

COMP28411 Computer Networks

Nick Filer - Link Layer and Physical Layer

Monday 28/11 LL - 1 of 4 - Introduction

Tuesday 29/11 LL - 2 - Switches, Routing and

Medium Access

Andy still has 1 lecture to give, maybe 13/12.

1

Current PLAN!

Tuesday 5/12 LL - 3 - Framing, Ethernet, ATM Point-2-Point Protocol

Monday 12/12 PL - 4 of 4 - Physical Layer and Wireless

Tuesday 13/12 - Exam Question & Answer

5 and 6/12 - Error Detection Workshop

Some material from:

Kurose & Rose - Chapter 5 + Slides

24/11/2016 COMP28411 Link Layer



Lecture Summary

- · Link Layer
 - Services
 - · Where implemented
 - Packet Encapsulation reminder
 - Flow Control
 - · Link layer addressing
 - Mapping IP to/from MAC addresses.
 - · Hubs and Switches

24/11/2016

COMP28411 Link Layer

2



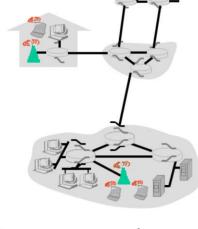
Link Layer Overview

Terminology

- Hosts and Routers/Switches/Hubs are NODES.
- Communications channel between nodes is a LINK.
- Layer 2 packets are a **FRAME**

Responsibility

- Node to adjacent node transfer of layer 3 datagram over link.
- Uses Layer 1 (Physical) methods to move frames between directly interconnected hosts.



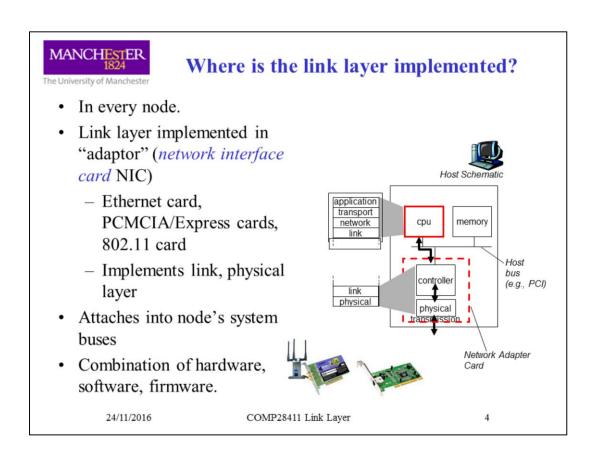
24/11/2016

COMP28411 Link Layer

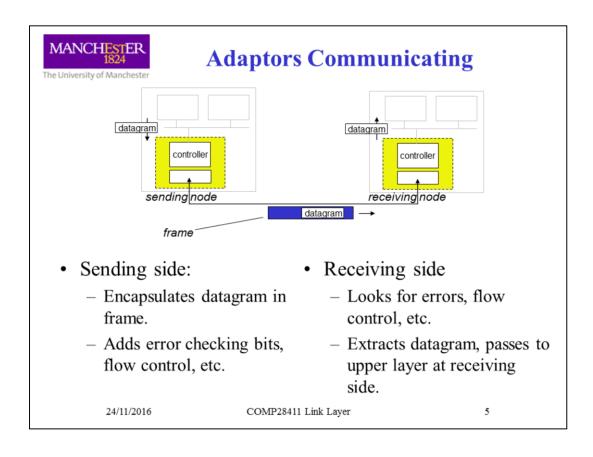
3

I'm not very good at sticking to the right terminology. In networking, we often call any unit of data sent between two devices a packet and use the term ambiguously to mean lots of different things that are things we send between devices.

At the link layer we are looking at **a graph of nodes** connected via **links.** In between nodes, down links we send frames, these frames encapsulate datagrams which may be UDP or TCP or some other layer 3 or above protocol's data.



In most devices, the link layer is implemented in a network adaptor or Network Interface Card (NIC). There are lots of different types of NIC. Each NIC will have its own on-board mini computer with a processor (often several of these now to support normal and digital signal processing tasks), memory for buffering incoming, outgoing and transient data. The NIC will have at least two interfaces, one to the host computer or device and the other to the network's physical layer.



