TELE 9301 review

The TELE 9301 exam

70% of the marks for this course
3 hours, attempt all questions
7 questions, but each has approx 4 parts (about 5 marks ea)
Answers are free-form/"short answer" (no multiple choice)
Covers all lectures (weighting roughly as per number of lectures)

Style of questions:
Test understanding of key concepts, not memorisation of protocol details.
Mainly descriptive: "Describe one technique for ..." "What is ..." "Discuss the relative merits of ..." "Why does ..."
Some applications of principles to specific problems (20%?).

Exam suggestions slides

Related events

Consultation
Tuesdays 5-6pm, each week until the exam
Starting in lecturer’s office (341), possibly moving to tutorial rooms on 2nd floor.

Special Consideration
“You must make formal application for Consideration for the course/s affected as soon as practicable after the problem occurs and within three working days of the assessment to which it refers.”
But note previous warning about supplementaries.

Review lectures

Nothing new this week
Highlights some of the more important topics.
Approx. 70-80% of the exam
Obviously not a replacement for the original lectures.

The slides are copied verbatim from those used in “earlier years”
• verbatim ⇒ references such as “later” & “next week” are often obsolete
• “earlier years”: Where they differ from those used this year, base your study on this year’s slides.
Switched networks

Most traffic is directed and bursty

Switches
- Forward traffic only towards its destination(s)
- Multiplex traffic from multiple sources

Advantages:
- Economical for large scale
- Relatively secure
- Reliable
- Simple to upgrade ⇒ supports heterogeneity

Caveats:
- Switches cost
- Switches may get congested or “block”
- Switches introduce delay

“Switch” defined

Switch: Any device with multiple I/O ports that aims to direct input traffic only to the port(s) that leads to the destination(s).
- The traffic may have multiple destinations, e.g. for multicast traffic.

What isn’t a switch:
- A multi-port device that directs input traffic to all ports. Call it a hub, combiner, etc.
- A single-port device (one input, one output). Call it a filter or switching element

We’ll consider detailed definitions of types of switches (routers, bridges, etc) next week.
Switching at various layers

Physical: all-optical networks: WDM, MEMS
Data link: bridging
Network: routing

Transport:
Strict interpretation: Switching between processes, e.g. Network Address Port Translation & load balancing.
Loose interpretation: Network layer that is aware of transport layer protocols, e.g. giving telnet priority over FTP

Application: Content Distribution Networks, caching

Ties between switching and reliability

Integrity: Affects switch forwarding modes
Completeness:
• Packet discard strategies to avoid unnecessary retransmissions.
• Loss priorities for layered video (baseband + enhancement)
Uniqueness: Negligible impact on switches.
Sequence:
• Switches may misorder when rearranging or recirculating traffic
• Deflection networks may misorder
• Preservation of sequence is a benefit of connection-oriented switching
Relevance: Protect addressing information in headers with CRCs
Delivery: Middleboxes may break end-to-end semantics (e.g. wireless Access Point may send TCP ack before segment received by end-system)

Spectrum of switching forms

Packet vs circuit switching

Packet switching:
• Bits of data to be switched are organised into discrete units of information, called “packets”.
• Switching decisions are influenced by the content of the packets†
  e.g. info from port P1 frequency F1 should go to output port P2 frequency F2

Circuit switching:
• Switching decisions are predetermined and are not influenced by the content of the information being switched.
• Examples of predetermining factors: port, frequency, time of arrival.
  e.g. from port P1 frequency F1 should go to output port P2 frequency F2

† Packet switching may also be influenced by other predetermined factors, e.g. a router discarding packets that have a source address that is invalid for the port from which they arrived.
Implications of circuit switching

Switch behaves predictably, independently of information being switched.

Circuit switching tends to (correlations, not requirements):
- Have little buffering within the network
- Operate in a connection-oriented manner (more later)
- Assign a fixed capacity to the circuit for the duration of the transfer (rather than re-negotiating).

⇒ Source must be able to predict its requirements.
- Offer coarse granularity of transmission rates.
- Preserve sequence of transmitted information.

Possible packet header information†

Address of end-system: Allows packets to be self-contained “datagrams”.
- “best effort” service = “I’ll do my best but make no guarantees”
- “best effort” ≠ better than other service.

Prototypical examples: Ethernet, Internet Protocol.

