NSWI090: **Computer Networks**

http://www.ksi.mff.cuni.cz/~svoboda/courses/202-NSWI090/

Lecture 7

Internetworking I

Martin Svoboda
svoboda@ksi.mff.cuni.cz

19. 4. 2021

Charles University, Faculty of Mathematics and Physics

Lecture Outline

Internetworking

- Motivation and objectives
- **Internetworking at `L1`, `L2`, and `L3`**
    - **Principles** and assumptions
    - Interconnection **devices**
    - Main **functions**
    - Features and consequences

Internetworking

Internetworking

- **Narrower** meaning
  - Practice of **interconnection of networks**
    - I.e., interconnection of whole networks at L3 using routers
  - Result of this process = **internetwork**
    - Alternatively also system of networks, internet, catenet
- **Broader** meaning
  - Practice of **interconnection of networks or parts thereof**
    - I.e., also building the internal structure of the individual involved networks at L1 and L2 layers

Internetworking

Ultimate **objective**

- Interconnection of a **set of end nodes** via passive and active **network elements** to enable their **mutual communication**

Different **points of view**

- **Bottom-up** (practical)
  - **Mutual composition of smaller units into larger ones**
    - I.e., how they should be defined and then interconnected
  - In order to achieve their coexistence and cooperation
- **Top-down** (logical)
  - **Decomposition of larger units into smaller ones**
    - I.e., how they should be divided and then interconnected
  - In order to attain certain desired properties and effect

Internetworking

Aspects to consider

- Tackling the **limited range of transmission media**
- **Optimization of data flows** and load balancing
- Definition of **access and other permissions**
- **Ensuring security** and protection against attacks
- Increasing overall **potential of network use**
- ...

Lower Layers

Lower layers and their tasks

- `L1`: **Physical** Layer
  - **Transmission** of individual **bits** via a given **physical medium**
- `L2`: **Data Link** Layer
  - Sending of **blocks** of data between network **interfaces** of particular **nodes** within a **local network**
- `L3`: **Network** Layer
  - **Routing and forwarding** of packets across an **internetwork** to the target **node** of the **final** intended recipient

Network Elements

Observation

- Internetworking at **different layers**…
  - Uses **different** **devices**, follows **different** **rules**, fulfills **different** **tasks**, supports **different** **protocols**, and so has **different** **properties** with **different** **consequences**

Types of **network elements**

- **Active** elements
  - Powered devices that **actively work with the transmitted data**
    - Buffer, route, forward or otherwise process at higher layers
    - Amplify and shape electrical signals at `L1`
  - E.g.: **repeater**, **switch**, **router**, …
    - Device names depend on layers they are used at
- **Passive** elements

Network Elements

Passive network elements

- Cables, connectors, splitters, sockets, …
- **Racks**
  - Standardized frame or enclosure for mounting various electronic equipment modules
- **Patch panels**
  - Device or unit with higher number of connectors allowing convenient and flexible interconnection of cables
    - E.g., RJ-45 registered jacks and twisted pairs for Ethernet
- **Structured cabling**
  - Systematic cabling within an administrative or other building
    - Using twisted pairs for computer networks as well as telephony
    - Installed in advance

Network Elements

Active network elements

- L1: **repeater**
  - **Amplification and shaping** of the transmitted signal
- L2: **bridge** or **switch**
  - **Filtering and forwarding** of frames within a local network
- L3: **router**
  - **Routing and forwarding** of packets between networks
  - Alternatively also L3 switch / L4 switch / L7 switch
- L7: **gateway**
  - Advanced functionality related to firewalls, NAT, …

Basic Terminology

Internetworking at L1

- Interconnects individual **end nodes** or groups of end nodes
  - Using **repeaters**
- Result = **segment**

Internetworking at L2

- Interconnects individual **segments**
  - Using **bridges** or **switches**
- Result = **network**

Internetworking at L3

- Interconnects individual **networks**
  - Using **routers** or other devices
- Result = **internetwork**

Internetworking at L1

Physical Layer

- Transmission of individual digital bits via analog unmodulated or modulated transmission through a given physical medium
  - Guided and unguided **physical transmission paths**
    - Metallic (twisted pairs, coaxial cables), optical (optical fibers)
    - Wireless
  - Various forms of **electromagnetic waves**
    - Electrical signals, light pulses, radio, infrared, or other waves
  - **Baseband** or **passband** transmissions

