Datalink Layer - Sharing Resources

COMPUTER NETWORKS A.A. 24/25

Leonardo Maccari, DAIS: Ca' Foscari University of Venice,
leonardo.maccari@unive.it

Venice, fall 2024

License

Sect. 1 Beyond One Link: Network Topologies

More Than Two Hosts

- So far we focused on a single link between two hosts, but we generally want to connect more
- This situation mainly occurs in Local Area Networks (LAN). A LAN is a network that efficiently interconnects several hosts (usually a few dozens to a few hundreds) in the same room, building or campus.
- Consider for example a network with five hosts. Any of these hosts need to be able to exchange information with any of the other hosts.
- How do we connect them?

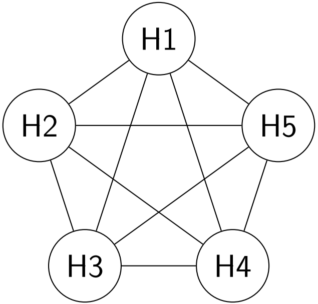
A Full Mesh Network

Pros:

- Maximum Capacity: every communication has one dedicated link
- Maximum robustness: resists to failure

Cons:

- A total of $\frac{n(n-1)}{2}$ physical links
- $n-1$ network interfaces per host



Hardly Ever Used

A Bus Network

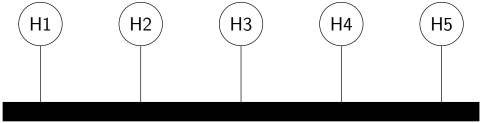
Pros:

- Very cheap, only one physical link

Cons:

- Not scalable: all communications share the same physical link
- Not robust: if the bus breaks the network is broken in two parts

Used in the early days of Networks, or in some specific scenarios (for instance, the so-called CAN-bus of your car)



A Ring Network

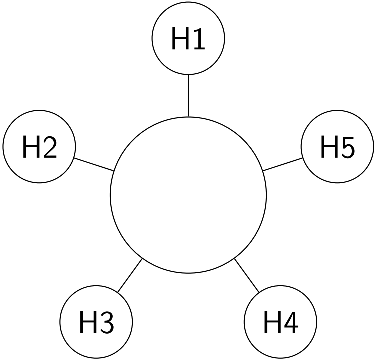
Pros:

- A more reliable link, it tolerates one failure

Cons:

- Still not scalable

Used in metropolitan networks to add redundancy when cabling is very expensive



A Star Network

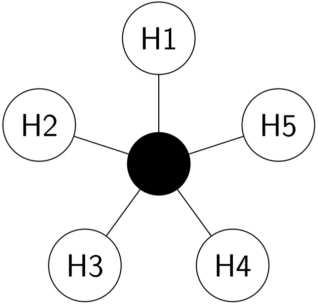
Pros:

- Best performance/cost trade-off: one dedicated link per host: $(n-1)$ physical links
- Better manageability, one point controls all
- (Almost) resistant to link failures

Cons:

- Concentrated: if the star-center dies, the network stops working

Used in LANs, for instance an Ethernet Switch.



A Tree Network

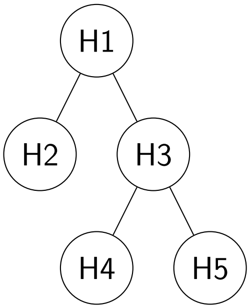
Pros:

- Extends a star network
- Best performance/cost trade-off: one dedicated link per host, $(n-1)$ physical links

Cons:

- Not resistant to link failures, if one link dies the network is disconnected

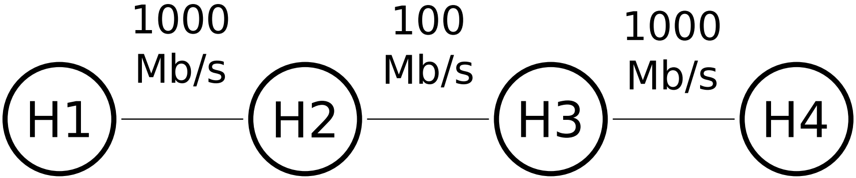
Used to extend LANs with switches in cascade, and in many other circumstances.