Addresses of intermediate systems = “source routing”

Labels/tags: Distinguish this flow of packets from others, e.g. IPv6
- FlowID, ATM VC/VPI.
  - Switch consults table as to what to do with the flow.
  - Switches can swap tags ⇒ label on a packet may change as it propagates.

† for routing. Headers may contain other info to indicate type of service, and for other functions.

Switch classifications

We’ll consider multiple ways of classifying switches:
1. By location in hierarchical network
2. By functionality
3. By modularity of implementation
4. By form of switching fabric (details later)
Switch trends as location in hierarchy changes

How do switches change as you move into the network core?:

1. **Availability** becomes increasingly important
   - High-reliability components
   - Redundancy in power supplies, even redundant fabrics
   - Hot swapping of line interfaces & power supplies
   - May employ “protection switches” to bypass failed switches (low switching rate, high throughput)

2. **Throughput** becomes increasingly important
   - (though load may be less variable)

3. **Reduced functionality**, e.g. NAT, DHCP servers, firewalls, QOS tend to be implemented in workgroup switches but not core switches.

Switch classification 3:

**By modularity of implementation**

- **Bounded systems**: fixed, pre-determined configuration.
- **Stackable switches**: intra-stack connection:
  - high-speed port (e.g. Gigabit Ethernet)
  - Low Voltage Differential Signaling (LVDS)
- **Chassis switches**: Increasing:
  - cost
  - performance
  - flexibility

Switch classification 4:

**By structure of implementation**

- **Switching Architecture**
  - **Space division**: Single stage, Connsular Switch Exchange
  - **Layer division**: Single path, Barriers
  - **Time division**: Multiple path, Augmented banyan
  - **Shared medium**: Data multiplexers, Clos
  - **Ring bus**: Internal medium, External medium

Figure from H. Peyravi

We’ll consider some of these structures (e.g. crossbar, Banyan, time-division) shortly...
Course outline

Section outline

Introduction to Switching
- Fabrics
  - Generic switch architecture
  - Criteria for evaluating switches
  - Switch classification 4: By structure of implementation
    - Time-division switches
    - Multiplexing and demultiplexing
    - Shared transmission medium
    - Shared storage medium
  - Space-division switches
    - Single-stage: Crossbar
    - Multi-stage: Clos, Banyan
  - Optical switching
- Functions surrounding fabrics
- Distributed switching and protocols

Generic switch architecture

- Line interface cards (input & output)
- Port processors (input & output, may be different)
- Switching fabric (internal to switch)
- Control processor
- LICs x 2
- Fabric

We’ll consider these in depth shortly.

Fabric usually has own clock, and port processors are synchronized to it (e.g. with buffering). Some fabrics include internal buffering.

Blocking

Blocking: When traffic can’t pass through the switch.

Why blocking may occur:
- output blocking: output port is not available – in use by other traffic
- internal blocking: there is no path through the switch to get to the output

Where: output blocking may occur internally within the switch, e.g. in Banyan
Blocking: Response

Possible responses to blocking:

- **Queue** the request, e.g. buffering.
  - A later lecture will examine buffering at input and output ports to deal with such blocking.
  - This lecture will examine some fabrics that use buffering for the switching process

- **Discard** the request (sometimes called “clearing” blocked calls)

- **Schedule** the request for a time when resources are available

- **Try again**: Require the initiator to make the request again, e.g. recirculation

For example, the Session Initiation Protocol (SIP) for signalling allows a server to tell a client to “Retry-After” a certain period.

Blocking: Types

Types of switch blocking:

- **Nonblocking**: Any desired connection can be established immediately.
  - e.g. Crossbar

- **Rearrangeable nonblocking**: Any desired connection can be established, possibly after rerouting existing connections.
  - e.g. Clos

- **Blocking**: There exist connection sets that prevent additional connections from being established.
  - e.g. Banyan

“Head-of-line blocking” covered later under buffering

Shared transmission media switches

Most commonly produced form of switch

(c.f. academic emphasis on space-division switching)

A single transmission medium shared by all input and output ports, i.e. broadcast-and-select within the switch.

Output ports pick off packets destined to them, based on address, time of arrival, etc.