The NIC when sending data receives the data from the network layer encapsulated as a datagram which it in turn encapsulates into a frame with its own header and tail as required by the link layer protocol being used. The frame is then transferred to the receiving NIC via the physical layer where the frame's headers and tail are stripped and checked. Normally, if the address in the frame matches the address the NIC knows it will then forward the datagram to the host node. If, for example the frame address does not match then the frame is normally thrown away. However, some NICs allow something called 'permissive mode' whereby all frames whoever they are for are sent to the host node for processing.



Virtual Links



Alternatives:

- My computer talks Ethernet (Layer 2 + 1).
- My laptop talks either Ethernet, WiFi (Layer 2 + 1) or Bluetooth.
- My Broadband (Last Mile Connection over telephone or cable) has:
 - IP input over Ethernet or WiFi. Both IEEE 802.xx standards.
 - ADSL2 physical layer with Asynchronous Transfer Mode (ATM) at the data-link layer.
 - VDSL (very high bit rate DSL) over fibre. Often PPPoE.
 - On ADSL the ATM (layer 2) carries Point to Point Protocol over ATM (PPPoA) – also layer 2!.
 - Outside Europe many use PPPoE (Ethernet)
 - · The PPP (layer 2) carries IP packets.
- We regard the PPPoA/E as providing a virtual link.
 - To higher layers it appears as a complete network.

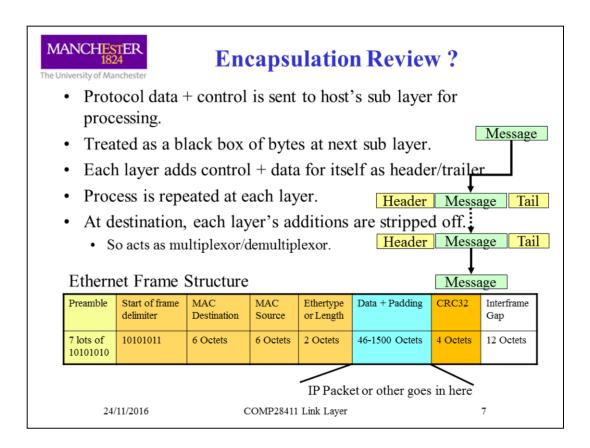
 24/11/2016 COMP28411 Link Layer

6

To make our lives as people studying networks worse there are also virtual links in many networks. Virtual links allow the network providers to layer one set of network links on top of another. A good example is a typical (as of 2015 still true) last mile connection between someone's home and their telephone/internet provider. Most last mile links in the UK use copper cables as the physical layer. These cables have their own physical layer protocols used for carrying mixtures of digital and analogue data i.e. Telephone calls which are often still analogue until they reach the telephone exchange and digital data encoded into analogue signals as voltages/current varying over time (frequency). The telephone cable could carry a wide variety of link layer frame types but mostly they use very small (48 bytes data + 5 header bytes = 53 bytes) Asynchronous Transfer Mode (ATM) frames carrying Point to Point Protocol (PPP) frames which themselves carry our IP or other datagrams. An alternative to ATM which is gaining in usage as connections get faster is PPP over Ethernet as the last mile link layer instead of ATM. Ethernet is more efficient and the frames are short enough once the data rates exceed about 622MB.

So why did I say "short enough" for the Ethernet frames. Carriers chose ATM because it has very small frames that include a channel identifier which allows multiple virtual channels (look like connections or single channels to end users) of data to be mixed up on the cable. This allowed, for example, real-time data to be easily mixed with best-effort data with very little extra jitter added. Typical 1,500 byte Ethernet frames would have added too much jitter at low data rates.

Another factor was the pervasive use of ATM within core networks and the fact that it easily allows cable sharing between multiple ISPs.



This should be simple to follow. As data drops through the layers of a network it gets encapsulated over and over again like the structure of an onion. Then as the same data rises up through the layers in a node it gets de-encapsulated layer by layer.

As an example of encapsulation an Ethernet frame is shown. The blue data portion might carry anything but often it carries an IP datagram.

To make the Ethernet work in its physical environment each frame starts with a preamble (why? We will cover this later), then a fixed bit pattern start of frame delimiter. There are two addresses, each 48 bits or 6 bytes in length which we know as MAC addresses where MAC stands for Medium Access Control. These addresses identify the end points of the frames journey through the local network.

