Transmission techniques and multiplexing hierarchies

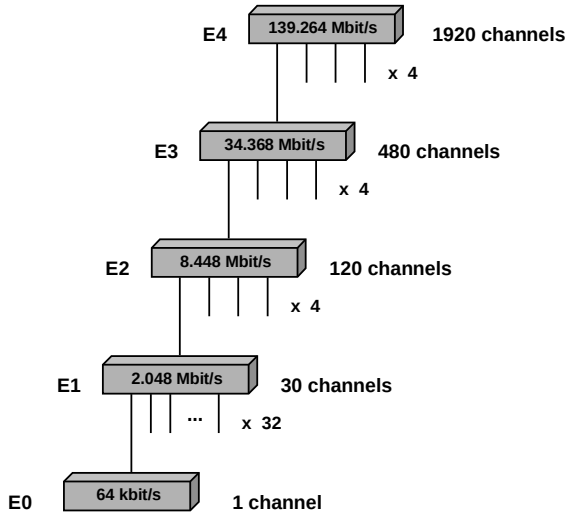
Switching Technology S38.3165
http://www.netlab.hut.fi/opetus/s383165

Transmission techniques

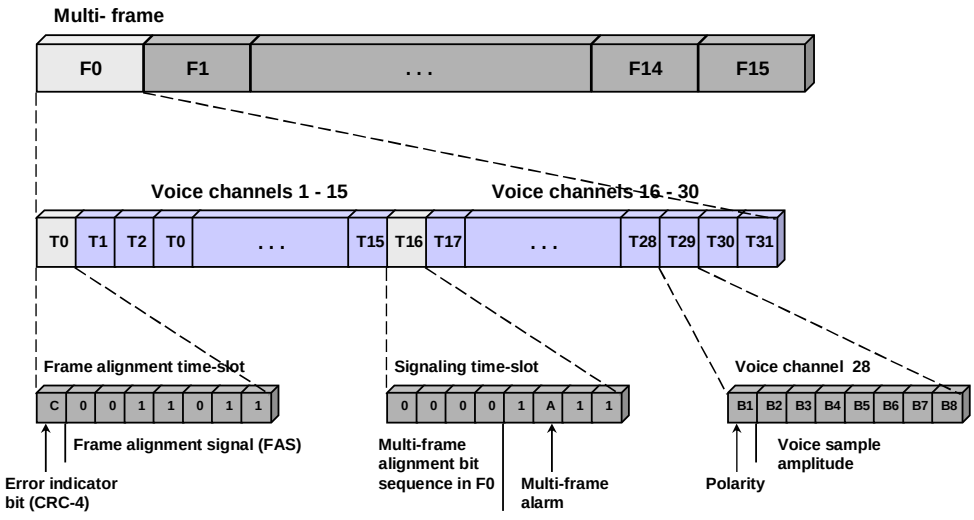
- PDH (Plesiochronous Digital Hierarchy)
- ATM (Asynchronous Transfer Mode)
- IP/Ethernet
- SDH (Synchronous Digital Hierarchy)
- OTN (Optical Transport network)
- GFP (Generic Framing Procedure)

Plesiochronous Digital Hierarchy (PDH)

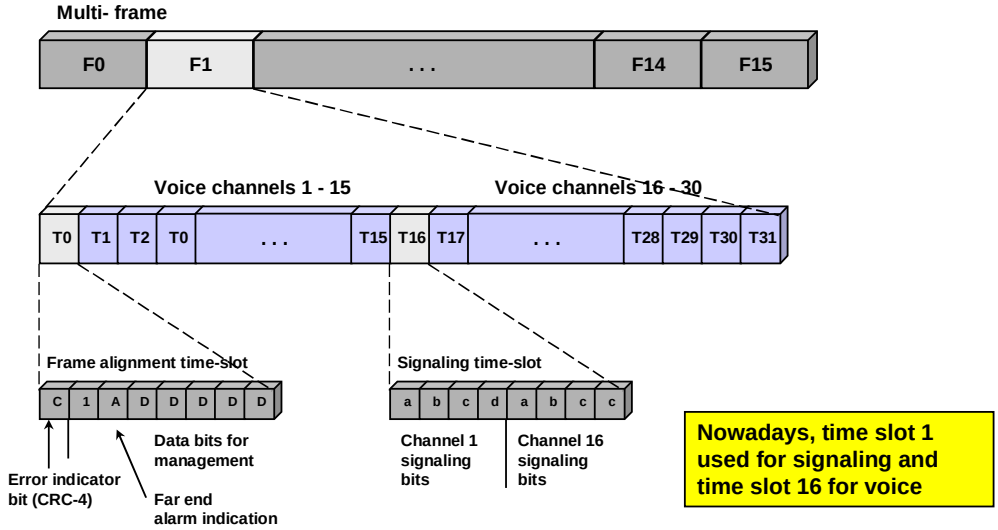
- Transmission technology of the digitized telecom network
- Basic channel capacity 64 kbit/s
- Voice information PCM coded
  - 8 bits per sample
  - A or μ law
  - sample rate 8 kHz (125 μs)
- Channel associated signaling (SS7)
- Higher order frames obtained by multiplexing four lower order frames bit by bit and adding some synchr. and management info
- The most common switching and transmission format in the telecommunication network is PCM 30 (E1)



PDH E1-frame structure (even frames)



PDH E1-frame structure (odd frames)



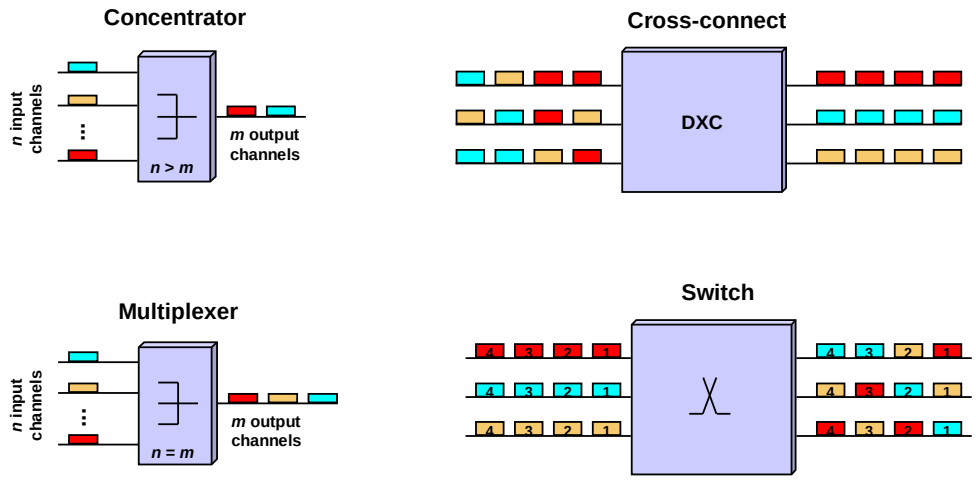
PDH-multiplexing

- Tributaries have the same nominal bit rate, but with a specified, permitted deviation (100 bit/s for 2.048 Mbit/s)
- Plesiochronous = tributaries have almost the same bit rate
- Justification and control bits are used in multiplexed flows
- First order (E1) is octet-interleaved, but higher orders (E2, …) are bit-interleaved

PDH network elements

- **concentrator**
  - *n* channels are multiplexed to a higher capacity link that carries *m* channels (*n* > *m*)
- **multiplexer**
  - *n* channels are multiplexed to a higher capacity link that carries *n* channels
- **cross-connect**
  - static multiplexing/switching of user channels
- **switch**
  - switches incoming TDM/SDM channels to outgoing ones