Limitation: high-speed filtering

Keshav Fig. 8.12

Time-Slot-Interchange circuit switch

1. Multiplexes inputs onto one medium (e.g. write port of memory)
2. Interchanges slot order by:
   1. Storing slots in a memory
   2. Reading slots out of the memory, but in a different order
   3. Demultiplexes output of memory to output ports.
Shared memory switches

Memory organized into linked lists, one for each output port
(Linked lists are useful in other switches to implement scheduling disciplines, e.g. fair queuing)
Incoming traffic is appended to the end of the list for the required output port
Single memory:
  Allows statistical gain: busy port can use memory not being used by an idle port
  Helps multicasting: Stored once in memory for multiple outputs (multiple links to payload)

Crossbar switch

At the intersection of each input and each output, there is a cross point, which can be selectively enabled, allowing communication from input to output.
*e.g.* Intel 470 switches, Cisco Catalyst 6500, 12000 series routers
Feasible to implement in VLSI (e.g. PMC-Sierra PM9312), problem is high-speed I/O to chip - pincount

Advantages of crossbar switches

- Simple structure
- Low latency – minimal number of connecting points between arbitrary input and output.
- No internal blocking (may have output port blocking)
- Electrical crossbars readily supports multicast
  Multicast is difficult with MEMS optical crossbars, since mirror will (hopefully) reflect all signal power to intended output port.

Disadvantages of crossbar switches

**Scalability:**
- Number of crosspoints \(N_x\) (complexity) grows with \(n^2\), where \(n\) is the number of ports
- Difficult to incrementally expand
- Output buffer speed increases in proportion to fabric, not line.

**Fanout:** Number of crosspoints on each line increases linearly with number of ports, increasing capacitive loading, slowing transmission

**No fault tolerance:** each crosspoint is needed for one connection or another.
Staged switches

Multistage switches are composed of "networks"† of smaller switches (e.g. crossbars or shared-media), often 2×2
Potential benefits:

- **Fewer crosspoints** than in a crossbar switch
- **Diversity of paths**

**How?** Selection of switch in first and last stage is determined by which input & output are being connected

**Why?**: Intermediate stages can offer choice of switch

- **Lower blocking probability**
- **Increased reliability**: can still connect input and output even if a component switch has failed

† Here we are interested with networks within the switch. Of course switches can also be interconnected to form external networks (the more common use of the word).

---

Clos switches

Each switch ("array") of a stage has one output feeding into each switch of the next stage (rectangular, not square).

Simplest Clos switches have 3 stages, but more stages are possible by replacing middle stages by 3-stage Clos structures.

Advantages: path diversity and fewer crosspoints...

\[
N_x = (n \times k)(N/n)(2 + ((N/n) \times (N/n))k = 2Nk + kN^2/n^2
\]

What are the optimal values for \(n\) and \(k\)?

---

Banyan switches

Self/source-routing using binary representation of output port

Direction for each stage specified by bit corresponding to that stage (MSb 1st)

⇒ can’t multicast

\(N\)-port switch has \(\log_2 N\) stages each with \(N/2\) 2×2 switches

\[N_x = 2 \log_2 N\]

"There is a very large, famous Banyan tree that occupies more than 3 acres of land in a park in Lahaina, on the island of Maui in the Hawaiian islands."

---

Types of Banyan blocking

**Internal blocking (†)**: Contention for an output port of a component switch in one of the early stages.

**Output port blocking (‡)**: Contention for an output port of a switch in the last stage.

Output port blocking may even occur internally (§).
How the optical domain differs

High-rate transmission

Re-evaluate decisions made in electronic networks

Electronic: Packet switching uses processing (routing decisions) to save transmission capacity

Optical: Circuit switching is currently more cost effective

Processing

e.g. optical address matching

Electronic: Route along shortest path (to save transmission) and buffer when ∃ contention on that path (buffering is cheap).

Optical: Deflection routing: Use longer path to avoid buffering

Buffering

WDM technology

Wavelengths = “lambdas” (λ)

DWDM = Dense WDM, e.g. 0.8nm spacing

Example of state of the art (March 2002):
64 channels @ 40Gb/s (2.56Tb/s) over 4000km

Tunable transmitters (multiple or tunable laser)

Tuning times range from ns (Distributed feedback) to ms (mechanical/acousto)

Tunable receivers – Burst mode can synchronize quickly, unlike continuous mode.

