NSWI090: **Computer Networks**

http://www.ksi.mff.cuni.cz/~svoboda/courses/222-NSWI090/

Lecture 7

Internetworking

Martin Svoboda
martin.svoboda@matfyz.cuni.cz

2023

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
- **Broadcast** domains
- **Virtual LANs**
  - Motivation and deployment
- **Firewalls**

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
  - ⇒ 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

Terminology Overview

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

Internetworking Principles

80/20 rule

- Traditionally…
  - Usually ≈ **80% of traffic was local** within a given network
  - And only ≈ 20% was leaving such a network

20/80 rule

- Things significantly changed with the Internet…
  - Usually only ≈ 20% is still local
  - Even ≈ **80% of traffic crosses the border of a local network**
- Routers may no longer be able to handle increasing data flows
- **Solutions**
  - **Virtual Local Area Networks** (**VLAN**)
    - Harness fast interconnection at `L2`, but limit broadcast domains
  - `L3` **Switches**
    - Increase overall efficiency and throughput of traditional routers

Broadcast Transmissions

L2 broadcast

- Intended **recipients**
  - **All nodes within our local network** = **broadcast domain**
    - I.e., all nodes residing in the same network as the sender node
- Frame **destination address**
  - `FF:FF:FF:FF:FF:FF`
    - Special address with binary ones only
- **Delivery** process
  - **Bridges and switches**: forwarding based on **flooding**
  - **Routers** (in our network): further propagation is stopped
- Natural motivation
  - **Limiting the size** of broadcast domains

Broadcast Transmissions

Local L3 broadcast

- Intended **recipients**
  - Once again, **all nodes within a given local network**
    - Only this time in the context of IP datagrams at L3
- Datagram **destination address**
  - `255.255.255.255`
    - Once again special address with binary ones only
- **Delivery** process
  - **Sender**: IP datagram is requested to be sent using L2 **broadcast**
  - **Routers** (in our network): further propagation is stopped

Broadcast Transmissions

Targeted L3 broadcast (**Directed L3 broadcast**)

- Intended **recipients**
  - **All nodes within a given particular network**
    - Usually **foreign network** (but also works for the local one)
- Datagram **destination address**
  - E.g.: `192.168.1.255`
    - **Network prefix at the beginning**, binary ones at the end
- **Delivery** process
  - IP datagram is first **routed and forwarded** using **standard unicast delivery**
  - Once the **router** serving as the **entry point to the target network** is reached, local L2 **broadcast** is then utilized
- Security considerations
  - Incoming targeted broadcasts are usually ignored nowadays

Network Layer Devices

Possible alternatives for L3 interconnection devices

- **Router**
  - **Traditional** complex device allowing for **routing and forwarding**
  - Suitable for **transition** between **heterogeneous** environments
- **L3 Switch**
  - Newer **integrated device** combining L2 **and** L3 **functionality**
    - Standard L2 switch for local network
    - Simplified but more efficient L3 router
  - Suitable for **interconnection** of **homogeneous** environments
- **Multilayer switch**
  - Basically L3 switch allowing to take into account information from higher layers L4 and / or even L7 for **routing decisions**
    - In particular, **L4 Switch** and **L7 Switch**

Network Layer Devices

Router

- Optimized for **logical functions** (and not only the core ones)
  - **Routing and forwarding**
  - Network Address Translation (**NAT**)
    - Allows to use **private IP addresses** in private networks
  - Assignment of IP addresses (**DHCP**)
  - Security: **firewall**, access rights, …
  - **Monitoring**, **management**, …
  - …
- Speed and throughput are not critical
  - As router was originally designed for 80:20 environments
  - **Implemented at the software level**
    - On top of a dedicated operating system (Cisco IOS)

Network Layer Devices

Router (cont'd)

- Suitable for transition between heterogeneous environments
  - **Bigger routing tables**
  - Usually bigger buffers
  - Can have **physical interfaces** with **different technologies**
    - Ethernet, EuroDOCSIS, xDSL, SDH, …
  - Can support **multiple routing protocols**
- Used for **connection to other networks**
  - Usually smaller networks (LAN, MAN) to larger ones (WAN)
  - Emphasis is put on…
    - Adaptation, logical separation, correct decision-making, …