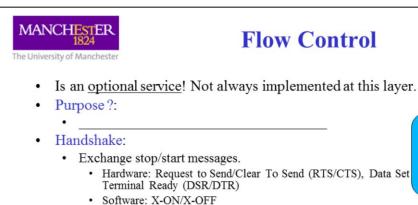
Because frames can vary in length there must be some way to work out the length of the data portion if the encapsulated protocol allows varying length. Note there are 16 bits here so in theory length might be as large as 65K. But these bits are shared with a type field. In 1997 it was agreed that values above 1536 (0x0600) would be used for types and below for lengths. 1,500 = 0x05DC, which leaves values between 1,501 and 1,535 undefined and unused. For Ethernet jumbo frames a special Ethertype value of 0x8870 is used. There is an incomplete list of Ethertype assignments for protocols see Wikipedia "Ethertype" for details. So Ethernet uses a length field for some data protocols but not all.

Other protocols might have fixed lengths or move the length field to the tail or use a data delimiter and count the size. A number of length determination schemes are used in other protocols.

The CRC field allows the correctness of the data to be almost certainly (probability) checked. There is a specified gap between frames of 96 bits in size.

MANCHESTER The University of Manchester Why is an Ethernet Frame like it is? Why Preamble at start of frame? Why is destination address near start? Why is length given at start? Why is length given at start? Why is CRC at the end? Why are other fields e.g. Modulation, speed... Missing? COMP28411 Link Layer 9

You need to either know or scribble your answers to these as I talk!



Flow Control

e.g. for realtime constant

10

bit rate traffic.

A2

- Open-Flow:
 - By prior reservation using Connection Admission Control (CAC)
 - To work must over resource redundancy.
- Closed-Loop:
 - Some way of reporting resource availability and resource needs.
 - Ethernet PAUSE frame sent to special multicast address 01-80-C2-00-00-01 with 16 bit time request in 512 bit time quanta's.
 - Asynchronous Transfer Mode (ATM) has Available Bit Rate (ABR) guarantees minimum bit rate and reports congestion.

24/11/2016 COMP28411 Link Layer

Flow control is all about making sure that destinations do not become overloaded with arriving traffic. If traffic arrives whilst the destination is busy processing the previous arrival it may not receive the data correctly or may simply treat it as garbage. There are various hardware and software ways to set up flow control. You have already seen an example of end-to-end software flow control in TCP!

A simple method is to have explicit signals reserved for flow start and stop. Serial RS232 cables used to have extra wires between two half-duplex (single direction at time) devices so hardware could set these lines to indicate that the sender has data to send and that the receiver is ready to process the arriving data (RTS/CTS). WiFi then adapted this scheme to deal with a problem called the 'hidden node problem' in wireless and other broadcast technologies. An RTS frame tell the destination that data is ready. A CTS frame tells the sender to send the data. But, at the same time the RTS and CTS tell anybody else who hears them to be quite for the expected duration of the data transaction. This helps to avoid collisions caused by other senders who may not be able to hear the data that is currently on air because they are too far away from the sender but which are close enough to the receiver to cause destructive interference if they transmit whilst the receiver is listening to the data from the original sender. The DRS and DTR signals were separate lines used for modem control intended to indicate the modem was in or not in use, some printers used it for hardware flow control back to computers. Serial, RS232 mouse required power, the RTS and DTR lines were often held high by mouse drivers to power otherwise unpowered mice.

Open-Flow is a protocol used as part of software Defined Networks (SDN) in order to traffic engineer packet routes through networks. To prevent congestion, traffic is flow controlled at network entry points before congestion occurs. Admission Control decides whether to accept of reject a connection based on its QoS requirements and the current load on the network as measured from varying traffic delays and loss probabilities. CAC is a QoS tool. Monitors demand and allocates resources. < continued on slide 11>



LAN/MAC Addresses

• 32-bit IP address:



- Network-layer address.
- · Used to get datagram to destination IP subnet.
- At least partially geographical and part ID.
- MAC (or LAN or physical or Ethernet) address:
 - Function: Get frame from one interface to another physically-connected interface (on the same network).
 - 48 bit (6 byte) MAC address (for most LANs)
 - · Moving towards 64 bit in IPv6 network cards. Compatibility?
 - Burned in Network Interface Card (NIC) ROM, also sometimes software settable.
 - 3 Bytes Organization Identifier + 3 bytes NIC Identifier.
 - 2⁴⁸ or 281,474,976,710,656 unique addresses, Expected to last until year 2100. Used in most IEEE 802 networks (Ethernet, WiFi), Bluetooth,
 - Does it matter if MAC addresses are re-used?