Bottleneck

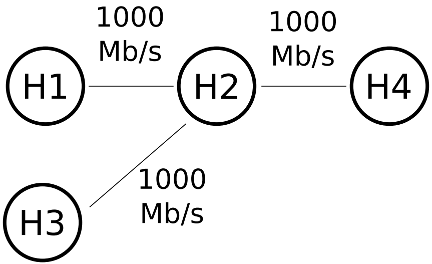
- When we extend a single link we say the communication is *multi-hop*
- In a multi-hop communication the traffic is repeated on all hops.
- In a multi-hop communication the bottleneck is the slowest of the links in the path.
- Traffic can be multiplexed and de-multiplexed, that is, traffic from several links can be scheduled in a single link *downstream*.

Bottleneck



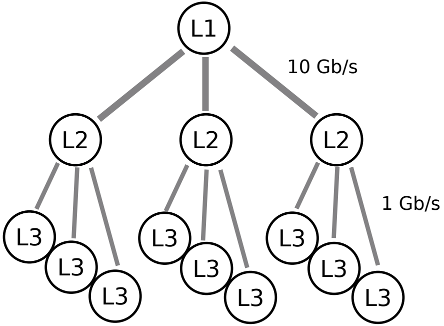
- The communication between H1 and H4 are limited to 100 Mb/s by the presence of the central link

Bottleneck



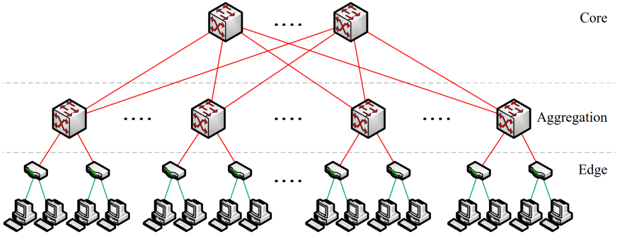
- The maximum bit-rate between H1 and H4 is 1Gb/s
- However when both H1 and H3 communicate with H4, their bit-rate is limited to 500 Mb/s, because they share the link H2-H4

Trees



In a hierarchical network (like a tree topology) when you move from the leaf nodes towards the root, the link capacity must increase.

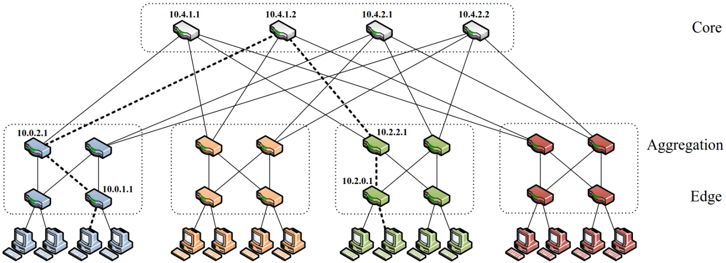
Example, a Data Center Network[1]



- A typical data-center network uses edge switches (low-end switches) to connect servers
- It uses high-end switches to connect the aggregation layer and core switches
- Fault tolerance is achieved by redundancy of core switches

[1] Images by Al-Fares et al. "A Scalable, Commodity Data Center Network Architecture", Copyright 2008 ACM, not released with CC license.

Example, a Fat-Tree[2]



- A fat-tree uses the same approach, but only with low-end switches.
- You need more switches but the overall cost is lower (a core switch can cost 100x a low-end one)
- You need multipath routing to distribute the traffic

[2] Images by Al-Fares et al. "A Scalable, Commodity Data Center Network Architecture", Copyright 2008 ACM, not released with CC license.