Important features

- **We do not understand the meaning** of transmitted data
  - All bits are treated equally and independently on each other
    - We cannot distinguish between them

Internetworking at L1

Internetworking **objectives**

- Increasing **range** (possible only to a limited extent)
- Physical interconnection and **branching**
  - Originally by **splitting coaxial cables** directly
    - Using Tee connectors (T-connectors) or splitters
  - Nowadays using **repeaters** as active network elements
    - Since direct branching is no longer possible in case of twisted pairs or optical fibers
  - Specifically in **Ethernet**, repeaters are, therefore, sometimes also referenced as **hubs** or, more precisely, **Ethernet hubs**
    - Since *hub* is just a generic name for a device that can be used at any layer for the purpose of physical / logical branching
    - E.g., L2 switches or L3 routers could also be viewed as kind of hubs in a broader sense (but they are not)

Repeater

Repeater = basically just a **digital amplifier and hub**

- Two structural designs
  - 2 ports only → increasing range and interconnection
  - 3 and more ports → interconnection and branching

Main function

- **Amplification and shaping** of the transmitted signal
  - Real-world physical transmission paths are never optimal
    - In terms of **attenuation**, **distortion**, **interference**, …
    - Impact of these phenomena needs to be compensated
  - **Received signal is recovered and instantly transmitted again**
    - At the hardware level (using **electronic circuits**)
  - There is **no buffer** that would allow to cache the incoming data

Basic Features

Direct consequences

- **Processing** of incoming data **cannot be deferred**
- **Latency is constant** and very small
  - Typically smaller than a bit period itself
  - Constant latency implies **zero jitter**
- **All ports must operate at the same rate**
  - We would not otherwise be able to compensate for mutual differences in such rates
- **Without congestion** possibility
  - No data is buffered, no decisions are to be made, …

Basic Features

Omnidirectional and neutral behavior

- **Received data must be propagated to all directions**
  - I.e., all the ports different from the incoming one
  - Simply because we cannot determine particular directions
    - Source and / or destination **HW addresses would be needed**
    - However, they are only accessible at L2, not L1
    - I.e., we have no idea about frame structure, header fields, etc.
- **All received data must actually be propagated**
  - Including **collisions** and L2 **broadcasts**
    - Because there is no way of even recognizing such situations
  - I.e., we have no other option than to **treat all bits equally**

Basic Features

Technological dependency

- **Repeaters** are always designed for a **particular technology**
  - More precisely, its particular variant, version, rate, …
- **Amplification and shaping would otherwise be impossible**
  - We must be aware of encoding specifics, bit interval lengths, or other characteristics of a given technology
    - So that we can produce the outgoing signal at all
  - Unfortunately, this **violates the principles of layered models**
    - In particular, a given layer should not depend on internal details of another layer, all the more not a higher one
- Nevertheless, **generic repeaters simply cannot exist**

Shared Capacity

All nodes in a segment share the same transmission capacity

- I.e., only two nodes can be communicating at a given time
  - More precisely, **only one node can be transmitting**
  - And so other nodes cannot engage in different communications
- **This holds even when** both the source and destination nodes are **separated by a repeater**
- What if multiple parallel communications were desired?
  - Different device than repeater would be needed
  - One, that would support targeted **filtering and forwarding**
    - So that the local communication is not further propagated
    - And the remote one is only forwarded to the right direction
  - However, **this is not possible at** `L1`
    - Simply because, once again, we are not aware HW addresses

Shared Capacity

Access methods in general

- Particular methods used at the MAC L2 sublayer to control the **interaction with the shared physical transmission medium**
  - **Exclusive access**
    - CSMA/CD in Ethernet
    - CSMA/CA in Wi-Fi
    - …
  - **Shared access**
    - CDMA or TDMA in mobile networks using multiplexing
    - …

Ethernet Collisions

CSMA/CD (**Carrier Sense Multiple Access with Collision Detection**)