24/11/2016

COMP28411 Link Layer

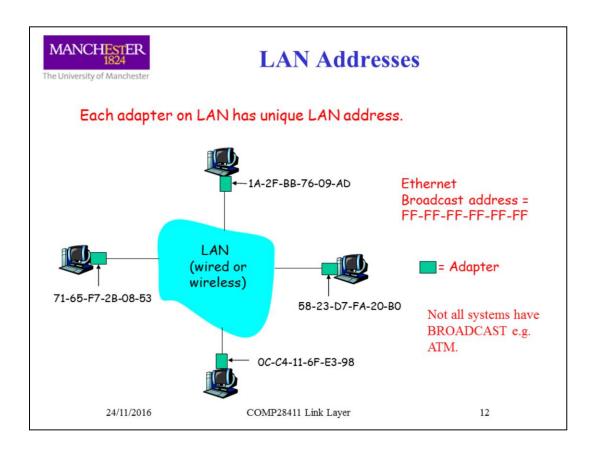
11

<continued from 10>

Closed loop flow control uses reports of resource demand and problems to control flow. ABR uses Resource Management Records (RM-cells in ATM) so that switches can set either absolute or relative (+/-) data rates for stream connections. The RM-cells are sent to sources and used to throttle forwarded traffic. However, this is a complex scheme so has had limited adoption.

The Internet is a multi-network communications device. IP was designed for geographical routing between nodes and hence the ID an IP address gives a device is location limited and of limited use for mobile devices like laptops, tablets, PDAs and mobile phones.

Hence, IP is not used for point-2-point connections at the link layer and below. Instead, a more static but non geographical address scheme is used based around having unique device addresses.



In a local network assuming we are using Ethernet there can be two or more computers which may share a single cable if using old pre 100MB speeds or be connected point-2-point for 100MB Ethernet and faster. 100MB Ethernet systems are usually connected using a switch which allows a number of machine to network connections to be interconnected to form a larger network. Switches can also be connected to one another via the same point-2-point cables used for machine to switch connections.

At the link layer the network then consists of a set of machines with fixed MAC addresses and switches that are transparent to the network in that they often have no address associated with the switch as a whole or with each switch connection point. For data monitoring and control modern switches may nowadays also have a network address which adds intelligence but this is not necessary for their operation.

The broadcast address can be used for the machines to talk to each other without knowing their addresses. However, ATM which is often used as a link layer protocol never agreed how broadcast would happen so it has never been implemented. As ATM is mainly used in 1:1 point-2-point virtual circuits, broadcast is not needed.



LAN Address (more)

- · MAC address allocation administered by IEEE
- Manufacturer buys portion of MAC address space (to assure uniqueness).
- Analogy:
 - (a) MAC address: like Passport Number or social security or National ID card.
 - (b) IP address: like postal address
- MAC flat address gives portability
 - · Can move LAN card from one LAN to another
 - · Switches auto-adapt to new card
- IP hierarchical address NOT portable
 - · Address depends on IP subnet where node is attached (geography/location).
 - · IP routing depends on attachment to network

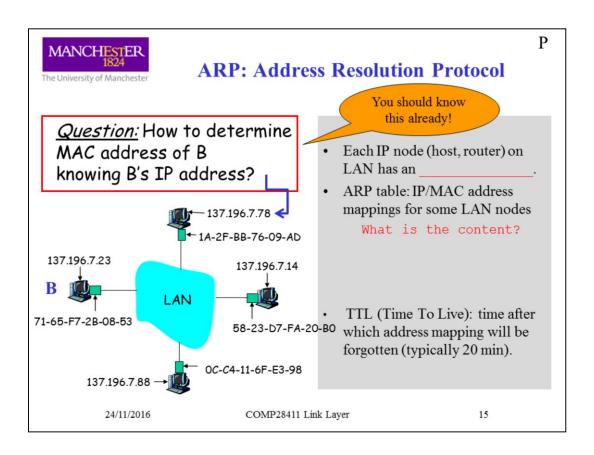
24/11/2016

COMP28411 Link Layer

13

The 48 bit MAC address is a true ID in that in theory there will never be another device with the same address. Of course, in practice, it is not that difficult to spoof or clone MAC addresses so it is not a perfect ID. I'm not sure any form of perfect ID exists!