High-Availability Topologies

- Data-centers require high capacity and high resiliency, and use some specifically crafted network topologies
- CLOS and Fat-Trees topologies are used in these circumstances, we will not get into the details, but at least you know they exist

Sect. 2 Sharing One Link: Media Access

Sharing the media on a Wired Star

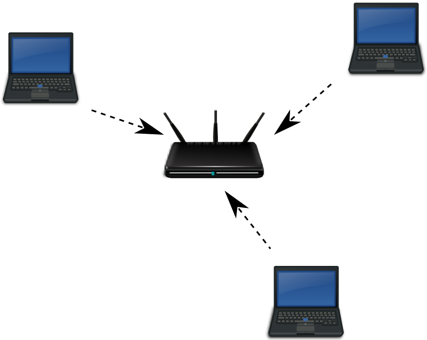
- In a star network collisions are normally not a problem.
- Modern Ethernet cables (more on this in future lessons) use two separate circuits for transmission and reception, one twisted pair per transmission and reception in one cable, and one dedicated cable per host
- This prevents collisions to happen.
- In previous version of Ethernet, this was instead a problem.

Sharing the Media on a Bus

- Consider the bus network. Every transmission passes on the same cable
- Transmissions then interfere with each other, if two hosts transmit at the same time, the receiver will not distinguish the two signals
- Both frames will be dropped.
- There must be a way to schedule the frames avoiding *collisions*.

Wireless Networks

- A wireless network is logically organized as a star, but physically works as a bus.
- The receiver has one single radio, if two hosts communicate at the same time, there is a collision.
- This is true for a Wi-Fi network, but also for cellular networks



We need to schedule the access to the physical media. This is one of the tasks of the Datalink layer.

Sharing Resources

- So with some network organization, the terminals will necessarily share the physical media
- When this happens they must find a way not to interfere with each other.
- There are several ways of obtaining this, FDMA, TDMA, CDMA, and they can be combined.

Sharing One Link: Media Access

↳2.1 Frequency Division

Frequency-Division Multiple Access: FDMA

- One way to avoid collisions is to use a different frequency for every communication.
- We have seen that at the physical layer the signal is *modulated* on a certain carrier.
- Radios are able to filter the communications that are not using a specific carrier frequency
- So if one carrier uses frequency X and another uses frequency Y, the communications can take place at the same time, even if the two receivers are next to each other
- This will not produce a collision

Wi-Fi Channels

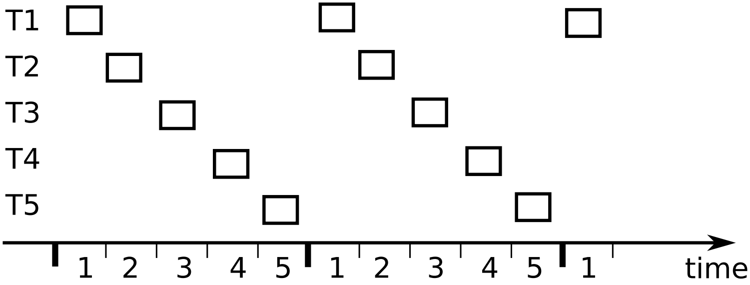
- Wi-Fi can operate on several frequencies, called *channels*
- When you turn on your Wi-Fi Access Point (AP) for the first time, it will search for a channel that is not used by other Wi-Fi networks
- This allows two networks to share the same physical space
- However, one single radio can use only one channel at a time
- So the hosts attached to one AP must use the same frequency, and thus, they must schedule their messages

Sharing One Link: Media Access

↳2.2 TDMA

Time-Division Multiple Access: TDMA

- The most intuitive way is Time-division Multiplexing (TDM).
- In a TDM, if the network has m terminals, time is divided into m *slots* of fixed size, with one slot assigned to one terminal
- slots are repeated cyclically, so terminal i can only transmit at frame i



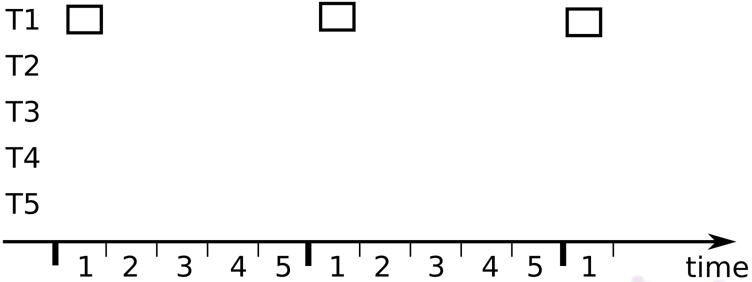
TDMA is Efficient at Saturation

- When all the stations always have something to transmit, the network is in what we refer to as *saturation*
- In that condition, TDMA is at maximum efficiency: every slot is occupied and there are no collisions.