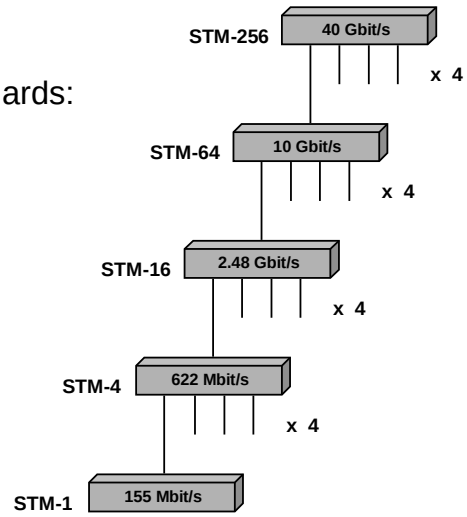
Example PDH network elements



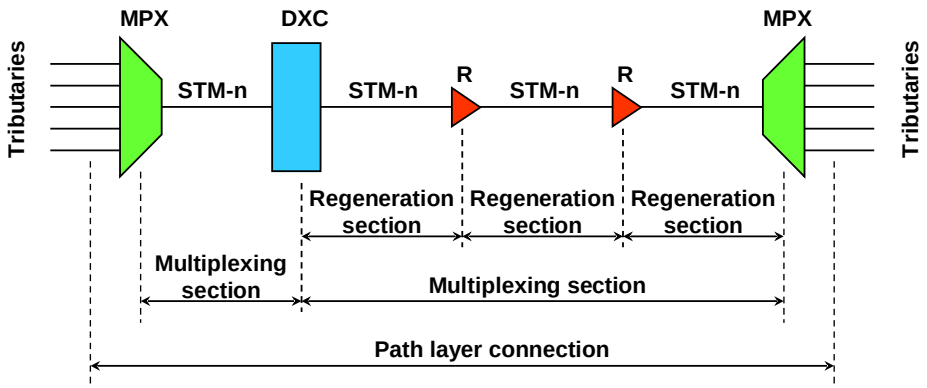
Synchronous digital hierarchy

Major ITU-T SDH standards:
- G.707
- G.783

Notice that each frame transmitted in 125 µs !



SDH reference model



- DXC Digital cross-connect
- MPX Multiplexer
- R Repeater

SDH-multiplexing

- Multiplexing hierarchy for plesiochronous and synchronous tributaries (e.g. E1 and E3)
- Octet-interleaving, no justification bits - tributaries visible and available in the multiplexed SDH flow
- SDH hierarchy divided into two groups:
  - multiplexing level (virtual containers, VCs)
  - line signal level (synchronous transport level, STM)
- Tributaries from E1 (2.048 Mbit/s) to E4 (139.264 Mbit/s) are synchronized (using justification bits if needed) and packed in containers of standardized size
- Control and supervisory information (POH, path overhead) added to containers => virtual container (VC)

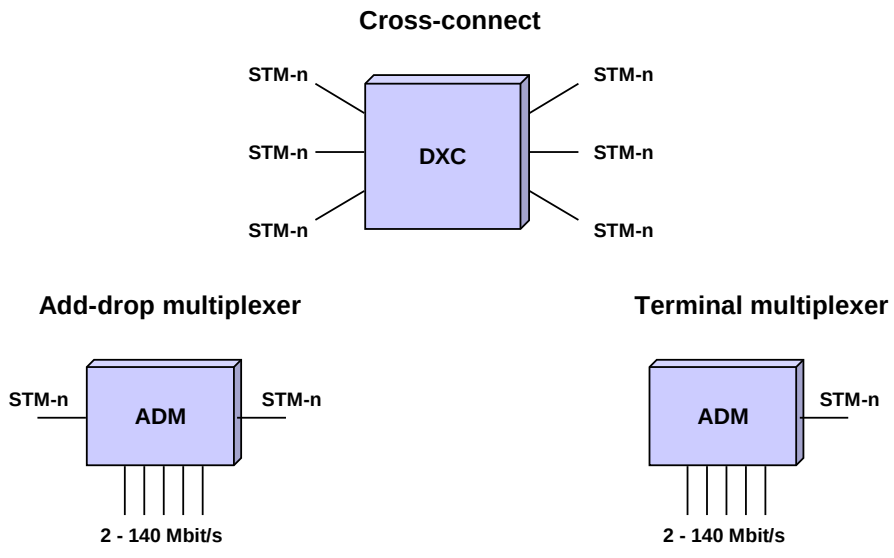
SDH-multiplexing (cont.)

- Different sized VCs for different tributaries (e.g. VC-12/E1, VC-3/E3, VC-4/E4)
- Smaller VCs can be packed into a larger VC (+ new POH)
- Section overhead (SOH) added to larger VC
  => transport module
- Transport module corresponds to line signal (bit flow transferred on the medium)
  - bit rate is 155.52 Mbit/s or its multiples
  - transport modules called STM-N (N = 1, 4, 16, 64, ...)
  - bit rate of STM-N is Nx155.52 Mbit/s
  - duration of a module is 125 μs (= duration of a PDH frame)

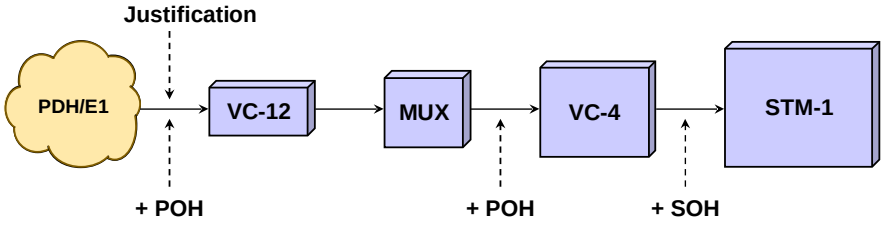
SDH network elements

- regenerator (intermediate repeater, IR)
  - regenerates line signal and may send or receive data via communication channels in RSOH header fields
- multiplexer
  - terminal multiplexer multiplexes/demultiplexes PDH and SDH tributaries to/from a common STM-n
  - add-drop multiplexer adds or drops tributaries to/from a common STM-n
- digital cross-connect
  - used for rearrangement of connections to meet variations of capacity or for protection switching
  - connections set up and released by operator

Example SDH network elements



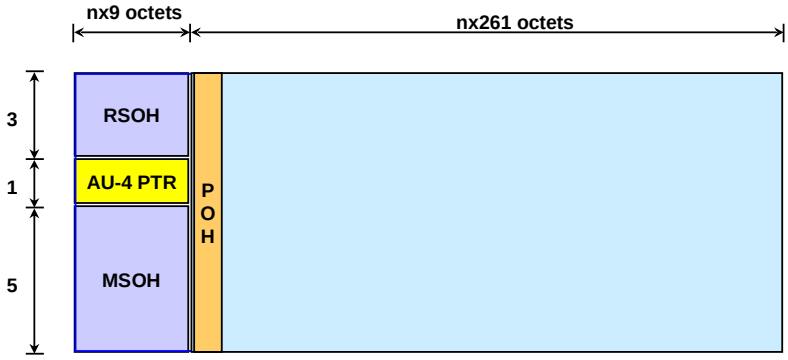
Generation of STM-1 frame



STM-*n* frame

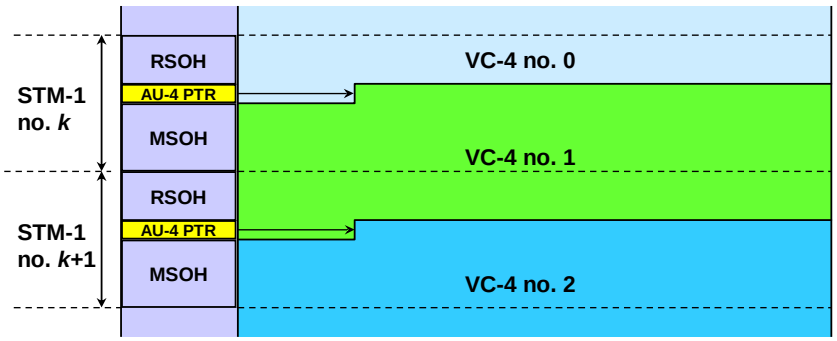
Three main fields:

- Regeneration and multiplexer section overhead (RSOH and MSOH)
- Payload and path overhead (POH)
- AU (administrative) pointer specifies where payload (VC-4 or VC-3) starts



Synchronization of payload

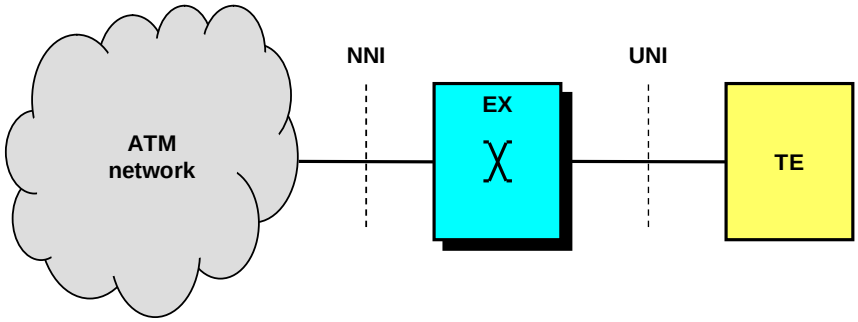
- Position of each octet in a STM frame (or VC frame) has a number
- AU pointer contains position number of the octet in which VC starts
- Lower order VC included as part of a higher order VC (e.g. VC-12 as part of VC-4)