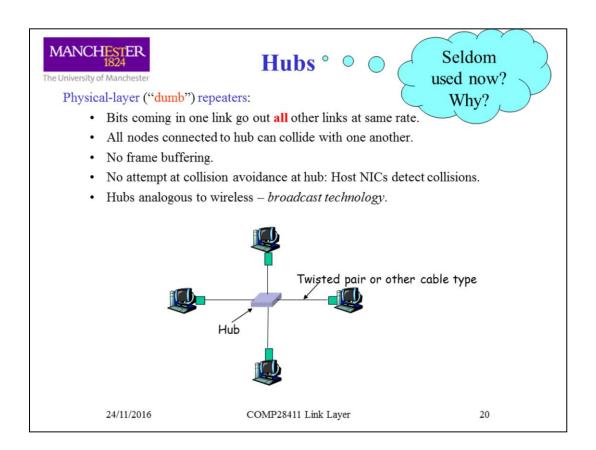
The big advantage of the ID nature of MAC is that it is 100% portable. MAC is therefore much more useful in some ways than IP. However, we still need IP as imagine the size of the routing tables that would be needed if you could plug in any MAC address device anywhere. Because there is no geography built in you would need router tables with several billion entries in today's Internet!



In wired networks a long time to live (TTL) makes sense because the entropy of the network is low – nodes seldom move. In environments where laptops and tablet computers are used the time a given device is connected to the network is fairly low. In order to efficiently cycle through sets of dynamically allocated IP addresses (DHCP) the TTL may be much shorter. In fact, it makes sense for the TTL to be dynamically adjusted as patterns of usage vary greatly as does demand for IP addresses. Imagine the lower first floor of the Kilburn building. Overnight there are almost no users and no movement so a TTL that is long is perfectly acceptable and reduces the amount of ARP maintenance traffic. From about 8am the pattern changes. Then from around 9 till about say 17:30 the space is very busy with very transient visitors between 45 and 00 each hour. Between 00 and roughly 45 each hour quite a lot of visiting machines will be static.

In IPv6 ARP is renamed the Neighbour Discovery Protocol (NDP). The new protocol has very similar features but adds more states to handle mobile and wireless nodes where loss of connectivity can be transient due to interference or may be more permanent. The table structure is more complex.

ARP and NDP traffic is seldom a big problem on networks but it does use CPU and power resources. The more mobile nodes there are the higher the ARP/NDP traffic load is likely to be and with low power devices this might significantly reduce there battery life. Imagine a network with say 100,000 devices per square meter (not that high a density) and the costs may/will require application and scenario specific tuning?



These are now almost never seen and were only used in Ethernets working up to a maximum of 10Mbps. In such networks, the Ethernet allowed many machines to be connected to a single wire connection which was therefore shared. Hubs were a device for allowing several different wires to be interconnected transparently into a single Ethernet. In effect, a hub was simply a method of shorting the wires together to form a circuit. Some hubs also had filters to clean the analogue waveforms and amplifiers to increase the signal strength but all of these were analogue components. The digital data was not extracted as a frame!

Hubs are superseded by switches which are digital devices meaning that frames are placed into registers or RAM memory and a CPU or other hardware decides what to do with the frame.





Switch & Hubs



- Older networks were Half-Duplex, wires were shared by many machines. They used Hubs.
- Modern networks are Full-Duplex over unshared point-2point wires. They use Switches.
 - Switches are Link-layer device: Smarter than hubs, take active role.
 - · Store + forward frames.
 - Examine incoming frame's MAC address, selectively forward frame to one-or-more outgoing links when frame is to be forwarded on segment, uses collision avoidance techniques to access segment.
 - Transparent
 - · Hosts are unaware of presence of switches.
 - · Plug-and-play, self-learning
 - · Switches do not need to be configured.
 - · Remember that routers do need configuration!



24/11/2016

COMP28411 Link Layer

21

In older networks a single wire was shared by many machines being attached to the wire. Being a single connection, for successful use just one machine could write to the wire at a time though all the machines could read the content of data on the wire. This one transmitter at a time configuration is called half-duplex.