Datalink Layer Efficiency (informally): The fraction of time the datalink is used to communicate useful data.

Time-Division Multiple Access: TDMA

- The limitation of TDM is that each terminal will have an opportunity to transmit every $\frac{1}{m}$ slots, regardless of what the others do.
- If station i has a frame to send at time $i+1$, it will have to wait a full cycle of m slots before it can transmit, even if the other nodes have nothing to transmit.
- TDM then introduces an non zero average delay.



Low load: inefficient MAC and large access delay

TDMA

- So a TMDA is efficient when the network is at congestion, and it is inefficient when the network is unloaded
- Note that when the network is congested, it generally means you have *more* traffic to send than what you can send, so you will drop some traffic.
- Practically speaking, if you have an expectation of the load, you must dimension the capacity of your network so that it is rarely congested.
- Therefore, having a system that works well in the worst case and works badly in the normal condition (as a TDMA) may be inconvenient.

TDMA

- However, TDMA is easy to implement and deterministic
- In systems in which you pay for the service (for instance, cellular networks, satellite networks) a fixed schedule allows all the users to have at least a bit of resources even at saturation.
- Even in piconets (like Bluetooth networks) TDMA can be useful because the exchanged data quantity is generally little

Exercise on TDMA

Consider a TDMA system in which the bit-rate is $b = 100\text{Mb/s}$, there are n=5 terminals, each one is a assigned a slot of length $s = 1$ ms

- What is the maximum amount of data a terminal can send in one slot?
- What is the maximum amount of data the terminal can send in one second?
- Assume terminal 1 needs to send 250kb, what is the time t it takes to completely send the frame, starting from the moment the data is received from the upper layer? (hint: there is not only one answer but one best, worst and average case)

Solutions

- $b = 100\text{Mb/s} = 10^8$ b/s; $s = 10^{-3}$sec

Solutions

- $b = 100\text{Mb/s} = 10^8$ b/s; $s = 10^{-3}$sec
- The maximum data per slot is: $s * b = 10^5 = 100\text{kb}$

Solutions

- $b = 100\text{Mb/s} = 10^8$ b/s; $s = 10^{-3}$sec
- The maximum data per slot is: $s * b = 10^5 = 100\text{kb}$
- $n = 5$ so in 1 second there are 1000/5 slots for each terminal, so the maximum bit-rate per terminal is $10^5 * 200 = 2 * 10^7 = 20.000.000 = 20\text{Mb/s}$, that is, one fifth of the available capacity

Solutions

- $b = 100\text{Mb/s} = 10^8$ b/s; $s = 10^{-3}$sec
- The maximum data per slot is: $s * b = 10^5 = 100\text{kb}$
- $n = 5$ so in 1 second there are 1000/5 slots for each terminal, so the maximum bit-rate per terminal is $10^5 * 200 = 2 * 10^7 = 20.000.000 = 20\text{Mb/s}$, that is, one fifth of the available capacity
- Sending 250kb takes 2.5 ms, so 2.5 slots. However, each slot arrives every 5 ms, so in total it takes 10.5ms *from the first available transmission opportunity*

Solutions

- $b = 100\text{Mb/s} = 10^8$ b/s; $s = 10^{-3}$sec
- The maximum data per slot is: $s * b = 10^5 = 100\text{kb}$
- $n = 5$ so in 1 second there are 1000/5 slots for each terminal, so the maximum bit-rate per terminal is $10^5 * 200 = 2 * 10^7 = 20.000.000 = 20\text{Mb/s}$, that is, one fifth of the available capacity
- Sending 250kb takes 2.5 ms, so 2.5 slots. However, each slot arrives every 5 ms, so in total it takes 10.5ms *from the first available transmission opportunity*
- However. . .

