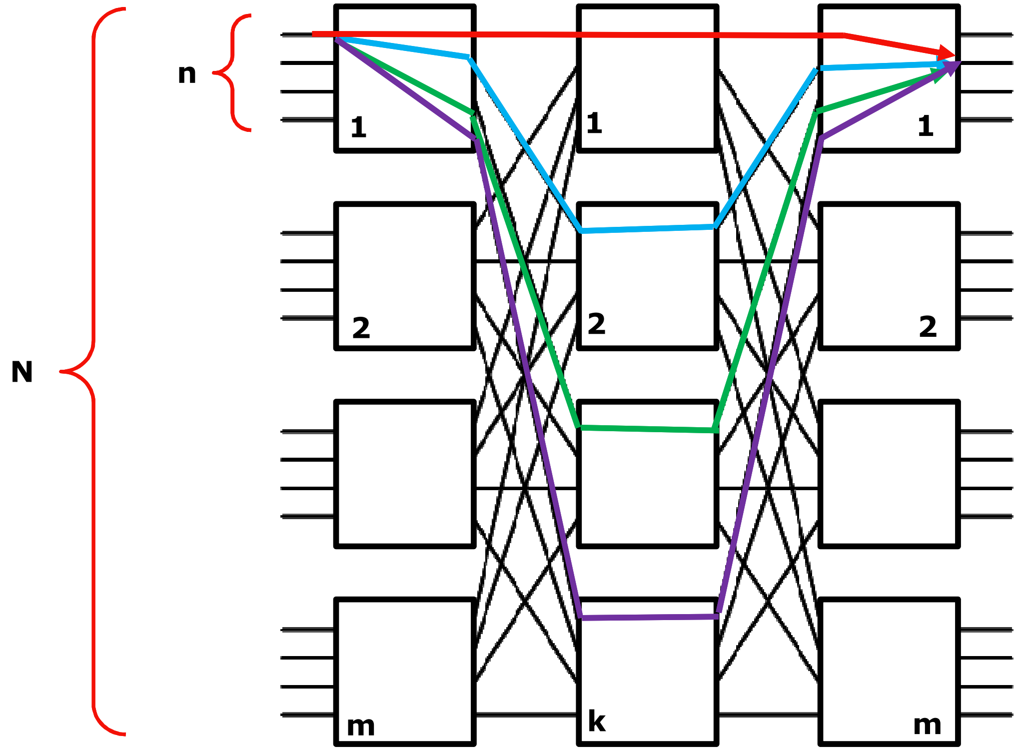


- 2 configurations: straight or crossover



Multi-stage switching networks

- N×N multistage network made of smaller switching matrices organized in multiple layers or *stages*
  - Let us consider a 3 stage network
- Let us decompose N=m·n
- Input stage: m (n×k) switches
- Output stage: m (k×n) switches
- Intermediate stage: k (m×m) switches
- If the number of intermediate switches k is not sufficiently high, a blocking condition may occur



Example:

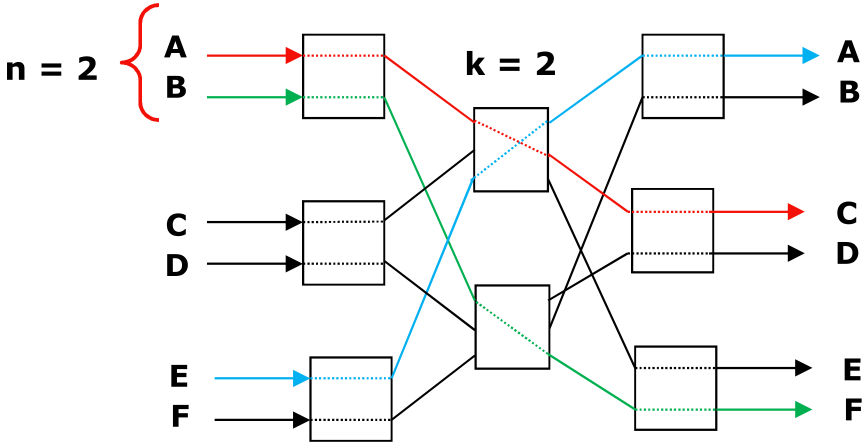
3-stage network made of 12 (4x4) switches replacing a single (16×16) switch