Asynchronous Transfer Mode (ATM)

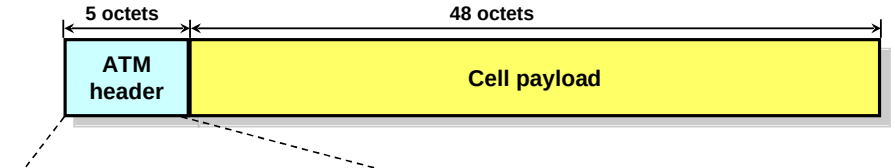
- **cell**
  - 53 octets
- **routing/switching**
  - based on VPI and VCI
- **adaptation**
  - processing of user data into ATM cells
- **error control**
  - cell header checking and discarding
- **flow control**
  - no flow control
  - input rate control
- **congestion control**
  - cell discarded (two priorities)

ATM reference interfaces



NNI - Network-to-Network Interface
UNI - User Network Interface
EX - Exchange Equipment
TE - Terminal Equipment

ATM cell structure



ATM header for UNI

GFC	VPI	
VPI	VCI	
VCI		
VCI	PTI	CPL
HEC		

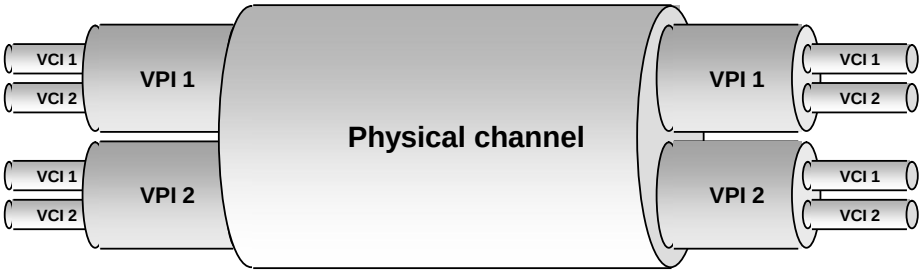
ATM header for NNI

VPI		
VPI	VCI	
VCI		
VCI	PTI	CPL
HEC		

UNI - User Network Interface
NNI - Network-to-Network Interface
VPI - Virtual Path Identifier
VCI - Virtual Channel Identifier
GFC - Generic Flow Control
PTI - Payload Type Identifier
CPL - Cell Loss Priority
HEC - Header Error Control

HEC = 8 x (header octets 1 to 4) / ($x^8 + x^2 + x + 1$)

ATM connection types

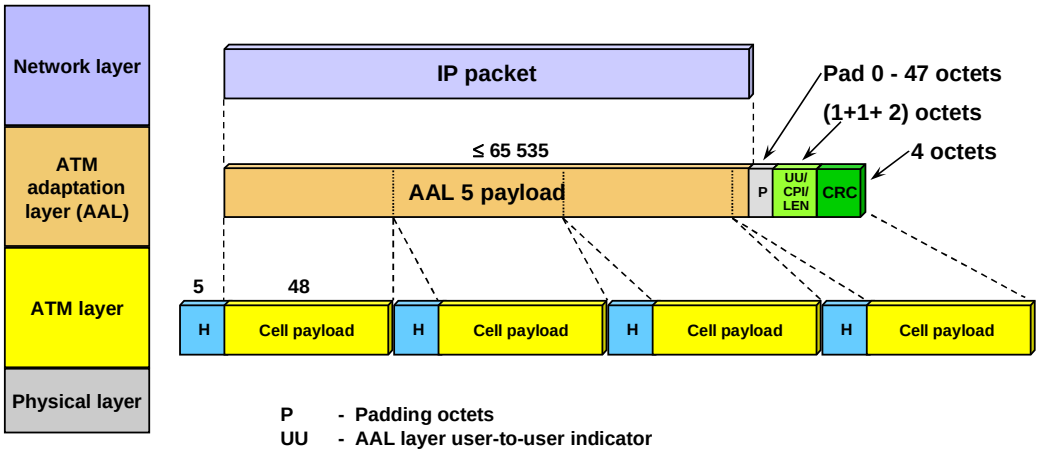


VCI *k* - Virtual Channel Identifier *k*
VPI *k* - Virtual Path Identifier *k*

Physical layers for ATM

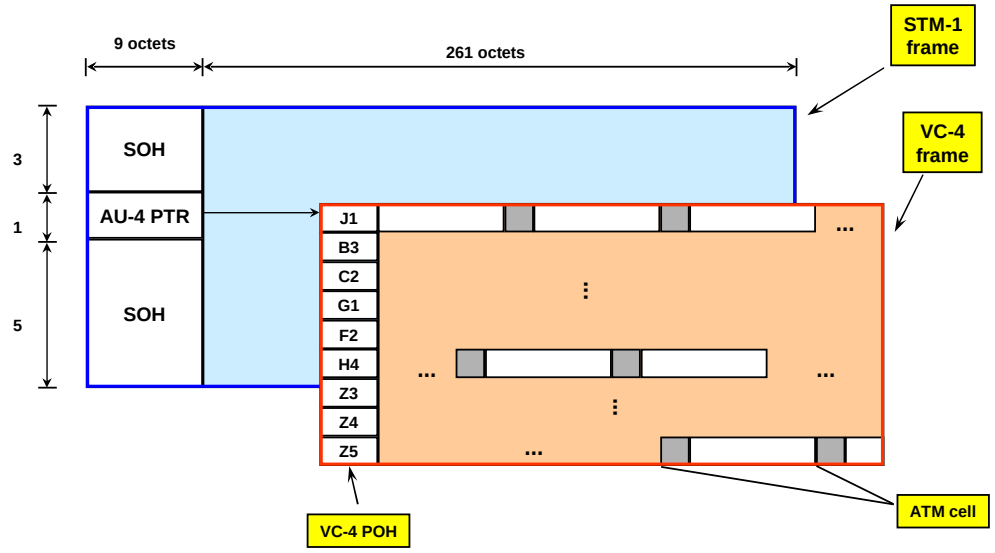
- **SDH (Synchronous Digital Hierarchy)**
  - **STM-1 155 Mbit/s**
  - **STM-4 622 Mbit/s**
  - **STM-16 2.4 Gbit/s**
- **PDH (Plesiochronous Digital Hierarchy)**
  - **E1 2 Mbit/s**
  - **E3 34 Mbit/s**
  - **E4 140 Mbit/s**
- **TAXI 100 Mbit/s and IBM 25 Mbit/s**
- **Cell based interface**
  - **uses standard bit rates and physical level interfaces (e.g. E1, STM-1 or STM-4)**
  - **HEC used for framing**

Transport of data in ATM cells

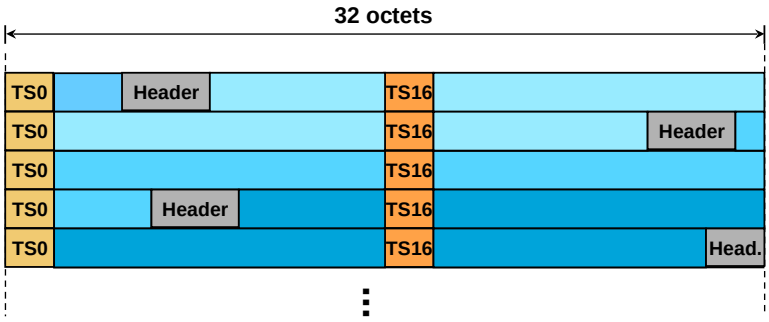


P - Padding octets
UU - AAL layer user-to-user indicator
CPI - Common part indicator
LEN - Length indicator

ATM cell encapsulation / SDH



ATM cell encapsulation / PDH (E1)