Solutions

We make a simplifying assumption: the terminal must receive the frame from the application **before** its slot begins, not after the slot begun.

- best case: $t = 10.5$ms (the frame arrives right before the slot begins)

Solutions

We make a simplifying assumption: the terminal must receive the frame from the application **before** its slot begins, not after the slot begun.

- best case: $t = 10.5$ms (the frame arrives right before the slot begins)
- worst case: $t = 10.5 + 5 = 15.5$ms (the frame arrives right after the slot begins)

Solutions

We make a simplifying assumption: the terminal must receive the frame from the application **before** its slot begins, not after the slot begun.

- best case: $t = 10.5$ms (the frame arrives right before the slot begins)
- worst case: $t = 10.5 + 5 = 15.5$ms (the frame arrives right after the slot begins)
- average case: $t = 10.5 + (5 + 0)/2 = 13$ms

Sharing One Link: Media Access

↳2.3 Random Access

Random Access

- Another way of accessing the media is through random access.
- In random access, terminals try to transmit, and if a collision occurs, the collision is detected and the frame is sent again.
- Conceptually they are simple schemes, but studying their performance is not easy.
- The father of all Random Access schemes is ALOHA.

ALOHA

- ALOHA is a system proposed in the 70s to avoid the need for coordination, and it is the ancestor of the CSMA/CA used in Wi-Fi today (more on this in future lessons)
- There are two versions of it, we look at the "slotted" one in which terminals need to be synchronized, and time is divided in slots

S-ALOHA

- The difference with TDMA is that every terminal can transmit in every slot.
- With S-ALOHA when a node needs to send a frame, it waits for the beginning of the next slot, and just transmit it.
- S-ALOHA admits a probability of collision, collisions make the transmission fail.
- Let's look at the basic assumptions that make the protocol mathematically tractable.

S-ALOHA: Assumptions

1. There are m synchronized terminals that transmit to a single receiver.
2. Being synchronized, terminals know the beginning of every slot
3. Transmissions are always shorter than a slot time

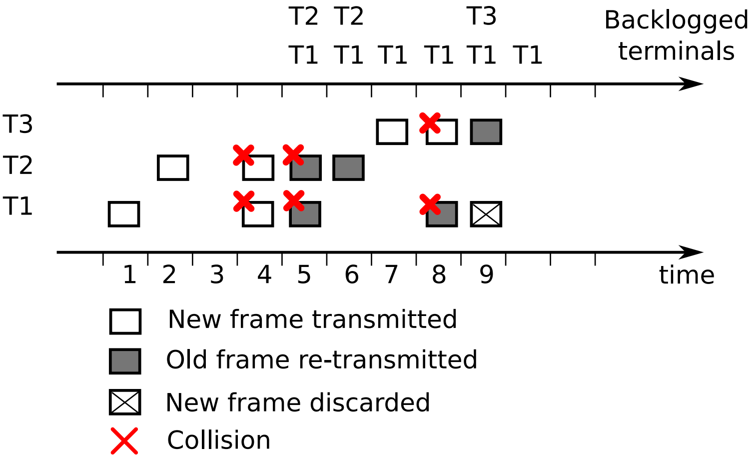
S-ALOHA: Assumptions (2)

4. Two contemporary transmission make the reception fail. We assume the terminals immediately know that this happened.
5. In the real world the receiver has a dedicated slot in which nobody else can transmit. It uses it to send acknowledgments of frames
6. If a collision happens, the sender will not receive an ACK and can detect the failure with some delay.