Network Layer Devices

L3 Switch

- Optimized for **speed and throughput**
  - As L3 switch was originally designed for 20:80 environments
  - **Implemented at the hardware level**
    - So that it can match the wire speed
  - Focuses only on the **core functionality**
    - I.e., routing and forwarding
- Suitable for interconnection of homogeneous environments
  - Usually smaller routing tables and smaller buffers
  - Usually **Ethernet** physical interfaces only
- Used for **interconnection of related networks** (LAN, MAN)
  - Also allows to limit broadcast domains
    - Analogously to routers, but more efficiently

L4 and L7 Switches

L4 Switch

- L3 switch which can take L4 **information** into account
  - I.e., **routing decisions** can also be based on…
    - Transport **protocols** (TCP, UDP, …) and / or **port numbers**
- Different kinds of traffic can thus be treated differently
  - E.g., port 80 (HTTP requests), port 53 (DNS queries), …

L7 Switch (Content Switch)

- L3 switch which can take L4+L7 **information** into account
  - I.e., **routing decisions** can also be based on L4 and…
    - Application **protocols** (HTTP, SMTP, …) and their data
- Analogous utilization as above
  - E.g., port 80 HTTP requests to specific URLs in GET headers, …

L4 and L7 Switches

Use cases: **diversified routing**

- **Distribution** of requests
  - Requests to different services (e.g., HTTP, FTP, …) are in fact forwarded to different servers each providing just one of them
- Simulation of **anycast** transmissions
  - Requests to the same service are in fact split between multiple standalone serves (stickiness may be required)
- **Load balancing**
  - Exploitation of more different routing paths
- **Transparent caching**
  - HTTP requests are redirected to a dedicated cache server
- **Redirection** of DNS queries
- …

L4 and L7 Switches

Use cases: **traffic management**

- Traffic **prioritization**
  - Multimedia data may be handled preferentially
- Traffic **blocking**
  - Certain kinds of traffic may be strictly prohibited
    - E.g., VoIP communication, …
- Traffic **limitation**
  - Introduction of **volume quotas** for various kinds of traffic
    - E.g., **Fair Use Policy** (**FUP**)

Virtual Local Area Networks

Motivation

- L3 **network** = **set of end nodes** residing in one or more L2 segments interconnected using bridges / switches
  - All involved nodes are **mutually visible** and **directly reachable**
    - And so all L2 **traffic is also visible** to the entire network
  - This is not always desirable
    - Especially in buildings with systematic cabling deployed
    - Since individual users (end nodes) may not be related at all
- And so what if **membership of end nodes to networks** would be determined differently?
  - I.e., **independently on physical locations**
  - Separate switches and physical rewiring could then help
    - But this approach is not flexible enough
  - And so the concept of **VLAN** was introduced

Virtual Local Area Networks

VLAN (Virtual LAN)

- Principle: coexistence of **multiple different virtual networks on top of one physical `L1+L2` infrastructure**
  - Allows to decouple…
    - **Physical users locations** from **logical network memberships**
  - And so individual VLANs can reflect different…
    - Organizational needs, groups or categories of users, access or other privileges, usage of services and servers, …
- Whole concept is generic
  - Both older proprietary and newer standardized solutions exist
  - **Implemented in several technologies**
    - **Ethernet**, ATM, …

VLAN Principles

Requirements

- **Additional logic** needs to be added into the infrastructure
  - Primarily **VLAN-aware** **switches** at L2
  - But also **routers** at L3
- Practical expectations
  - **End nodes should remain ignorant** to the whole concept
    - I.e., they should not need to know what VLAN they are part of, nor whether VLANs are being deployed and utilized at all
    - Thus their interfaces / software do not need to be upgraded
  - $\Rightarrow$ only network administrators should concern themselves
- Fundamental requirement
  - **Traffic belonging to a given VLAN** must stay within that VLAN
    - I.e., it must be guaranteed that it will not leak to a different one
    - And so **VLAN hopping** must be avoided