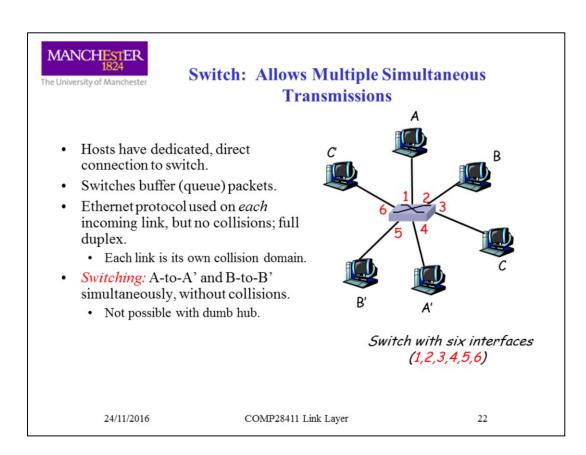
Modern networks use either 2 wires one for transmitting and one for receiving for each connection or sometimes the two directions are multiplexed on to a single connection using e.g. different frequencies to separate the upload and download directions. This is called full duplex.

Modern Ethernets operating at 100Mbps or greater speeds use unshared point 2 point wires where each end of the wire can transmit and receive at the same time over independent channels. This is full 2 way communication.

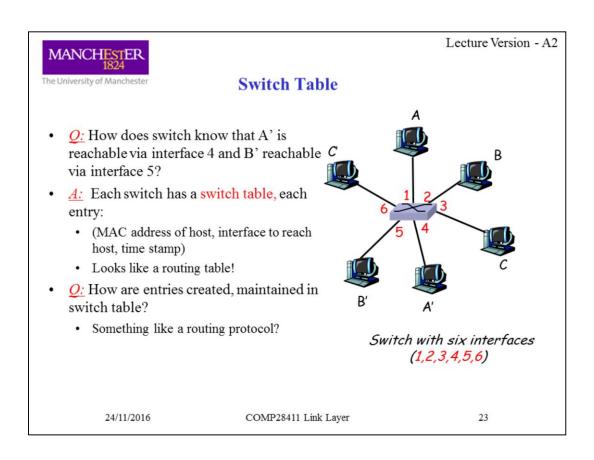
Some digital systems can appear to be full duplex by swapping rapidly from upload to down load but using a half duplex physical connection.

Most existing mobile phone systems have specifications for either full or half duplex. In full duplex wireless systems the uplink and downlink will operate on different frequencies at the same time or at offset times. Time offsets are used in many wireless systems to make them cheaper by having a single transceiver which is either transmitting or receiving at any given time. By switching from transit to receive quickly users will still appear to have a full duplex service.

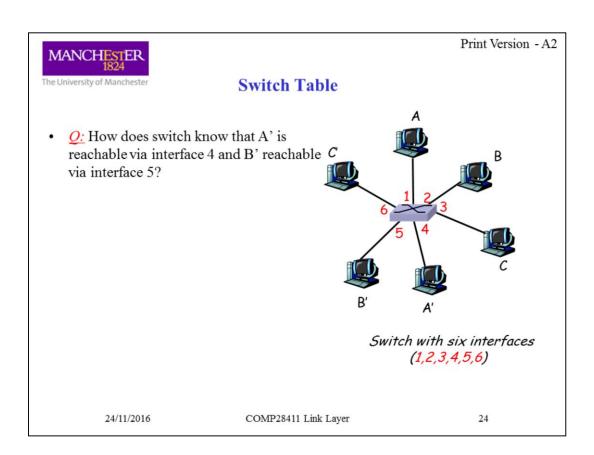
Switches are designed to be plug-and-play. They do not any external configuration. Anything they need to know about their environment is learned.