- When we want to **start transmitting**…
  - We must make sure that the **shared medium** (MA) is **currently not in use** by anyone else (CS)
    - If it is, we wait until it is not (1-persistence)
    - If not (or no longer), we immediately start transmitting
- While transmitting, we must **detect potential** collisions (CD)
  - I.e., despite the CS step, another node or even several nodes could have independently started transmitting as well
  - If **collision is detected**…
    - We cease transmitting the originally intended data
    - And instead start transmitting a **special** **jam signal**
    - So that we help other nodes detecting the collision as well
    - After a random waiting time, we make another attempt

Ethernet Collisions

CSMA/CD (cont'd)

- **Collision domain** = segment
  - We must make sure that collisions can reach all the nodes ⇒
    - **Maximal segment size** must be limited
    - **Minimal frame size** must be introduced
  - **Repeaters must propagate collisions**
- **Collision window**
  - **Period of time during which collisions can appear**
    - Time needed until the signal propagates to the whole domain
- Example: **10 Mb/s Ethernet**: 5-4-3 rule
  - 5 parts, 4 repeaters, 3 inhabited parts
- Observations
  - CSMA/CD is **no longer needed in newer versions of Ethernet**
    - Just one node resides in a segment and full-duplex is possible

Communication Principles

Vertical / horizontal communication at `L1`

- **Sender** node
  - `L2` requests to send a frame to a given HW address
  - `L1` transmits its individual bits in a form of signals
- **Repeater** node
  - All received signals are amplified, shaped, and transmitted to all the remaining directions
- **Recipient** node
  - Received signals are interpreted as individual bits
  - Stream of these bits is provided to `L2`

Summary

Internetworking at L1

- **Segment** = one or more nodes connected by repeaters (if any)
  - Segment **size** is limited, transmission capacity is **shared**
- **Repeaters**
  - **Invisible** for the communicating nodes
  - **All incoming data is propagated to all directions**
    - Including **collisions** and L2 broadcasts
  - **No buffering**
    - **Small and constant latency**, zero jitter
    - Congestion is not possible

Conclusion

- All in all, not the most efficient form of internetworking
  - But the only possible at L1

Internetworking at L2

Data Link Layer

- Sending of **blocks** of data between network **interfaces** of particular **end nodes** within a **local network**
  - Each network interface is associated with its **hardware address**
    - Must be unique within a given network

Important assumptions

- **Illusion** provided to end nodes
  - All nodes are mutually visible and reachable
    - I.e., they can **communicate with each other directly**
- **Reality**
  - Internal network structure may be more complicated
    - I.e., there can be **multiple interconnected segments**
  - **End nodes are not aware of this structure**, though

Internetworking at L2

Internetworking **objective**

- **Forming internal network structure**
  - I.e., more sophisticated **range extension and interconnection**
    - Note that all nodes within the resulting network will have **IP addresses from the same range** at L3
- Data flow optimization, …

Available **devices**

- **Bridges** and **switches**
  - Very similar devices as for the main aspects
  - Yet different in many particular details

Main **functions**

- **Filtering** and **forwarding**
  - Required source / destination node hardware addresses

Filtering and Forwarding

Default behavior

- **Incoming blocks are forwarded to all remaining directions**
  - As if the **flooding** principle was applied
    - With all its advantages and disadvantages
    - I.e., not entirely efficient and loops must be treated
- Only necessary when no **topology information** is available

Filtering

- **Local communication** within a given segment can be **filtered**
  - I.e., will not be further forwarded

Forwarding

- **Remote communication** will only be forwarded to the **right direction**, i.e., not all the remaining ones

Filtering and Forwarding

Consequences

- **Overall transmission capacity can be used more efficiently**
  - Since capacity of non-involved segments remain unused and so available for other potential concurrent communications

Topology knowledge

- At least certain topology knowledge is necessary
  - **Reachability of nodes** via neighboring segments (ports)
- **Static** configuration provided by network administrators
- **Dynamic** techniques
  - Allow bridges / switches to be used as Plug&Play devices
  - **Backward Learning** in **Ethernet**
    - Loops are treated using Spanning Tree Algorithm (STA)
  - **Source Routing** in **Token Ring**