MEMS switches

2D

3D

Figures from Chu02
**Electro-optic packet switching**

*Payload* travels over a “lightpath” from source to destination; no optoelectronic conversion within network switching elements. *Control* information may incur optoelectronic conversion, may even flow on auxiliary electronic control network.

---

**Optical buffer implementation**

Fibre delay line
- Capacity ∝ length ∝ loss

Programmable fibre delay line
- Enables control of holding time
- Loss from splitting signal
- (Time-slot-interchange switch)

Active recirculating delay line
- Enables control of holding time
- Amp noise limits recirculations

---

**Outline**

Introduction to Switching Fabrics
Functions surrounding fabrics
Packet classification
Buffering
Scheduling
Distributed switching and protocols
Packet classification: applications

When "packets"† arrive at "switch"‡ inputs, need to use packet fields to look up state information:

Which output port?
- Switch input to the correct output. "Correct": port leads to destination
- Filtering (no output port): Prevent flow of traffic From certain addresses, e.g. from users who haven’t paid
- To certain addresses/services, e.g. firewall

What class of service?  
- "appropriate service" again

Record usage (billing, conditioning, etc)
- If packet can’t be classified, then switch may take some default action, e.g. discard it, send it to a default router, send it out on all ports, etc.

Packet fields that may be used for classification:

- Source and destination address fields
  - Destination address is the most important
  - But source address may be relevant:
    - for filtering when recording usage
    - for "load balancing" traffic from multiple sources to one destination follow different paths to spread load. Prefer all traffic from one source follow the same path to preserve sequence
    - for bridged networks† – learn station locations by observing source addresses
  - MAC: address fields are 48b
  - Network layer: address fields are 32b (IPv4) or 128b (IPv6)

For virtual circuit identifiers:
- 16b for ATM (8 or 12b VPI), 12b VLAN ID
- Fields indicating required Class of Service:
  - IPv4: Type of Service/DiffServ (IPv4), Traffic Class (IPv6)
  - IPv6: 20b Flow Label
  - 8b: Protocol (IPv4), Next Header (IPv6), e.g. TCP, UDP or ICMP

⇒ key width K. Classification is simple for small K (e.g. ≤ 32b), harder for larger K (e.g. 48b+)

Keys: inputs to classification (2)

Packet fields that may be used for classification:

- Virtual Circuit Identifiers: 16b for ATM (8 or 12b VPI), 12b VLAN ID
- Fields indicating required Class of Service:
  - 8b: Type of Service/DiffServ (IPv4), Traffic Class (IPv6)
  - IPv6: 20b Flow Label
  - 8b: Protocol (IPv4), Next Header (IPv6), e.g. TCP, UDP or ICMP

Classification variations

Do all stored rules apply to the same set of key digits?
- Yes is the simplest case, e.g. MAC bridges
- No can be more complicated, e.g.
  - IPv4 classful addresses: limited set of key digits
  - IPv6 classless addresses: Longest prefix matching
  - Firewalls: Different rules may apply to different fields

Can a key match multiple rules?
- No is the simplest case.
- Yes with Longest prefix matching and exceptions
Exact-match classification

e.g. for MAC addresses; IP addresses without aggregation exceptions

Techniques:
- Lookup tables
- Lists
- Hashing
- Content Addressable Memories
  + techniques that can classify by partial match

Content Addressable Memories

Each word of memory contains:
- $K$ bits of label storage and comparison logic

Figure shows the essence of a CAM with
- 4 words with 3b labels
- $D \leq 4$, $K=3$

Matching process:
- Key is distributed to all words simultaneously.
- Comparison operations are done in parallel.
- Words with matching labels generate match signal.

CAM example

Note that data doesn’t have to be stored in any particular order

Recall also, the RTL 8308 chip, used in the D-link DES-1008D shared memory switch, has a 128-entry CAM to accommodate hash-bashing.

Associating data with keys in a CAM

1. Store in CAM words
   Useful when key bits are programmable with a mask – remaining bits can store data.

2. Encode matching CAM word and use to index a RAM
Ternary CAMs

Ternary comparison: 0, 1, X ("Don’t care" = 0 or 1)

Don’t care bits generally specified as a mask (e.g. 1="Do care", 0="Don’t")

Word mask
word-specific, stored with label

Key mask
apply to all words; iterative search

E.g.
192.16.001xxxx.X port P
(192.16.32/19)

192.16.0010001.x.X port R
(192.16.34/23)

and prioritise match lines
s.t. lowest rule has priority

CAMs are good when excluded bits needn’t form a suffix, e.g. firewall

Tries

Trie (from retrieval):
"a tree where each node corresponds to a string that is defined by the path to that node from the root.