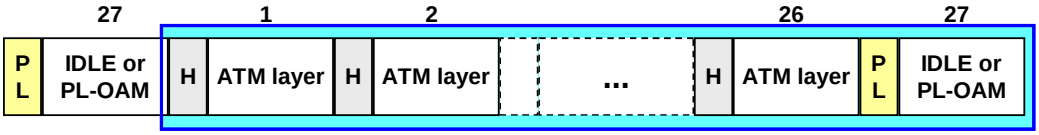


TS0
- frame alignment
- F3 OAM functions
  - loss of frame alignment
  - performance monitoring
  - transmission of FERF and LOC
  - performance reporting

TS16
- reserved for signaling

Cell based interface

Frame structure for cell base interfaces:



- PL cells processed on physical layer (not on ATM layer)
- IDLE cell for cell rate adaptation
- PL-OAM cells carry physical level OAM information (regenerator (F1) and transmission path (F3) level messages)
- PL cell identified by a pre-defined header
  - 00000000 00000000 0000000 00000001 (IDLE cell)
  - 00000000 00000000 0000000 00001001 (phys. layer OAM)
  - xxxx0000 00000000 0000000 0000xxxx (reserved for phys. layer)

H = ATM cell Header, PL = Physical Layer, OAM = Operation Administration and Maintenance

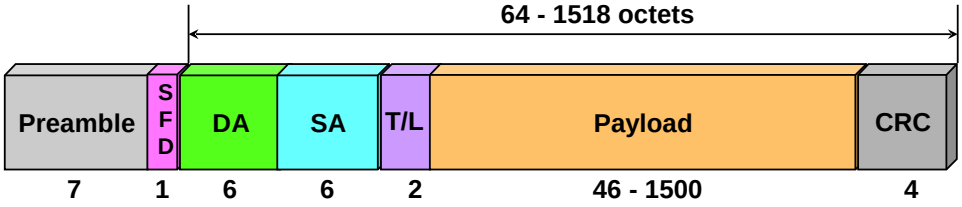
ATM network elements

- Cross-connect
  - switching of virtual paths (VPs)
  - VP paths are statically connected
- Switch
  - switching of virtual channel (VCs)
  - VC paths are dynamically or statically connected
- DSLAM (Digital Subscriber Line Access Multiplexer)
  - concentrates a larger number of sub-scriber lines to a common higher capacity link
  - aggregated capacity of subscriber lines surpasses that of the common link

Ethernet

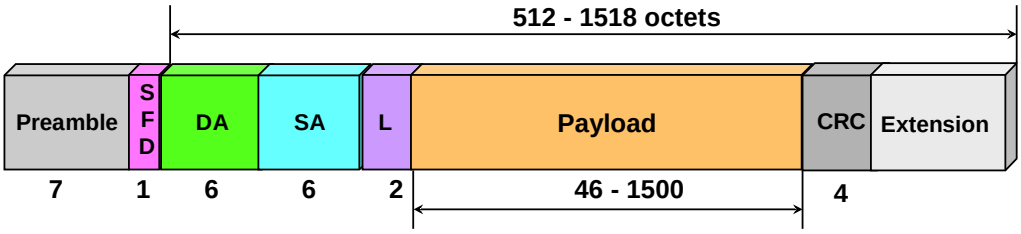
- Originally a link layer protocol for LANs (10 and 100 MbE)
- Upgrade of link speeds
  => optical versions 1GbE and 10 GbE
  => suggested for long haul transmission
- No connections - each data terminal (DTE) sends data when ready - MAC is based on CSMA/CD
- Synchronization
  - line coding, preamble pattern and start-of-frame delimiter
  - Manchester code for 10 MbE, 8B6T for 100 MbE, 8B10B for GbE

Ethernet frame



Preamble - AA AA AA AA AA AA AA (Hex)
SFD - Start of Frame Delimiter AB (Hex)
DA - Destination Address
SA - Source Address
T/L - Type (RFC894, Ethernet) or Length (RFC1042, IEEE 802.3) indicator
CRC - Cyclic Redundancy Check
Inter-frame gap 12 octets (9,6 µs /10 MbE)

1GbE frame



Preamble - AA AA AA AA AA AA AA (Hex)
SFD - Start of Frame Delimiter AB (Hex)
DA - Destination Address
SA - Source Address
T/L - Type (RFC894, Ethernet) or Length (RFC1042, IEEE 802.3) indicator
CRC - Cyclic Redundancy Check
Inter-frame gap 12 octets (96 ns /1 GbE)
Extension - for padding short frames to be 512 octets long

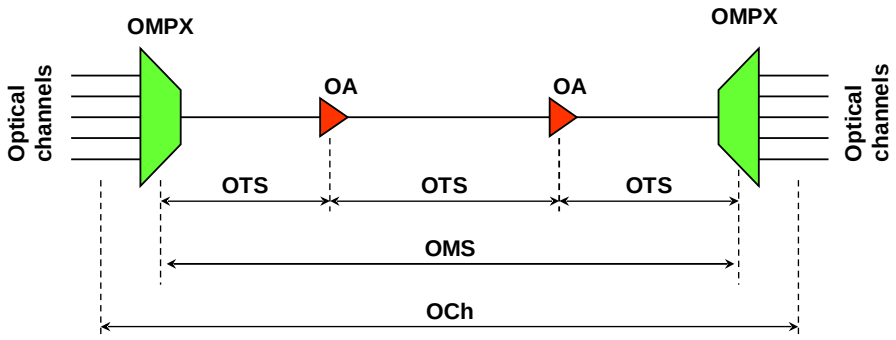
Ethernet network elements

- **Repeater**
  - interconnects LAN segments on physical layer
  - regenerates all signals received from one segment and forwards them onto the next
- **Bridge**
  - interconnects LAN segments on link layer (MAC)
  - all received frames are buffered and error free ones are forwarded to another segment (if they are addressed to it)
- **Hub and switch**
  - hub connects DTEs with two twisted pair links in a star topology and repeats received signal from any input to all output links
  - switch is an intelligent hub, which learns MAC addresses of DTEs and is capable of directing received frames only to addressed ports

Optical transport network

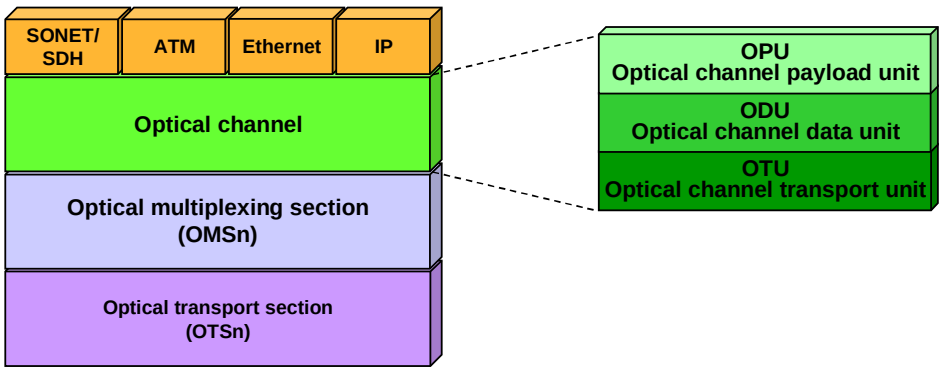
- Optical Transport Network (OTN), being developed by ITU-T (G.709), specifies interfaces for optical networks
- Goal to gather for the transmission needs of today's wide range of digital services and to assist network evolution to higher bandwidths and improved network performance
- OTN builds on SDH and introduces some refinements:
  - management of optical channels in optical domain
  - FEC to improve error performance and allow longer link spans
  - provides means to manage optical channels end-to-end in optical domain (i.e. no O/E/O conversions)
  - interconnections scale from a single wavelength to multiple ones