Communication Principles

Vertical / horizontal communication at L2

- **Sender** node
  - L3 requests to send an **IP datagram** to a given **HW address**
  - L2 frame is prepared using **encapsulation** and **framing**
    - **Source** address corresponds to the interface HW address
    - **Destination** address was requested and provided by L3
  - L1 is then requested to transmit the frame contents
- **Recipient** node
  - L2 frame is recognized from the stream received by L1
  - When the **destination HW address** corresponds
    - Frame is **unpacked** and its **payload** (IP datagram) given to L3
    - Note that **broadcast** and **multicast** addresses must also be accepted beside the standard single **unicast** address
  - Otherwise a given frame is ignored (thrown away)

Communication Principles

Vertical / horizontal communication at L2 (cont'd)

- **Bridge** / **switch** nodes
  - L2 frame is recognized from the stream received by L1
  - It is then processed using the **filtering and forwarding** rules
    - I.e., sent to a given output L1 port (if any) / or all of them
  - Unless this frame was intended for the bridge / switch itself
- Observation
  - **Bridges / switches** work in the so-called **promiscuous mode**
    - It means they capture and process all the incoming frames
  - **End nodes** work in the standard **non-promiscuous mode**
    - They only capture and process their frames
    - However, this behavior can be changed to allow packet sniffing

Buffering Mechanisms

Buffering mechanisms

- Allow to temporarily cache the incoming frames
  - So that they can actually be processed
    - Since filtering and forwarding require knowledge of addresses
    - And so at least a certain portion of headers must be received
- In fact, **each port** has its own **incoming / outgoing queue**
- **Two basic approaches** are possible
  - **Store&Forward**
    - Incoming frames are first fully received
    - Only then their processing is initiated
  - **Cut-Through**
    - Incoming frames are **processed** and possibly also **transmitted immediately** after the necessary frame headers are available
    - I.e., **without waiting** for the entire frame to be even received

Buffering Mechanisms

Store&Forward

- Advantages
  - **Segments with different rates** can be connected
    - However, still within one particular technology
    - Since frames themselves are kept untouched
  - **Damaged frames** are not further disseminated
- Disadvantages
  - **Higher latency**
    - Higher than time needed for frame contents transmission

Buffering Mechanisms

Cut-Through

- Advantages
  - Significantly **lower latency**
- Disadvantages
  - **Damaged frames** cannot be detected and stopped
    - Because **checksums** are usually placed at the end of frames
    - And transmission is started before their are fully received
  - **Segments with different rates** cannot be connected

Basic Features

Collisions are not propagated

- I.e., they are not disseminated out from the segment where they appeared
  - And so **traffic in other segments remains intact**
  - As well as the bridge / switch operation as a whole
- This is only possible because of **buffering**
  - In case a frame is intended to be delivered to a segment with a currently ongoing collision, its forwarding to this segment is simply **postponed until the collision ceases**

L2 **broadcast is propagated**

- Since its recipients are all the nodes within a given network

Network Segmentation

Network segmentation

- Decomposition of a given network into individual segments
- **Transmission capacity within a segment is shared**
  - Nodes must compete with each other to gain medium access
    - They may even not be successful at all
  - Anyway, **the more nodes in a segment**, **the higher the probability of collisions**
- Possible solutions
  - **Single large segment**
    - All network nodes reside only inside a single segment
    - I.e., there are no bridges nor switches
  - Microsegmentation
    - **Each segment contains only a single end node**
  - Of course, any solution between these two is possible as well

Network Segmentation

Microsegmentation

- There is **no longer any competition** inside any segment
  - Under the assumption that **full-duplex is possible**
  - Available segment capacity is dedicated solely for a given node
- Brings the effect of **exclusive transmission capacity**
  - **Each segment can engage in its own communication**
    - Local inside a given segment (enabled by filtering)
    - Remote between a pair of segments (enabled by forwarding)
  - I.e., multiple communications can be in progress at any time
- Necessary condition: **sufficient transmission capacity**
  - Only possible in case of switches, not bridges
    - More precisely, **non-blocking switches**
  - Internal computation capacity must correspond to the sum of transmission capacities of all the segments
    - Otherwise such a switch would represent a bottleneck