Rule Prefix

<table>
<thead>
<tr>
<th>Rule Prefix</th>
<th>Key in</th>
<th>Data out</th>
</tr>
</thead>
<tbody>
<tr>
<td>a 0*</td>
<td>10</td>
<td>0</td>
</tr>
<tr>
<td>b 01000*</td>
<td>11</td>
<td>1</td>
</tr>
<tr>
<td>c 011*</td>
<td>110</td>
<td>0</td>
</tr>
<tr>
<td>d 1*</td>
<td>11</td>
<td>0</td>
</tr>
<tr>
<td>e 100*</td>
<td>110</td>
<td>0</td>
</tr>
<tr>
<td>f 1100*</td>
<td>111</td>
<td>0</td>
</tr>
<tr>
<td>g 1101*</td>
<td>1110</td>
<td>1</td>
</tr>
<tr>
<td>h 1110*</td>
<td>1111</td>
<td>0</td>
</tr>
<tr>
<td>i 1111*</td>
<td>11110</td>
<td>0</td>
</tr>
<tr>
<td>j 11111*</td>
<td>111110</td>
<td>0</td>
</tr>
</tbody>
</table>

Application to partial match: Locate “don’t care” at end of key.
OK for CIDR and address aggregation, but not for firewalls

Section outline

Should a switch buffer?: Forwarding modes
How to buffer
What if the buffer overflows?

Forwarding modes

Forwarding could start as soon as packet has been classified;
DA in header+fast classification

could start forwarding before finish receiving, e.g. Ethernet frame:

<table>
<thead>
<tr>
<th>Address</th>
<th>Address</th>
<th>Type</th>
<th>(Data)</th>
<th>(Pad)</th>
<th>CRC</th>
</tr>
</thead>
<tbody>
<tr>
<td>6B</td>
<td>6B</td>
<td>2B</td>
<td>0-1500B</td>
<td>0-46B</td>
<td>4B</td>
</tr>
</tbody>
</table>

could start now

classification delay could start now

cut-through forwarding

can start now

check store-and-forward

can start now

Forwarding modes:

Store-and-forward: Store complete frame before forwarding. Saves transmission capacity.
Cut-through: Start forwarding frame as soon as know output port and port is free. May reduce delays.
Discard strategies

According to position in queue
Loss priorities
  - Source marking
  - Network marking
According to packet length
Discarding when buffers aren’t full:
  - Partial Packet Discard
  - Early Packet Discard
  - Random Early Detection

try to manage congestion

Explicit Congestion Notification

Idea is for routers to explicitly mark packets to indicate those that have experienced congestion, rather than source having to infer congestion from loss.

Don’t want to create extra network traffic during times of congestion ⇒ rather than router directly telling source, router sets bit in packet header, and destination (if it receives the packet) indicates congestion through field in normal feedback (ACKs)

Presence of ECN indicates congestion, but absence of ECN doesn’t prove non-congestion (congestion might have resulted in ECN signal being lost) ⇒ still have to respond to loss as a congestion indicator ⇒ still have ambiguity between congestion/transmission error

Token regulator

Packet† passage requires a token‡
If no tokens are stored, the packet can’t pass (may be buffered or discarded ⇒ “Regulator types” slide)
When a packet passes, a token must be removed from the bucket.
Periodically, tokens “drip” into the bucket. If the bucket is full, token is discarded.
Bucket depth \( (B \text{ or } \sigma) \Rightarrow (\text{but } \neq) \) maximum burst size.
Transmission of packets permitted by bucket will take time, and during that time, further tokens may be earned, allowing the burst to continue.
Drip rate (of tokens entering bucket) \( (R \text{ or } 1/\rho) \) determines long-term mean

Section outline

Scheduling
Goals:
  - Fairness
  - Protection
  - Ease of implementation
  - Ability to guarantee performance
  - Efficient admission control
Mechanisms:
  - First Come First Served
  - Round Robin
  - Generalised Processor Sharing
  - Fair Queueing
Fair Queueing

Rounds: A byte is sent from each active connection in each "round". Round number (R) can be fractional (part way through a round).