OTN reference model



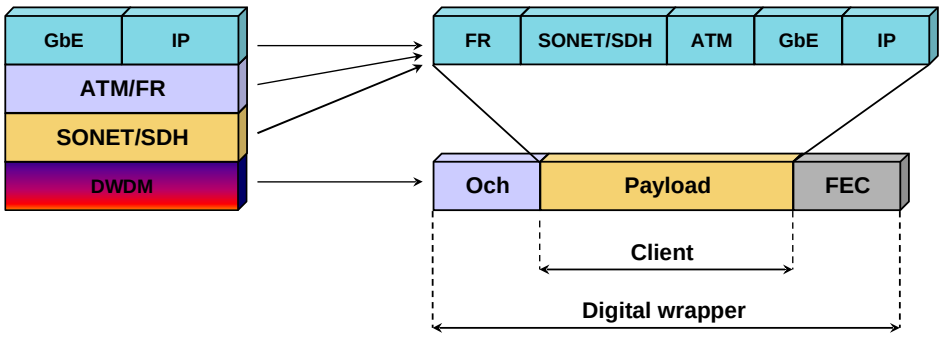
- OCh Optical Channel
- OA Optical Amplifier
- OMS Optical Multiplexing Section
- OMPX Optical Multiplexer
- OTS Optical Transport Section

OTN layers and OCh sub-layers

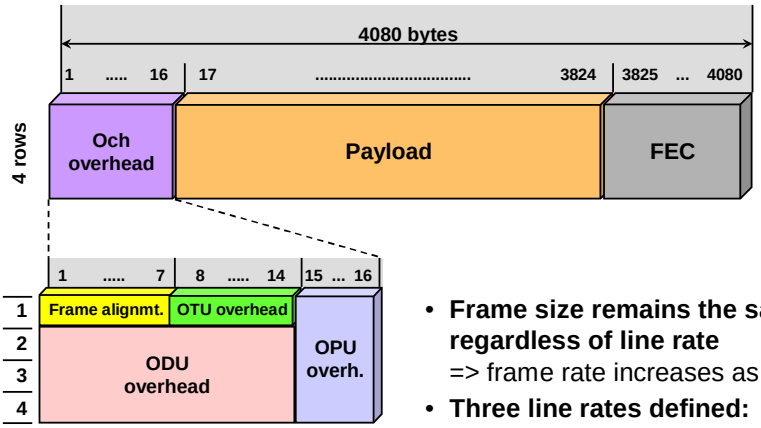


OTN frame structure

- **Three main fields**
  - Optical channel overhead
  - Payload
  - Forward error indication field



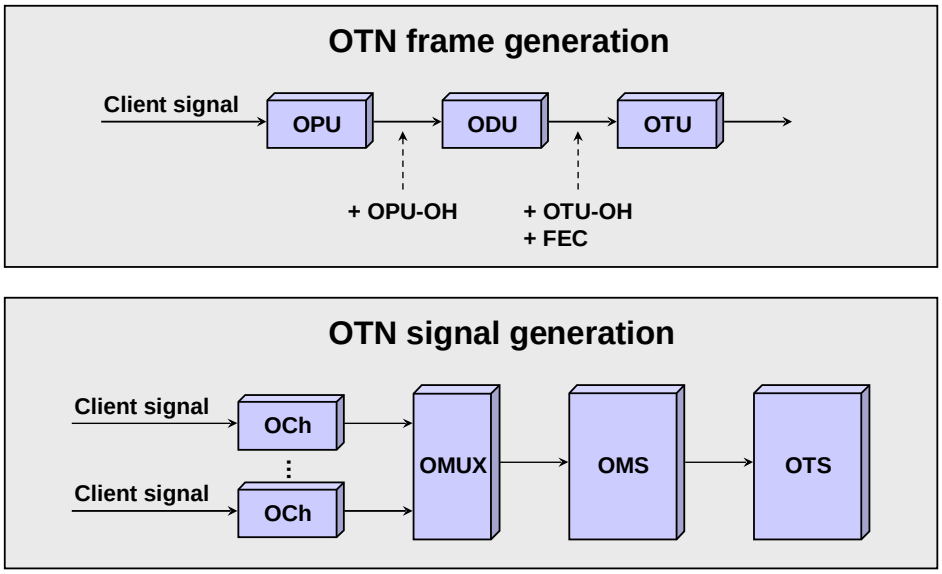
OTN frame structure (cont.)



OTU - Optical transport unit
ODU - Optical data unit
OPU - Optical payload unit
FEC - Forward error correction

- **Frame size remains the same (4x4080) regardless of line rate**
  => frame rate increases as line rate increases
- **Three line rates defined:**
  - OTU1 2.666 Gbit/s
  - OTU2 10.709 Gbit/s
  - OTU3 43.014 Gbit/s

Generation of OTN frame and signal



OTN network elements

- **optical amplifier**
  - amplifies optical line signal
- **optical multiplexer**
  - multiplexes optical wavelengths to OMS signal
  - add-drop multiplexer adds or drops wavelengths to/from a common OMS
- **optical cross-connect**
  - used to direct optical wavelengths (channels) from an OMS to another
  - connections set up and released by operator
- **optical switches ?**
  - when technology becomes available optical switches will be used for switching of data packets in the optical domain

Generic Framing Procedure (GFP)

- Recently standardized traffic adaptation mechanism especially for transporting block-coded and packet-oriented data
- Standardized by ITU-T (G.7041) and ANSI (T1.105.02) (the only standard supported by both organizations)
- Developed to overcome data transport inefficiencies of existing ATM, POS, etc. technologies
- Operates over byte-synchronous communications channels (e.g. SDH/SONET and OTN)
- Supports both fixed and variable length data frames
- Generalizes error-control-based frame delineation scheme (successfully employed in ATM)
  - relies on payload length and error control check for frame boundary delineation

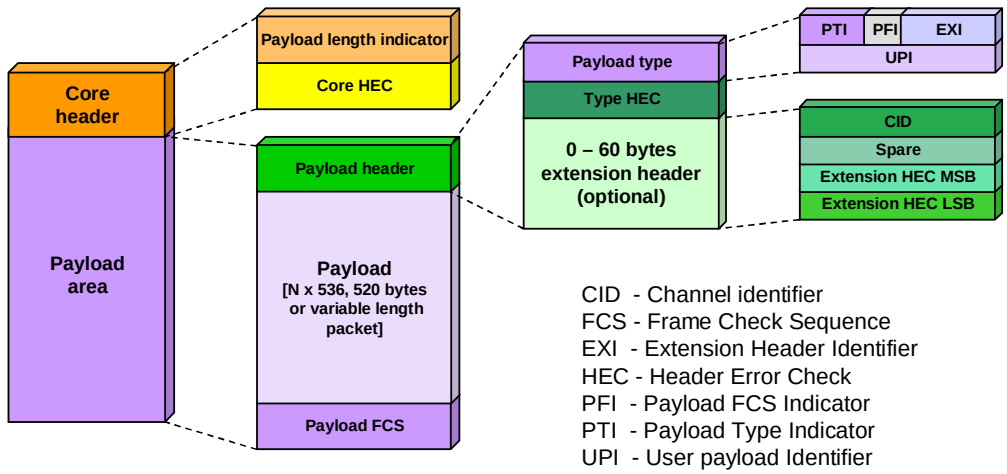
GFP (cont.)

- Two frame types: client and control frames
  - client frames include client data frames and client management frames
  - control frames used for OAM purposes
- Multiple transport modes (coexistent in the same channel) possible
  - Frame-mapped GFP for packet data, e.g. PPP, IP, MPLS and Ethernet)
  - Transparent-mapped GFP for delay sensitive traffic (storage area networks), e.g. Fiber Channel, FICON and ESCON

GFP client data frame

- Composed of a frame header and payload
- Core header intended for data link management
  - payload length indicator (PLI, 2 octets), HEC (CRC-16, 2 octets)
- Payload field divided into payload header, payload and optional FCS (CRC-32) sub-fields
- Payload header includes:
  - payload type (2 octets) and type HEC (2 octets) sub-fields
  - optional 0 - 60 octets of extension header
- Payload:
  - variable length (0 - 65 535 octets, including payload header and FCS) for frame mapping mode (GFP-F) - frame multiplexing
  - fixed size Nx[536, 520] for transparent mapping mode (GFP-T) - no frame multiplexing

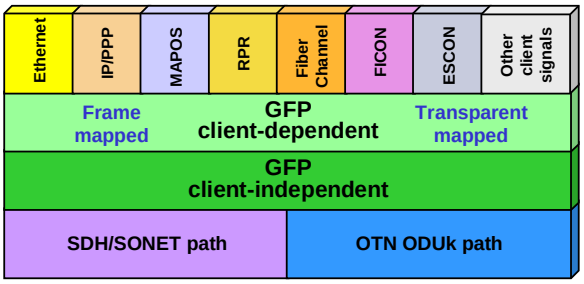
GFP frame structure



CID - Channel identifier
FCS - Frame Check Sequence
EXI - Extension Header Identifier
HEC - Header Error Check
PFI - Payload FCS Indicator
PTI - Payload Type Indicator
UPI - User payload Identifier