k different paths from each input i to any output j

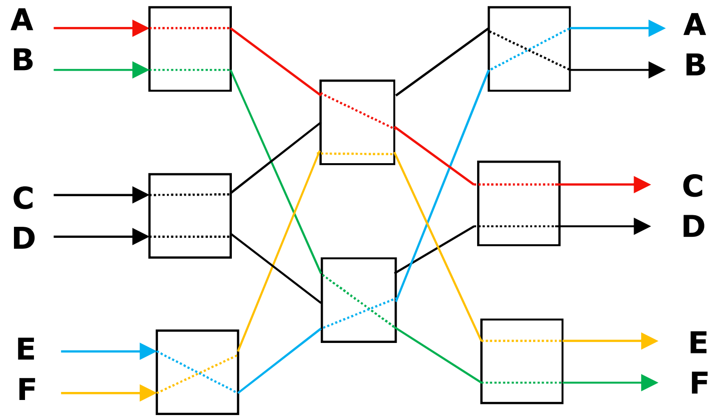
For a given intermediate switch v, only one path exists that connects a given input i to an output j through switch v

3-stage switching networks: Clos theorem

- If the number k of intermediate switches is not sufficiently high, a blocking condition may occur
- Non-blocking condition (re-routing may be nessary):
  - $k \geq n$
- Non-blocking condition without re-routing:
  - $k \geq 2n - 1$ (Clos theorem)



F cannot transmit to E: block

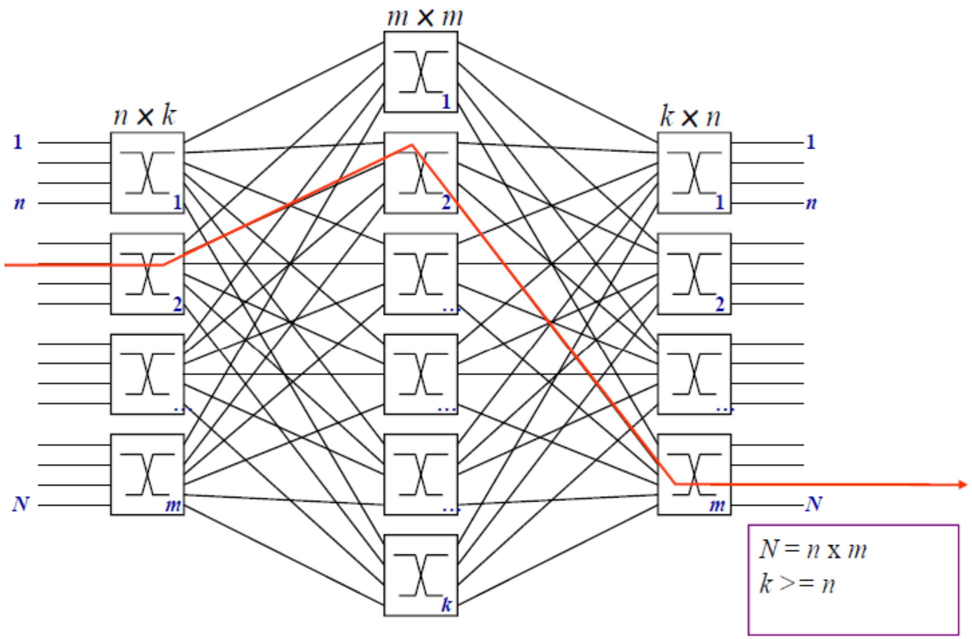


F may transmit to E if E→A transmission is re-routed

Charles Clos. "A Study of Non-blocking Switching Networks". Bell System Technical Journal, March 1953

Clos networks

- 3-stage switching networks
  - Non-blocking if k ≥ 2n-1
  - Block may be removed through re-rerouting if k ≥ n



- Example: N=1000, n=10 → m=100 switches into the first and third stage
  - Second stage with k ≥ 19 (100x100) switches
  - Number of connection points: 1000x1000=1000000 for a single crossbar, 100x(10x19)+19x(100x100)+100x(19x10) = 228000 for a Clos network

Multi-stage switching networks

- Increasing the number of elementary crossbars and the number of stages, it is possible to implement arbitrarily large switching networks