Finish number (F): Indicates order of packet transmissions finishing

Define a flow to be "active" when F flow's most current packet > R

dR/dt = 1/(# active)

Equations to determine Fi:

Packet of length Li arrives & is the i-th packet of:
An inactive flow: Fi = R + Li
An active flow: Fi = Fi-1 + Li

⇒ bursty arrivals have dispersed finish times

Parekh-Gallager theorem

g = least throughput the connection will receive at each WFQ scheduler on its path

Connection is Leaky-Bucket regulated such that # bits sent in time [t1, t2] ≤ ρ(t2 - t1) + σ

r(k) = rate of kth scheduler, g2p ∀ k

P = size of largest packet allowed

end to end delay ≤ σ + \sum_{i=1}^{P} P + \sum_{i=1}^{P} P / r(k)

⇒ WFQ can provide end-to-end delay bounds

But inefficient: low delay requires high throughput (g)
⇒ non-work conserving scheduling

Work conservation

Work conserving schemes are idle (not transmitting) only when they have no packets awaiting service

e.g. GPS

Conserving work means packet transmission time depends on presence of traffic from other flows (at output) ⇒ leads to delay variability.

Non-work conserving schemes can be idle when packets are queued

i.e. they may deliberately withhold a packet from transmission, when no other packet is being transmitted.
⇒ may not be able to fully utilise the output port.

e.g. rate-controlled scheduling: Preface the scheduler with a regulator (shaper) that delays packets until they are eligible (e.g. eligible according to a Leaky Bucket)
Course outline

Section outline

Introduction to Switching
Fabrics
Functions surrounding fabrics
Distributed switching and protocols
  Bridging
  Prioritisation and VLANs
  Signalling and SS7
  Frame Relay
  ATM
  MPLS, Intserv, Diffserv

Bridge process (Synopsis)

Whenever a frame arrives
1. Update database to record location of source
2. Filter out certain frames
3. Forward remaining frames
Learning by source lookup

Process:
Listen "promiscuously" (indiscriminately) to all frames, not just those traversing the bridge or destined to the bridge (e.g. management)

Observe the Source Address of these frames. Assuming that the links are bidirectional, source station should be reachable through port from which its traffic arrives.

Shouldn’t learn SA from erroneous frames, since SA may be errored, leading to later mis-direction of frames. When using cut-through, switch should defer learning until it has verified the frame CRC.

Ageing‡: Motivation

Even if the database could record information about all addresses in the catenet, it shouldn’t retain learned information indefinitely:

Station mobility: Mobile station may change location, unicasting to its old address could create unnecessary loss & the station won’t reveal its new location by responding.

Topological changes: links/bridges go down/up

Filtering

A bridge can filter frames based on multiple criteria:

If destination is known to be reachable through port from which frame arrived, then discard the frame (e.g. shared media LAN).

Custom filtering rules: May discard frames in order to enforce security/policy/etc, e.g.
- Wireless LAN: only forward traffic from known link layer addresses.
- Prevent specific users.
- Control propagation of sensitive traffic.
- Don’t forward traffic to certain multicast addresses used to disseminate Spanning Tree info (coming soon...).

Spanning Tree algorithm: Synopsis

Elect a root
For each link, elect a Designated Bridge that has the lowest cost to reach the root.
“Designated port” connects Designated Bridge to this link.
Each bridge identifies a Root Port that leads towards the root.
Bridges enable Root Port, but disable all other ports (except Designated ports)
Maintain the topology through periodic advertisements

‡ "a parameter known as 'ageing' time... (Actually, the 802 spec spells the parameter "ageing" because the editor of 802.1d is British and the country seems to have a surplus of vowels.)" [Perlman, p. 50]
Section outline

Out-of-band signalling: Advantages

**Higher speed:**
- For signalling: e.g. 56kb/s vs DTMF pulsing
- For payload: Separating signalling from simpler payload transfer may simplify payload processing
- Facilitates **signalling during the call**, rather than signalling only before the call. (assuming circuit-switching)
- Signalling and payload transfer **can evolve separately**, e.g. phone signalling advances
- Signalling & payload **paths can differ**, e.g. signals via databases not en-route to destination
- Easier to **ensure reliable transfer of signalling**, despite overload of payload, e.g. reduce misdelivery rate (reliable addressing of connection)