Source: IEEE Communications Magazine, May 2002

GFP relationship to client signals and transport paths

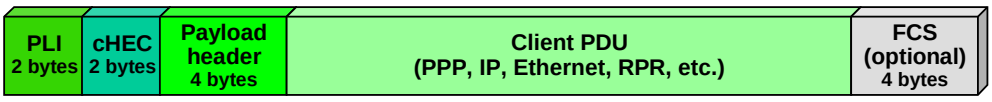


ESCON - Enterprise System CONnection
FICON - Fiber CONnection
IP/PPP - IP over Point-to-Point Protocol
MAPOS - Multiple Access Protocol over SONET/SDH
RPR - Resilient Packet Ring

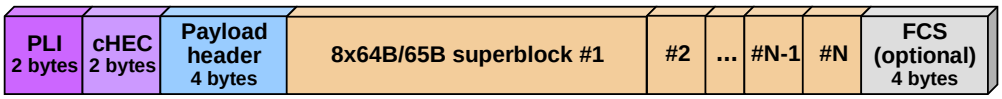
Source: IEEE Communications Magazine, May 2002

Adapting traffic via GFP-F and GFP-T

GFP-F frame

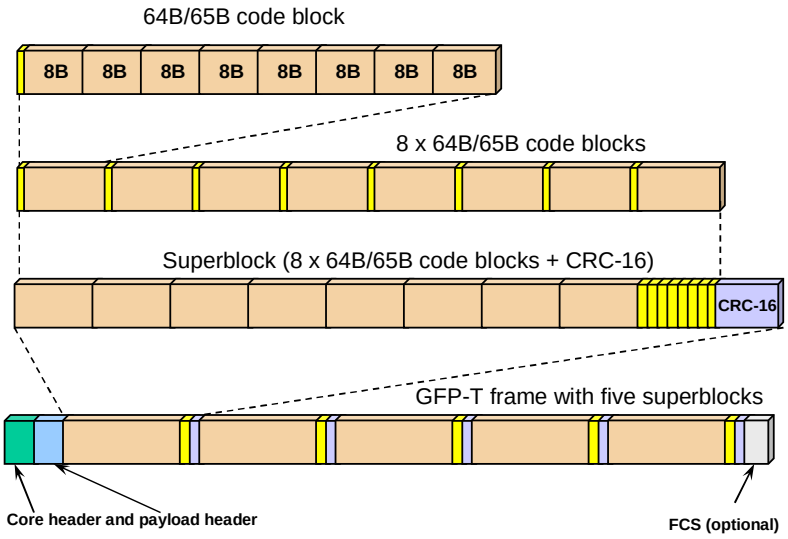


GFP-T frame



FCS - Frame Check Sequence
cHEC - Core Header Error Control
PDU - Packet Data Unit
PLI - Payload Length Indicator

GFP-T frame mapping



Switch Fabrics

Switching Technology S38.3165
http://www.netlab.hut.fi/opetus/s383165

Switch fabrics

- Basic concepts
- Time and space switching
- Two stage switches
- Three stage switches
- Cost criteria
- Multi-stage switches and path search

Switch fabrics (cont.)

- Multi-point switching
- Self-routing networks
- Sorting networks
- Fabric implementation technologies
- Fault tolerance and reliability

Basic concepts

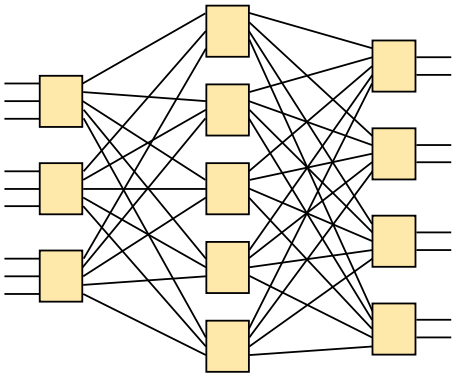
- Accessibility
- Blocking
- Complexity
- Scalability
- Reliability
- Throughput

Accessibility

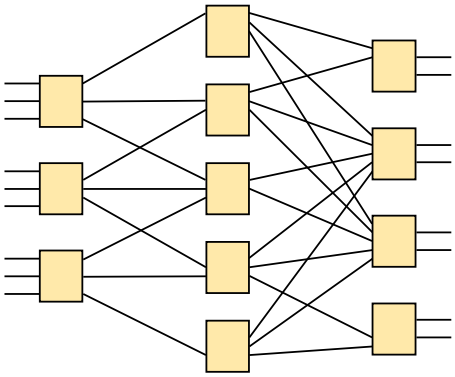
- A network has **full accessibility (= connectivity)** when each inlet can be connected to each outlet (in case there are no other I/O connections in the network)
- A network has a **limited accessibility** when the above given property does not exist
- Interconnection networks applied in today's switch fabrics usually have full accessibility

Accessibility (cont.)

Example of full accessibility



Example of limited accessibility



Blocking

- Blocking is defined as failure to satisfy a connection request and it depends strongly on the combinatorial properties of the switching networks

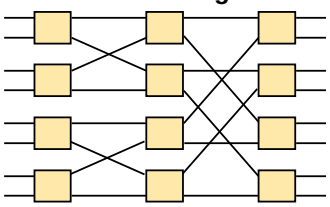
Network class	Network type	Network state
Non-blocking	Strict-sense non-blocking	Without blocking states
	Wide-sense non-blocking	With blocking state
	Rearrangeably non-blocking	
Blocking	Others	

Blocking (cont.)

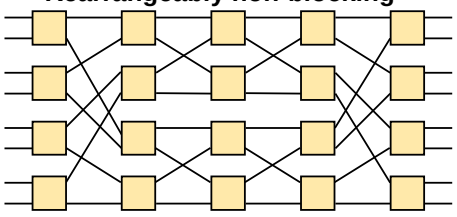
- **Non-blocking** - a path between an arbitrary idle inlet and arbitrary idle outlet can always be established independent of network state at set-up time
- **Blocking** - a path between an arbitrary idle inlet and arbitrary idle outlet cannot be established owing to internal congestion due to the already established connections
- **Strict-sense non-blocking** - a path can always be set up between any idle inlet and any idle outlet without disturbing paths already set up
- **Wide-sense non-blocking** - a path can be set up between any idle inlet and any idle outlet without disturbing existing connections, provided that certain rules are followed. These rules prevent network from entering a state for which new connections cannot be made
- **Rearrangeably non-blocking** - when establishing a path between an idle inlet and an idle outlet, paths of existing connections may have to be changed (rearranged) to set up that connection

Examples of different sorts of blocking networks

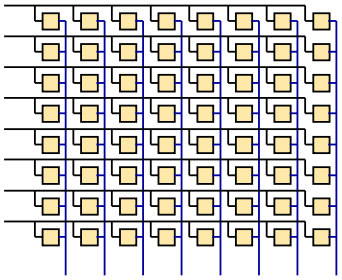
Blocking



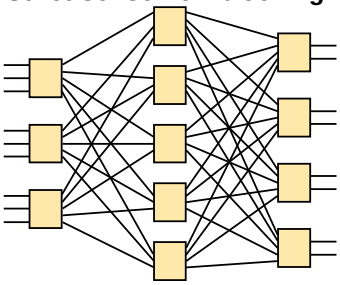
Rearrangeably non-blocking



Strict-sense non-blocking



Strict-sense non-blocking

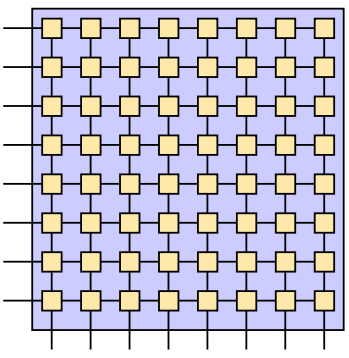


Complexity

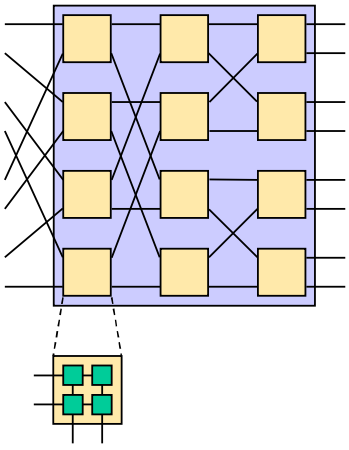
- Complexity of an interconnection network is expressed by **cost index**
- Traditional definition of cost index gives the **number of cross-points in a network**
  - used to be a reasonable measure of space division switching systems
- Nowadays cost index alone does not characterize cost of an interconnection network for broadband applications
  - VLSIs and their integration degree has changed the way how cost of a switch fabric is formed (number of ICs, power consumption)
  - management and control of a switching system has a significant contribution to cost

Complexity (cont.)