Bridges and Switches

Bridge

- **Older** kind of device
  - Almost no longer used nowadays
- Optimized for **filtering**
  - Even though forwarding is also supported
- Usually **lower number of ports** (even just 2)
  - And so intended for **lower number of usually larger segments**
    - Where local traffic prevails over the remote one
  - $\Rightarrow$ bridge is supposed to **separate**
- Can be implemented at the **software level**
  - Since filtering is not that demanding
  - And internal speed is not that important

Bridges and Switches

Switch

- **Newer** kind of device
  - And significantly more **complex**
- Optimized for **forwarding**
  - Filtering is, of course, also supported, but it may happen that it will actually not get a chance to be exploited
- Usually **higher number of ports** (even up to around 50)
  - And so for **higher number of usually smaller segments**
    - Even with always just a single node (**microsegmentation**)
    - I.e., supports the concept of **exclusive capacity** creation
  - $\Rightarrow$ switch is supposed to **connect**
- Implemented at the **hardware level** using electronic circuits
  - Since internal **speed is crucial**

Summary

Internetworking at L2

- **Network** = one or more interconnected segments
  - Network **size** is not directly limited
  - Transmission capacity is **shared** only inside a segment
- **Bridges / switches**
  - Still **invisible** for the communicating nodes
  - **Incoming frames are buffered**
    - **Higher and variable latency**, non-zero jitter
    - Congestion is possible
  - **Filtering** and targeted **forwarding**
  - **Collisions** are not propagated
  - L2 broadcasts are propagated

Internetworking at L3

Network Layer

- Delivery of packets across a **system of interconnected networks** to the target **node** of the **final** recipient

Important assumptions

- We are aware of the **existence of multiple networks** as well as the way they are **mutually interconnected**
  - Or at least to a certain extent
  - Even the sender itself must think about the first steps of routing
- Packets are delivered through individual **routers**, one by one

Internetworking at L3

Internetworking **objectives**

- **Interconnection of individual networks**
- Definition of access and other permissions
- Limitation of broadcast domains
- …

Available **devices**

- **Router**
- Alternatively also L3 switch / L4 switch / L7 switch

Main **functions**

- **Routing** and **forwarding**

Communication Principles

Vertical / horizontal communication at L3

- **Sender** node
  - L4 requests to send a block of data to a given **IP address**
    - I.e., **TCP segment** / **UDP datagram**
  - **Routing** (forwarding) **tables** are consulted
    - So that **local interface** is resolved in case of **direct delivery**
    - And both **local interface** and **gateway** (first-hop router) in our network is resolved otherwise (in case of **indirect delivery**)
  - **IP datagram** is prepared using **encapsulation**
    - IP address of the local interface is used as the **source** address
    - IP address of the final recipient is used as the **destination**
  - **HW address** of the L2 **local recipient** is resolved
    - Final node / first-hop router in case of direct / indirect delivery
  - Selected L2 interface is requested to send the IP datagram

Communication Principles

Vertical / horizontal communication at L3 (cont'd)

- **Recipient** node
  - IP datagram is unpacked from the received frame at L2
  - When the **destination IP address** corresponds
    - Datagram is **unpacked** and its **payload** (TCP / UDP) given to L4
    - Note that **broadcast** and **multicast** addresses must also be accepted beside the standard **unicast** address / addresses
  - Otherwise a given datagram is ignored (thrown away)
- **Router** node
  - IP datagram is unpacked from the received frame at L2
  - It is then processed using the **routing and forwarding** rules
    - I.e., sent to a given L2 interface (if any)
    - This interface will create its own frame to be sent
  - Unless this datagram was intended for the router itself

Summary

Internetworking at L3

- **Internetwork** = one or more interconnected networks
- **Routers**
  - **Visible** for the communicating nodes
  - Incoming datagrams are **buffered**
    - Higher and variable latency, non-zero jitter
    - Congestion is possible
  - Collisions are not propagated
  - L2 **broadcasts are not propagated** as well

Lecture Conclusion

Internetworking at L1

- Segment
- **Repeaters**: amplification and shaping
- Collisions

Internetworking at L2

- Network
- **Bridges** and **switches**: filtering and forwarding
- Microsegmentation

Internetworking at L3

- System of networks
- **Routers**: routing and forwarding