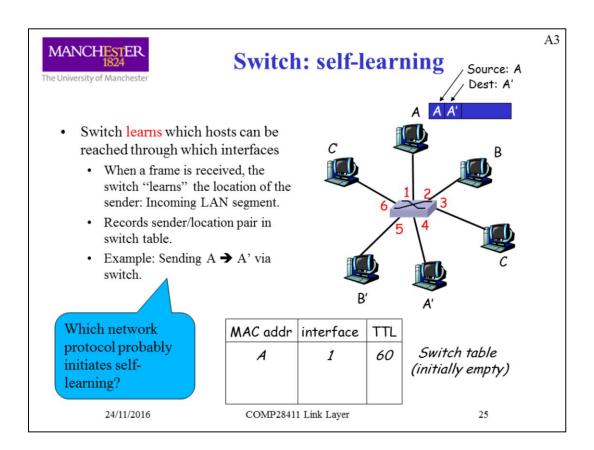


Each node connecting to a switch has separate uplink and downlink channels which usually share a single wire bundle though this wire bundle may have several different wires within it. A switch can queue (buffer) at least one inward and one outward frame per connection point it has. Sending a frame though a switch always has some delay due to the frame being at least partly decoded to extract the addresses being used and the frame being copied from buffer to buffer within the switch. Of course, some copying can be avoided by exchanging pointers to addresses in memory instead of copying the data at one address to the other address . This is much cheaper as frames may have thousands of bytes of data .

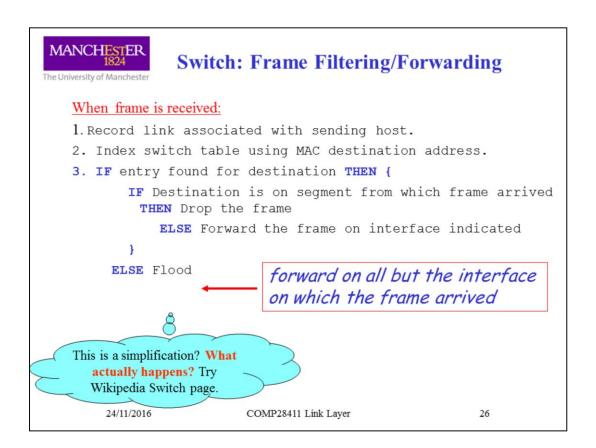


The switch builds a table of which node is connected to which interface on the switch. This looks very similar to an ARP table or a routing table.

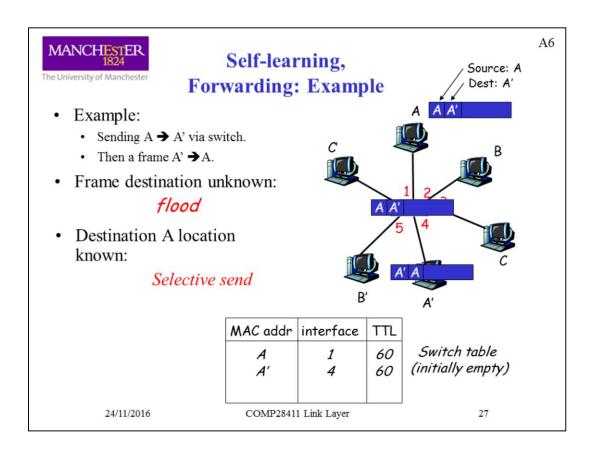




Just like ARP, the switch table includes a time to live (TTL) field. This field is updated on every frame exchange. But, if the TTL counts down to zero then the record in the table is deleted. The next frame exchange will re-learn the information. But, if the next frame exchange is a request to send a frame to an unknown MAC address, the switch may need to initiate a search for the unknown address on all its currently unlabelled links.



Switches have an in-built minimum spanning tree algorithm to search with.



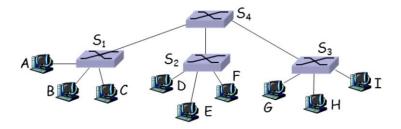
A sends a frame destined for A' to the switch on connection 1. We assume the routing table in the switch is currently empty. On receiving this first frame the routing table has a record inserted saying that A (or its MAC address) is reached via interface 1.

The switch table does not have a current entry for the MAC address A', therefore the switch send a a flood request for the address A' on all its interfaces except 1 on this occasion. A' recognizes its address and returns a confirmation frame to the switch. This is very like ARP. However, we have been told switches are transparent so they cannot send there own frames or expect relies destined just for the switch. What actually happens is that the frame from A being sent to A' is sent to all the unlabelled exit interfaces. The arriving frame is dropped by all the receivers except A'. However, the switch still does not know that A' is attached to interface 4. This knowledge is learned when A' sends a frame to A or anybody else.