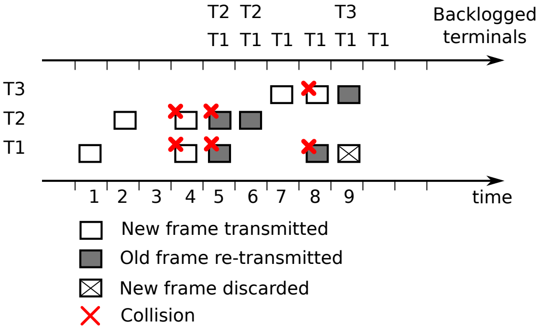
S-ALOHA Protocol Details

- When the datalink receives a SDU frame from the upper layer, it waits till the next slot, and transmit. We call q_a the probability a node has something to transmit.
- If another terminal transmits in the same frame, both enter a *backlogged* state.
- This means they will re-transmit in the next slots with probability $q_b > q_a$.
- A backlogged terminal goes back to normal state when it is able to transmit successfully.
- If a backlogged terminal receives another (new) frame to send, it discards it.
- q_a depends on incoming traffic while q_b is a network parameter.

S-ALOHA: Evolution

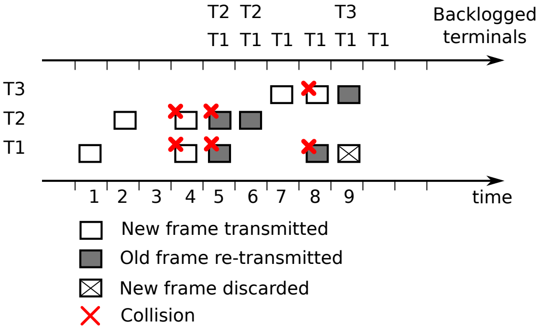


S-ALOHA: Evolution



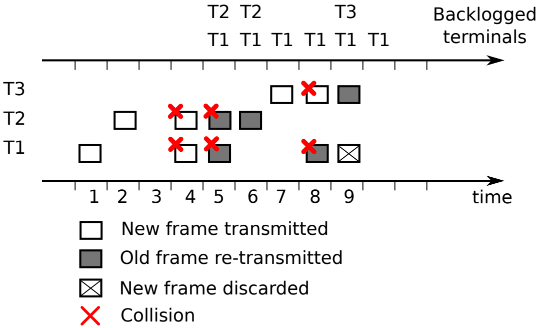
- At time 1 and 2, T1 and T2 send a new frame, successfully
- at time 3, no one is transmitting
- at time 4, a collision occurs. T1 and T2 enter the backlogged state.

S-ALOHA: Evolution



- At time 5, T1 and T2 try to re-transmit the old frame, but collide again.
- at time 6, T2 manages to re-transmit the old frame, exits the backlogged state
- at time 7, T3 transmits a new frame successfully, T1 is backlogged and does not transmit

S-ALOHA: Evolution



- At time 8, T1 and T3 collide, T3 also enters the backlogged state
- at time 9, T3 sends the old frame and exits the backlogged state.
- at time 9, T1 receives a new frame, but discards it without sending it because it does not have a queue.

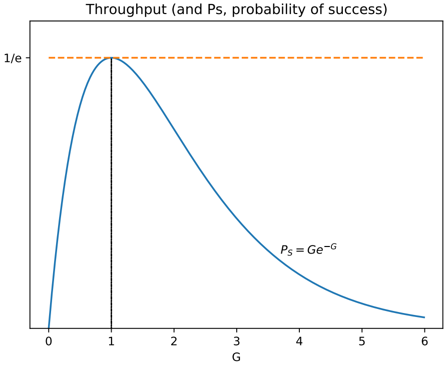
Aloha Performance, Static Analysis

- Assume G is the system load, that is, the average number of frames that need to be transmitted per slot, from all terminals
- The probability of correctly transmitting a frame P_s is given by the probability of having only 1 frame transmitted in a slot, so no collisions.
- It can be shown that (under certain assumptions of asymptotical behavior):

$$P_S = Ge^{-G}$$

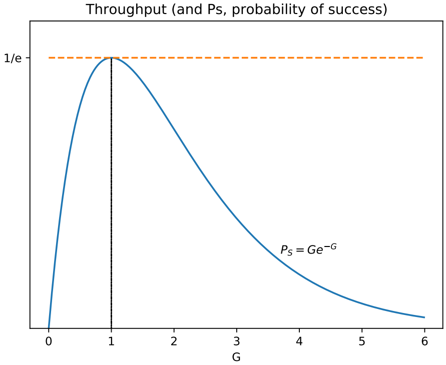
- Since we send one frame per slot, this is also proportional to the *throughput*: the maximum achievable bit-rate on the network and to the layer efficiency.