Cost index of an 8x8 crossbar is 64 (cross-points)



Cost index of an 8x8 banyan is 12x4= 48 (cross-points)



Scalability

- Due to constant increase of transport links and data rates on links, scalability of a switching system has become a key parameter in choosing a switch fabric architecture
- Scalability describes ability of a system to evolve with increasing requirements
- Issues that are usually matter of scalability
  - number of switching nodes
  - number of interconnection links between nodes
  - bandwidth of interconnection links and inlets/outlets
  - throughput of switch fabric
  - buffering requirements
  - number of inlets/outlets supported by switch fabric

Scalability (cont.)

Example of scalability

- a switching equipment has room for 20 line-cards and the original design supports 10 Mbit/s interfaces (one per line card)
- throughput of switch fabric is scalable from 500 Mbit/s to 2 Gbit/s
- when new line cards that each implement two 10 Mbit/s interfaces are introduced, the interface logic may have to be upgraded
- when new line cards that implement a 100 Mbit/s interface (one per line-card) are introduced, the switch fabric has to be upgraded (scaled up) to 2 Gbit/s speed and the interface logic has to be upgraded to 100 Mbit/s speed
- buffering memories need to be replaced by faster (and possible larger) ones
- larger number (>20) of line cards implies at least new physical design
- increase of line rates beyond 100 Mbit/s means redesign of switch fabric

Reliability

- Reliability and fault tolerance are system measures that have an impact on all functions of a switching system
- Reliability defines probability that a system does not fail within a given time interval provided that it functions correctly at the start of the interval
- Availability defines probability that a system will function at a given time instant
- Fault tolerance is the capability of a system to continue its intended function in spite of having a fault(s)
- Reliability measures:
  - MTTF (Mean Time To Failure)
  - MTTR (Mean Time To Repair)
  - MTBF (Mean Time Between Failures)

Throughput

- Throughput gives forwarding/switching speed/efficiency of a switch fabric
- It is measured in bits/s, octets/s, cells/s, packet/s, etc.
- Quite often throughput is given in the range (0 ... 1.0], i.e. the obtained forwarding speed is normalized to the theoretical maximum throughput

Switch fabrics

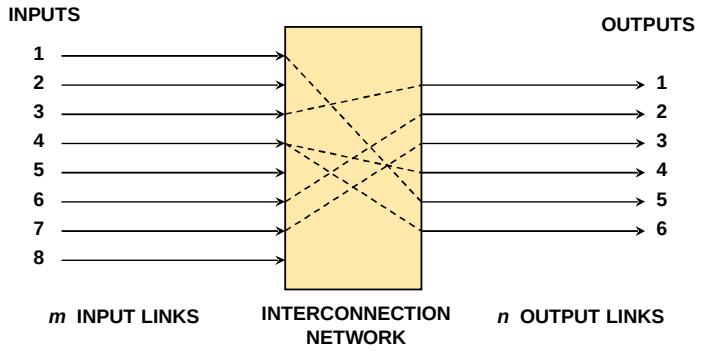
- Basic concepts
- **Time and space switching**
- Two stage switches
- Three stage switches
- Cost criteria
- Multi-stage switches and path search

Switching mechanisms

- A switched connection requires a mechanism that attaches the right information streams to each other
- Switching takes place in the switch fabric, the structure of which depends on network's mode of operation, available technology and required capacity
- Communicating terminals may use different physical links and different time-slots, so there is an obvious need to switch both in time and in space domain
- **Time and space** switching are basic functions of a switch fabric

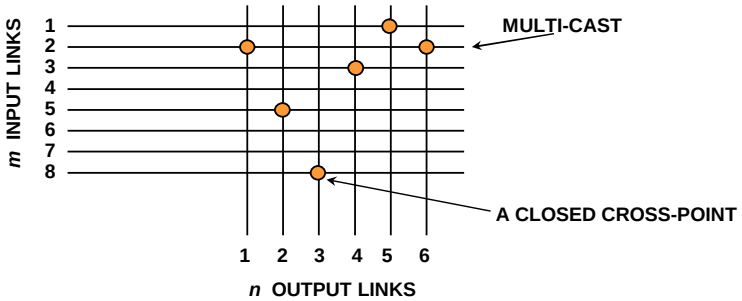
Space division switching

- A space switch directs traffic from input links to output links
- An input may set up one connection (1, 3, 6 and 7), multiple connections (4) or no connection (2, 5 and 8)



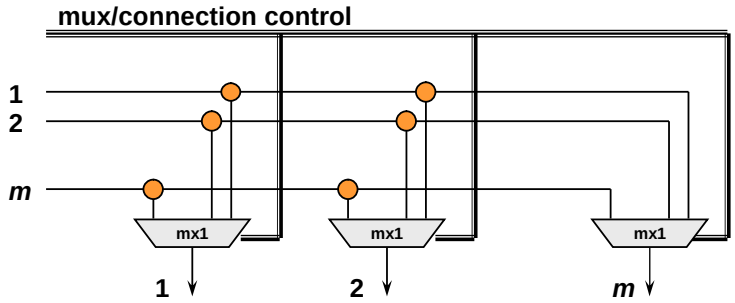
Crossbar switch matrix

- Crossbar matrix introduces the basic structure of a space switch
- Information flows are controlled (switched) by opening and closing cross-points
- ***m*** inputs and ***n*** outputs => ***mn*** cross-points (connection points)
- Only one input can be connected to an output at a time, but an input can be connected to multiple outputs (multi-cast) at a time



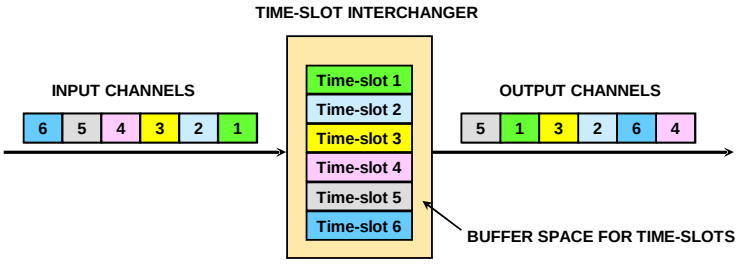
An example space switch

- *m* x1 -multiplexer used to implement a space switch
- Every input is fed to every output mux and mux control signals are used to select which input signal is connected through each mux

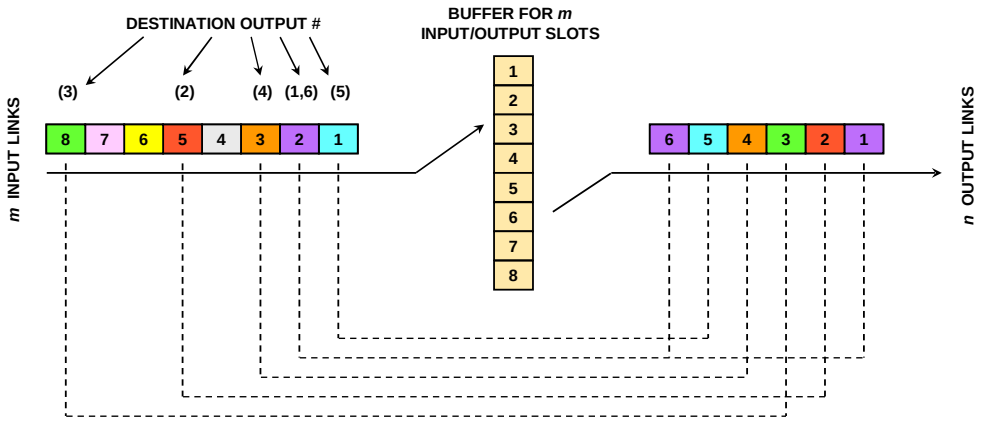


Time division multiplexing

- Time-slot interchanger is a device, which buffers *m* incoming time-slots, e.g. 30 time-slots of an E1 frame, arranges new transmit order and transmits *n* time-slots
- Time-slots are stored in buffer memory usually in the order they arrive or in the order they leave the switch - additional control logic is needed to decide respective output order or the memory slot where an input slot is stored