VLAN Principles

Consequences and features

- Limiting **broadcast domains**
  - Broadcasts and unknown unicasts are flooded everywhere
- Improving **security and privacy**, minimizing external threats
- Enabling **Quality of Service**
  - Kind of VLAN side-effect, based on traffic prioritizing
- Simplifying **network administration** and **fault management**

VLAN concepts

- Two basic types of virtual networks can be distinguished
  - **Local VLANs** and **End-to-End VLANs**
- They both differ in the primary motivation and objectives
  - However, their mutual **boundaries are not defined strictly**

VLAN Concepts

Local VLANs

- Aim at **separating geographically close nodes**
  - In the reach of just one switch (or a small group of switches)
  - This allows for easier implementation of the whole concept
- Primary goal: **limiting broadcast domains**

End-to-End VLANs

- More generic concept
- Aim at **interconnecting geographically remote nodes**
  - Individual nodes are dispersed throughout the whole network
  - And so **VLANs span multiple switches** across the network
    - Special links between the switches are therefore needed
    - So that they can carry traffic of several different VLANs at a time
- Primary goal: **grouping users with similar interests**

Logical Model

Set of VLANs, each associated with…

- Distinct **integer** VLAN Identifier (**VID**)
- Optional **name** allowing for user-friendly management

Types of **segments** involved in the infrastructure

- VLAN-unaware **segments**
  - Contain nodes from **exactly one VLAN**
    - Actually just a single node in case of microsegmentation
    - Transmitted frames do not need to be mutually distinguished
  - Correspond to **switch-to-host** links
- VLAN-aware **segments**
  - Carry traffic from **several different VLANs**
    - And so such frames must be tagged to be mutually recognizable
  - Correspond to **switch-to-switch** or **switch-to-router** links

Logical Model

Operation principles

- VLAN can actually be seen simply as kind of a projected network consisting of only segments where it is activated
  - From this point of view, everything works as expected
  - I.e., **filtering and forwarding**
    - Including Spanning Tree Protocol (STP), etc.

VLAN configuration

- Expressed via **association of switch ports to VLANs**
  - I.e., not directly in terms of the intended usage of segments
- In particular, **each port** is labeled with a **set of permitted VIDs**
  - Obviously, network administrator must ensure **consistency**
    - I.e., corresponding ports on switches containing a given segment must be configured identically

Types of Ports

Access port (untagged port)

- Connects a VLAN-unaware segment
  - Labeled with **exactly one VID**
    - If not specified, **default VLAN** is assumed (usually VID 1)
  - This very VID **determines the VLAN membership** of nodes
- All frames (are expected to) belong to this single VLAN
  - **Incoming frame** is altered by **tagging** it with a given **port VID**
    - So that it becomes prepared to enter VLAN-aware segments
    - Already tagged frame is only accepted if it matches the port VID
  - **Outgoing frame** is altered by removing its tag
- **Tagging mechanism** is required
  - Open standard **IEEE 802.1q** (**Dot1q**)
  - Proprietary approaches: Cisco ISL (Inter-Switch Link), …

Types of Ports

Trunk port (tagged port)

- Connects a VLAN-aware segment
  - Labeled with **one or more VIDs**
    - By default, **all VLANs**
    - Or enumeration of only **selected VLANs**
- Frames of all involved VLANs are carried alongside each other
  - And so they must be **tagged so that they can be distinguished**
  - **Incoming frame** is only accepted if it matches the allowed VIDs
- **Native VLAN** may optionally be specified
  - Its frames may **remain untagged**
    - This allows to have VLAN-unaware devices in trunks as well
  - Configured on a per-port and per-device basis
    - Must hence be consistent within the entire trunk segment
    - Typically the same value everywhere (for sanity)

VLAN Configuration

Static (port-based) approaches

- Each port is **configured manually** by network administrator
- Relatively small overhead, higher security, not flexible enough