Successful Transmission Probability (throughput)



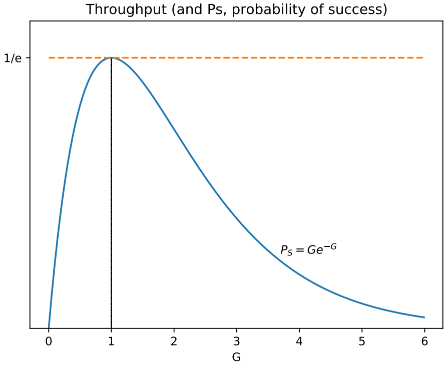
- The blue line shows P_S as a function of G.
- When G is lower than 1, we are under-utilizing the link
- Many slots are empty, so there are only a few collisions, efficiency (throughput) is low
- As expected, the link works well when the network is unloaded

Successful Transmission Probability (throughput)



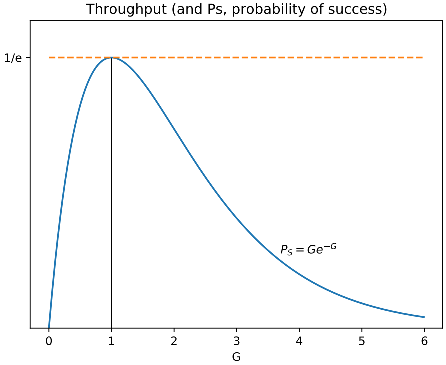
- When $G = 1$ the network is at saturation, because we can't send more than 1 slot per frame.
- We have the highest throughput: $1/e$ ($e =$ Napier constant).
- But when $G = 1$ the throughput is $1/e = 0.36 < 1 = G$.
- So if we try to send on average one frame per slot, 64% of them are lost!

Successful Transmission Probability (throughput)



- Note that if you have $n = 10$ terminals, the highest efficiency is when $G = 1$
- G is the average load on the network, so that it requires that in average every terminal sends one frame every 10 slots, $q_a = 1/10$

Successful Transmission Probability (throughput)



- If we increase the load, we decrease the throughput
- It means that we are increasing the collision probability, and so the efficiency of the link decreases

Congestion Collapse

- What you have observed is a pattern that you can find often in networks

Congestion collapse: Congestion collapse happens when the relative performance of the system degrades abruptly with the increase of the load.

Collapse of TDMA

- TDMA never collapses.
- Even if the network is overloaded (you have more traffic to send than what the network can handle) every terminal has its little share of resources
- If you increase the terminals, every one has a smaller share
- However, the network usage is 100%, the network is at full efficiency (even if frames are dropped)
- This is what you intuitively expect: high load $\rightarrow$ less resources for everyone.

Collapse of Random Access

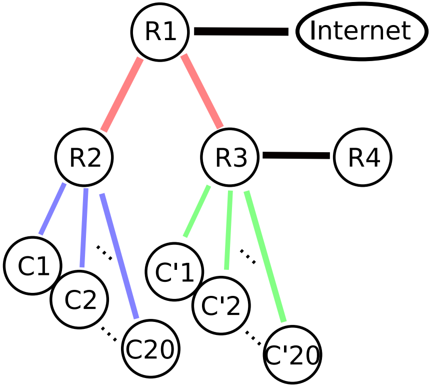
- Random access instead does not follow a linear, deterministic distribution of resources
- Not only the resources are reduced. The resources for every terminal simply go to zero: the system collapsed

ALOHA and Modern Networks

- ALOHA has other unwanted properties, for instance, it is proven to be instable.
- When we will study 802.11 wireless networks we will see how they improved ALOHA to and stabilize the network and reduce the probability of a collapse.