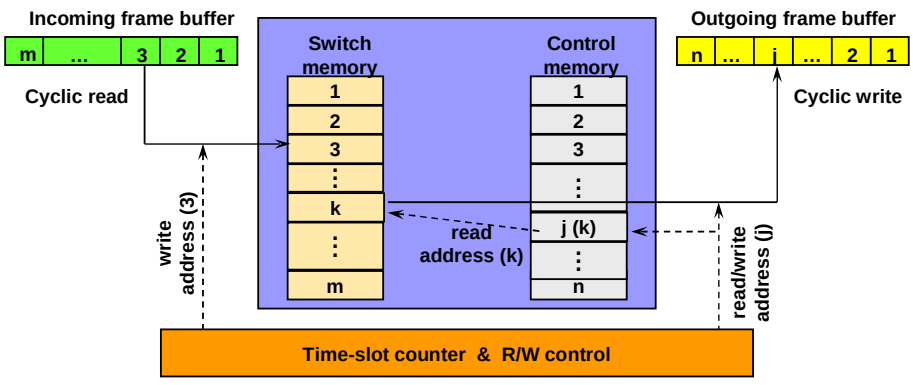


Time-slot interchange



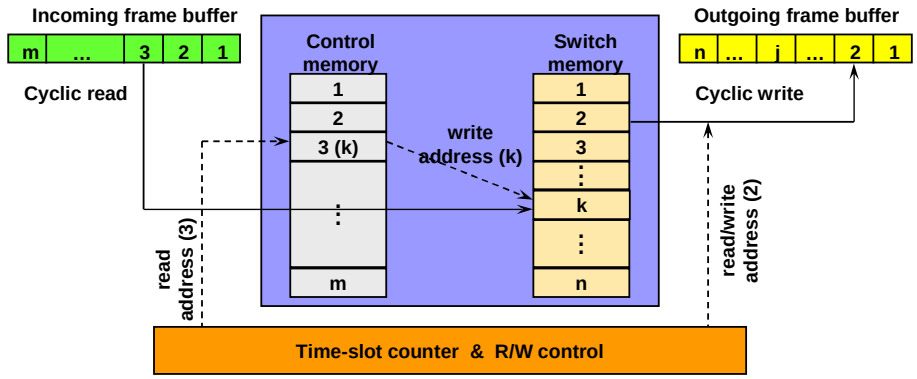
Time switch implementation example 1

- Incoming time-slots are written cyclically into switch memory
- Output logic reads cyclically control memory, which contains a pointer for each output time-slot
- Pointer indicates which input time-slot to insert into each output time-slot



Time switch implementation example 2

- Incoming time-slots are written into switch memory by using write-addresses read from control memory
- A write address points to an output slot to which the input slot is addressed
- Output time-slots are read cyclically from switch memory

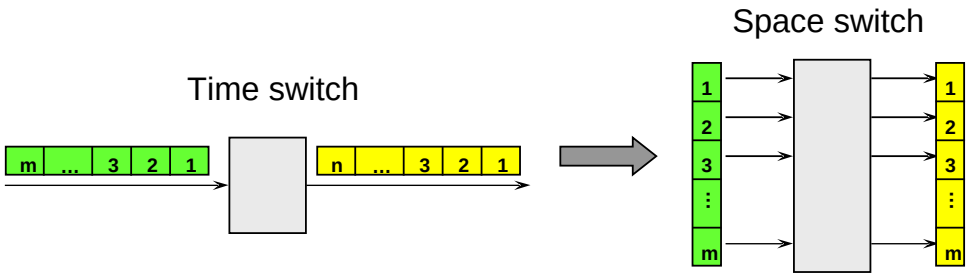


Properties of time switches

- Input and output frame buffers are read and written at wire-speed, i.e. *m* R/Ws for input and *n* R/Ws for output
- Interchange buffer (switch memory) serves all inputs and outputs and thus it is read and written at the aggregate speed of all inputs and outputs
  => speed of an interchange buffer is a critical parameter in time switches and limits performance of a switch
- Memory speed requirement can be cut by utilizing parallel to serial conversion
- Speed requirement of control memory is half of that of switch memory (in fact a little moor than that to allow new control data to be updated)

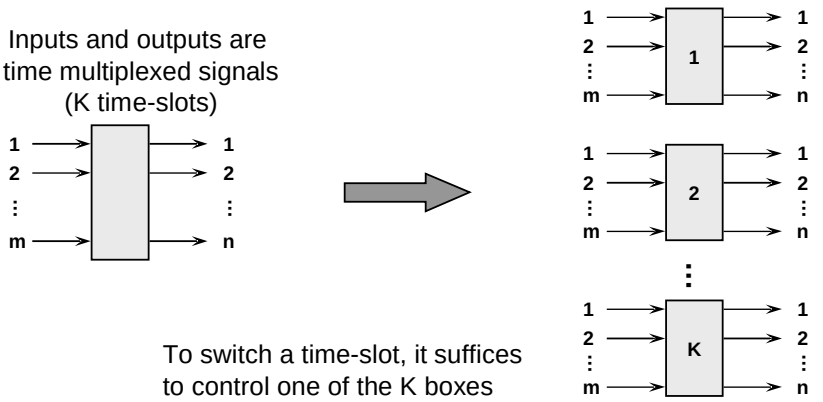
Time-Space analogy

- A time switch can be logically converted into a space switch by setting time-slot buffers into vertical position => time-slots can be considered to correspond to input/output links of a space switch
- But is this logical conversion fair ?



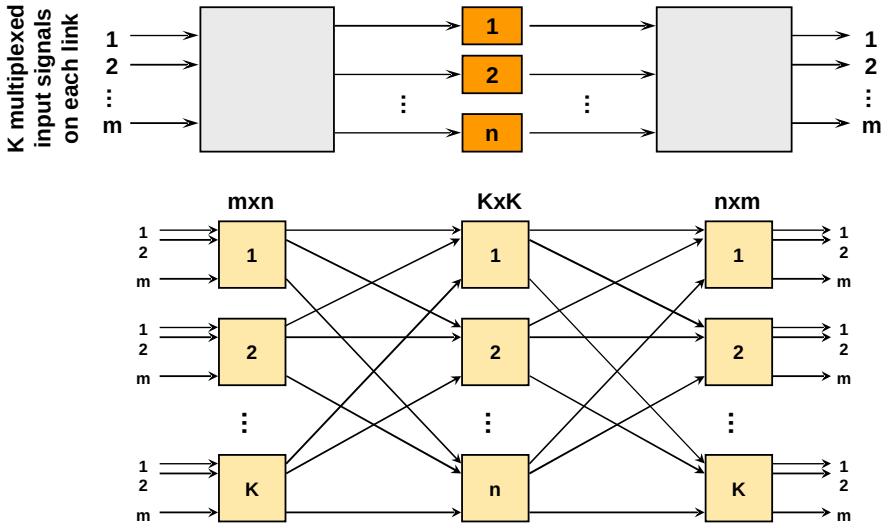
Space-Space analogy

- A space switch carrying time multiplexed input and output signals can be logically converted into a pure space switch (without cyclic control) by distributing each time-slot into its own space switch



To switch a time-slot, it suffices to control one of the K boxes

An example conversion



Properties of space and time switches

Space switches

- number of cross-points (e.g. AND-gates)
  - *m* input x *n* output = *mn*
  - when *m*=*n* => n^2
- output bit rate determines the speed requirement for the switch components
- both input and output lines deploy "bus" structure
  => fault location difficult

Time switches

- size of switch memory (SM) and control memory (CM) grows linearly as long as memory speed is sufficient, i.e. SM + CM + input buffering + output buffering
  = 2 x 2 x number of time-slots
- a simple and cost effective structure when memory speed is sufficient
- speed of available memory determines the maximum switching capacity