In-band signalling: Advantages

**Detect impairments on path** used for payload when signalling
- e.g. e2e signalling won’t succeed when payload can’t be transferred.
- Allows **piggybacking of payload** with initial signalling, reducing impact of delay
- e.g. TCP: SYN/ACK (signalling); unlikely to get through unless data segments can get through can carry payload (though data not delivered until 3-way handshake is complete)

SS7 network architecture

Quad = 2 mated pairs of STPs
- Network is hierarchical
SS7 protocol stack

<table>
<thead>
<tr>
<th>Application</th>
<th>Presentation</th>
<th>Session</th>
<th>Transport</th>
<th>Network</th>
<th>Data Link</th>
<th>Physical</th>
</tr>
</thead>
<tbody>
<tr>
<td>OMAP</td>
<td>AES-TCAP</td>
<td>ISUP</td>
<td>TUP</td>
<td>SCCP</td>
<td>MTP Level 1</td>
<td>MTP Level 2</td>
</tr>
</tbody>
</table>

User Parts

OMAP: Operations, Maintenance and Administration Part
e.g. to diagnose link problems & validate routing tables
ASEs: Application Service Elements

Network Services Part

Network Services Part is being replaced with alternative transport, e.g. ATM, SCTP over IP

ISUP call setup example

Left SSP selects an idle trunk to right SSP sends an Initial Address Message (IAM) to right SSP
2. Right SSP receives IAM, determines that target is available Sends ringing tone over trunk Returns Address Complete Message (ACM), indicating reverse path trunk Left switch connects caller to trunk, so they hear ringing

Network Services Part

User Parts

Backward ECN

**Backward ECN:** Switch that is, or is nearing, congestion sets this bit, assuming that dest of this frame may send frames through this switch.
Signal may experience less delay reaching source ⇒ faster response to congestion.
Signal may propagate when frames are being lost due to congestion There may not be any frames propagating to the source to carry the signal
Section outline

Minimising serialisation delays

Multiplexing onto a 2.048Mb/s output port

0.6ms

0.2ms

ATM cell structure: Key parts

Virtual Path Identifier

Virtual Channel Identifier

“Inter-ATM-user indication”, Congestion Experienced, OAM?

Cell Loss Priority

Header Error Correction – CRC-8

Virtual Channels and Paths

VCs identify individual flows of information between endpoints. Multitudes of flows through network core switches inefficient for core switches to handle individual flows.

e.g. 622Mb/s link carries 8,800 64kb/s voice channels

average call duration of 2 minutes \( \Rightarrow \) 150 connection establish/release requests per second = large signalling burden

\( \Rightarrow \) Virtual Path = aggregate of VCs

Commerce analogy:
core switches are like wholesalers, dealing in bulk (VPs)
edge switches are like retailers, dealing with individuals (VCs)

Resources may be assigned to the VP
good for predictable flows (set up VC without negotiating switch resources) – e.g. in core of network where individual variations annul one another may waste resources when flows vary
Switches translate identifiers

VP switching

VP & VC switching

VP switch/cross-connect

VCI 24
VCI 23
VCI 22
VCI 21

VCI
24
VCI
23
VCI
22
VCI
21

VPI 1
VPI 2
VPI 3
VPI 4
VPI 5

Switches translate identifiers

VP switch/cross-connect

VCI 24
VCI 23
VCI 22
VCI 21

VCI
24
VCI
23
VCI
22
VCI
21

VPI 1
VPI 2
VPI 3
VPI 4
VPI 5

MPLS terminology

Label Edge Router (LER) – pushes (+) / pops (-) labels
Label Switched Router (LSR) – switches using MPLS labels
Label Distribution Protocol – enables an upstream router (closer to source) to be allocated a label by a downstream router.
Label Switched Path (LSP) – similar to ATM VC/VP
Forwarding Equivalence Class (FEC): A set of packets that are all treated the same way by a router. Not necessarily by all routers – e.g. packets with different DAs may belong to same FEC in the midst of the network.
Conventional IP routers determine FEC at each hop, MPLS routers (LERs) determine FEC only at entry to domain.

Pushing & popping multiple labels

Transit network (middle): pushes on label (blue) as packets enter this label is used for routing within this network pops off label as packets depart
null label

RSVP and Diffserv

Covered only last week - not covered in review