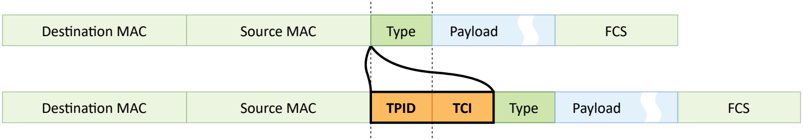
Dynamic approaches

- **VLAN membership** is resolved dynamically
  - Based on **MAC addresses** (deprecated, not a good idea anyway)
  - Or **IEEE 802.1X authentication** (based on user credentials)
- Information needs to be shared between switches
  - **Multiple VLAN Registration Protocol** (**MVRP**) (IEEE 802.1ak)
    - L2 protocol allowing to de/registers VIDs on ports, …
  - Proprietary approaches: Cisco VTP (VLAN Trunking Protocol)
- Greatly **simplifies network design and deployment**

Ethernet Frames

IEEE 802.1q (Dot1q **tagging**)

- **VLAN tag** is added into the original **Ethernet frames**
  - Between Source MAC and Type / Length header fields
  - TPID = Tag Protocol Identifier = `0x8100`
    - So that tagged frames can be distinguished from untagged ones
  - TCI = Tag Control Information
    - Contains 12-bit long **VLAN Identifier** (**VID**) ≈ **4094 VLANs**
    - Certain values are reserved (at least `0x000` and `0xFFF`)
- Adding and removing tags also involves recalculating the CRC



Routing Between VLANs

Observation

- **IP traffic between VLANs must normally go through routers**

Routing options

- **VLAN-unaware router** with **separate physical interfaces**
  - One separate port is needed for each VLAN on the router
  - They are all connected to different **access ports** on a switch
  - Obviously working, but not efficient enough and scales poorly
- **VLAN-aware router** with **sub-interfaces**
  - Physical interface is split up into multiple virtual sub-interfaces
    - Each corresponds to one particular VLAN
    - Frames outgoing from the router are tagged appropriately
  - Connected to a **trunk port** on a switch
- **VLAN-aware L3 switch**

Firewalls

Firewall

- General security system permitting to **monitor and control** both **incoming and outgoing** network traffic
  - Allows to **block unauthorized** / **allow authorized** access
    - So that users (their traffic) can only get where they are allowed
- Forms a **barrier** between a **trusted** and an **untrusted network**
  - I.e., between the inner (LAN) and outer (Internet) networks

Firewalls

Possible **deployments**

- **Network-based firewall**
  - Protects the whole inner corporate / school / home network
    - And so all its nodes / users
- **Host-based firewall** (individual, personal)
  - Protects just a single node / user

Possible **implementations**

- **Dedicated device** (combination of hardware and software)
- Purely **software solution**
- Set of **organizational measures**

Firewalls

Possible **strategies**

- **Prohibited unless permitted**
  - Everything is by default prohibited
  - Only something is explicitly permitted via **positive exceptions**
    - Having the nature of permissions
  - Approaches
    - Demilitarized Zones, Packet Filters
- **Permitted unless prohibited**
  - Everything is by default permitted
  - Only something is explicitly prohibited via **negative exceptions**
    - Having the nature of prohibitions
  - Approach
    - Packet Filters

Demilitarized Zones

Demilitarized Zone (**DMZ**) (Perimeter Network)

- Physical or logical **network acting as a barrier separating the inner and outer networks / zones**
  - Serves as kind of a gateway to the public Internet
    - Neither as secure as the inner zone,
      nor as insecure as the outer zone
  - Provides additional security especially from external attacks
- Permitted traffic
  - **Outer zone** ↔ **inner zone**
    - This kind of communication is **entirely prohibited**
    - I.e., **no traffic can directly pass through DMZ**
  - **Outer zone** ↔ **DMZ** and **DMZ** ↔ **inner zone**
    - Possible in principle
    - But can also be partially restricted if need be

Demilitarized Zones

Demilitarized Zone (cont'd)