Ethernets do not have acknowledgement frames. The frame types are listed here http://tinyurl.com/oel8gq4.



· Switches can be connected together



- Q: Sending from C to I →
 How does S₁ know to forward frame destined to I via S₄ and S₃?
- A: Self learning! (works exactly the same as in single-switch case!)

24/11/2016 COMP28411 Link Layer 28

Large networks may contain many switches. Sending a frame across a network of switches uses the same self learning algorithm that we have already seen! But, there are now more stages.



Lecture Version - A4

Self-learning Multi-Switch Example

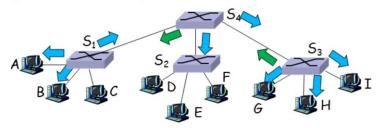
Suppose C sends frame to I, I responds to C – How learn how to do it?

S₁ has never heard of I so floods

Flood reaches S_4 , also not heard of I so floods

Flood reaches S_3 , also not heard of I so floods.

S₃ now knows I, tell S₄ which tells S₁.



Q: Show switch tables and packet forwarding in S₁, S₂, S₃, S₄

24/11/2016

COMP28411 Link Layer

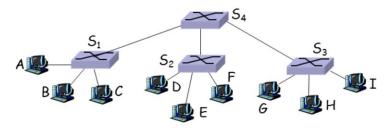
29



Print Version - A4

Self-learning Multi-Switch Example

Suppose C sends frame to I, I responds to $C - \underline{Q}$: How learn how to do it?

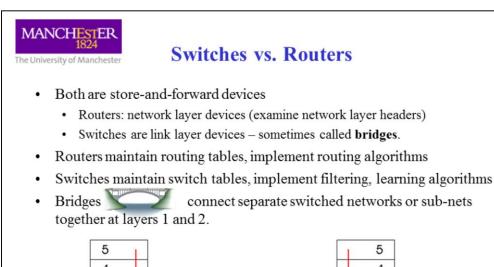


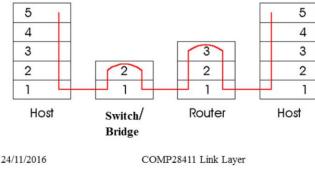
Q: Show switch tables and packet forwarding in S₁, S₂, S₃, S₄

24/11/2016

COMP28411 Link Layer

30





As said, switches are physical layer and data link layer devices. However, switches do not have there own data link layer addresses, simply internal interface names used for forwarding frames. Routing is a network layer function, switches do not have a network layer.

32

Several groups of nodes can be interlinked by connecting them at the physical layer (hb) or data link layer (switch). Devices/nodes at these layers can be transparent in which case they do not need or use source or destination addresses.

These devices can be used to bridge or interconnect different groups of nodes. Bridges can also be transparent or visible as links. Interestingly, some bridges are semi-transparent. Wfi has a special bridge mode whereby a frame can be addressed to another device via an intermediate bridge which itself has two ends each with its own address. These cross bridge frames therefore have 4 addresses.



Lecture Summary

- · Link Layer is concerned with node to node communications.
- It is normally implemented in hardware but may be soft coded
 - · Services
 - · Where implemented
 - Packet Encapsulation reminder
 - · Flow Control
 - · Link layer addressing
 - · Address Resolution
 - · Hubs and Switches
- Next:
 - More on addresses and Sharing a network Link Layer Routing (between switches)
 - Multiple access sharing a network.
 - · Collisions and channel partitioning.
 - · Random Access protocols

24/11/2016

COMP28411 Link Layer

33



Questions?

- · What information passes between:
 - · The network layer and the link layer?
 - · The link layer and the physical layer?
- If the transport layer can do error checking, why is error detection and correction important at the link layer?
- The network layer has the full address for the datagram, why doesn't the link layer use network addresses?
- What extra hardware is needed to implement a full-duplex link compared to a half-duplex link?
- · If data is buffered at other layers, why does the link layer also do buffering?
 - · Why does the network layer buffer?
 - · Do you think the physical layer buffers? Why or why not?
- Why do you think IPv4 addresses were 32 bits whereas Ethernet MAC addresses were 48 bits?
 - Why are IPv6 IP addresses 128 bits whereas the IPv6 MAC address equivalent is 64 bits?
- What is "Inverse ARP" (InARP)? Why is it used?

24/11/2016

COMP28411 Link Layer

34