- Means of **implementation**
  - Simply via appropriate **configuration of routing tables** in both the routers separating the zones (i.e., at L3)
    - **Only traffic commencing / terminating in DMZ is allowed**
    - Which is detectable using **source / destination IP addresses**
- **DMZ contains…**
  - **Public servers** providing services to external users
    - E.g.: HTTP, SMTP, POP3, DNS, …
    - These are the hosts that are most **vulnerable to attacks**
    - And so when any of them gets compromised, inner zone is still likely to remain protected
  - **Application Gateways**
    - **Mediate** otherwise impossible **outer** $\leftrightarrow$ **inner communication**

Demilitarized Zones

Application Gateway (L7 Gateway, Application Proxy)

- **Server mediating communication with the outer zone**
  - E.g.: HTTP Proxy Gateway for requesting web pages, …
- Principle
  - (1) Inner node sends an intermediate request to the gateway
    - I.e., not directly to the intended target node
    - And so the sender must be aware of the gateway existence!
    - ⇒ application **gateways are not transparent**
  - (2) Gateway then generates and sends its own request
  - (3) Response from the target node is received by the gateway
  - (4) It is then forwarded to the original node in the inner zone
- Observation
  - Gateways are always **application-dependent**
    - I.e., specifically designed for a given particular L7 protocol

DMZ Architectures

Dual Firewalls (**Back-to-Back DMZ**)

- **Two routers** (firewalls) are needed
  - **Front-end** (**perimeter**) between the outer zone and DMZ
  - **Back-end** (**internal**) between DMZ and the inner zone
- **Higher security**
  - Because two devices would need to be compromised at a time
    - Especially when devices from **different vendors** are used
    - Since it is not likely they would have the same vulnerabilities
- **Relatively costly** solution
  - And so suitable only for larger corporate networks

DMZ Architectures

Single Firewall (**Three-Legged DMZ**)

- **Only one router** (firewall) with (at least) **3 network interfaces**
- Represents a single point of failure
  - Since it must be able to handle all of the traffic

Integrated DMZ

- DMZ on a **software basis** without even a single router device
  - I.e., within a node directly separating the outer / inner zones

DMZ Host – not a true DMZ!

- Solution frequently appearing in **small home routers**
  - One server in the inner network can be specified
    - It then receives all unrecognized incoming traffic
  - This server is not isolated from the inner network at all
    - And so this solution has nothing to do with the DMZ concept

Packet Filters

Packet Filter

- Inspects and **filters** both **incoming and outgoing traffic** based on a set of configured **rules**
  - Works at L3
    - In terms of both blocking and permitting
    - In contrast, DMZ blocks at L3 and permits at L7
- Both **positive and negative** strategies are possible
  - Individual **rules** are described via **Access Control Lists**
- Available information
  - **Source / destination IP addresses** by default
  - But also information from higher layers
    - Such as **transport protocols** or **port numbers** at L4, …

Packet Filters

Modes of operation

- **Stateless Packet Inspection** (**Static Packet Filtering**)
  - Each packet is treated **independently** on each other
  - Easier to implement
  - Less computationally demanding
- **Stateful Packet Inspection** (**Dynamic Packet Filtering**)
  - Each packet is treated with regard to the recent **history**
    - I.e., also with respected to the previously handled packets
  - And so **more undesirable situations can be detected**
    - Especially various **concurrencies**
    - Can help to prevent DOS / DDOS attacks

Packet Filters

Access Control List (**ACL**)

- **List of rules** to be applied
  - Based on **positive permissions** or **negative exceptions**
- Standard **ACL**
  - Only **source IP address** is considered
  - Recommended deployment
    - Usually as **close to the target nodes** as possible
- Extended **ACL**
  - Other information is considered as well
    - **Destination IP address**, port number, …
  - Recommended deployment
    - Usually as **close to the source nodes** as possible

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

Lecture Conclusion

Broadcasts

- L2, local L3, targeted L3
- Broadcast domains

L3 **interconnection devices**

- Routers, L3 / L4 / L7 switches

VLANs

- VLAN-aware / VLAN-unaware segments
- Access (untagged) / trunk (tagged) ports
- Static / dynamic configuration

Firewalls

- **Demilitarized zones**, application gateways, **packet filters**