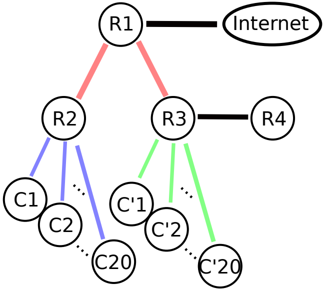
Exercise: Putting Everything Together

- Links of same color share same physical medium
- The red links share a phy with: B=20MHz, M=8, TDMA with 3 slots.
- Blue links use S-ALOHA with B=20MHz, M=16
- Green links use S-ALOHA with B=5MHz, M=8
- Black links are dedicated links, never a bottleneck
- Assume perfect links (use Nyquist)
- Assume S-ALOHA behaves as described (asymptotical behavior, if more than one terminal transmits)



Putting Everything Together: Q1

- Assume C1 is the only client that is generating some traffic.
- C1 generates a Constant Bit-Rate (CBR) traffic of 40 Mb/s to the Internet.



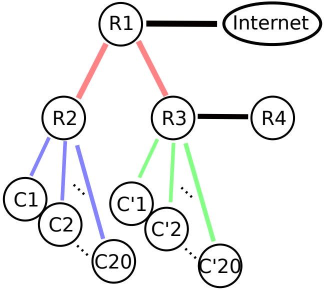
? | **Is C1 able to successfully deliver all its traffic?**

Answer 1

- The red network has a capacity $C_r = 2 * 20 * log_2(8) = 120Mb/s$
- There are three stations (R1,R2,R3) so TDMA provides 1/3 of the capacity for each: 40Mb/s
- The blue network has a capacity $C_b = 2 * 20 * log_2(16) = 160Mb/s$
- S-ALOHA has an efficiency of $1/e$ when multiple terminals are on. However, C1 is the only one transmitting, so it can just use all the capacity.
- Then yes, C1 delivers all its traffic.

Putting Everything Together: Q2

- Now together with C1 also all terminals C2-C20 send CBR traffic of 1 Mb/s, with final destination R2.



Are all clients able to successfully deliver all the traffic?

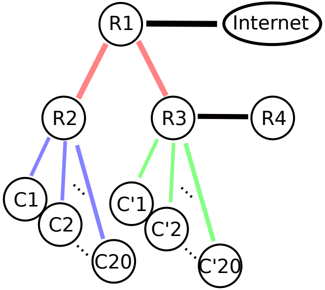
Answer 2

- The traffic from C2-C20 stays in the blue network.
- The maximum efficiency of S-ALOHA is $\frac{160}{e} \simeq 59Mb/s$
- Overall the blue network generates 40+19, so S-ALOHA can deliver it

Putting Everything Together: Q3

- Same as before, but terminals C2-C20 sends CBR traffic to the Internet.

Are all clients able to successfully deliver all the traffic?



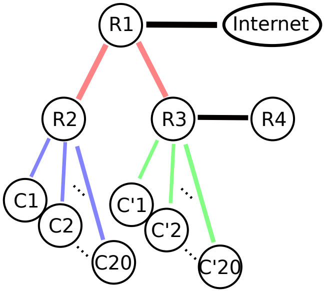
Answer 3

- The aggregated traffic in link R2-R1 sums to 59 Mb/s.
- This is more than the allowed 40 Mb/s, so R2 will drop 19 Mb/s
- So the answer is No because the red link R2-R1 is a **bottleneck**

Putting Everything Together: Q4

- Now C1 generates a 40 Mb/s CBR traffic to C'1. All other terminals are not generating traffic.

? **Is the communication possible? are there bottlenecks?**



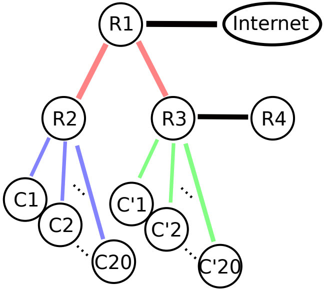
Answer 4

- The links C1-R2 has capacity 59Mb/s, R1-R2 and R2-R3 have capacity 40Mb/s
- The green network has a capacity $C_g = 2 * 5 * log_2(8) = 30Mb/s$
- Only R3 is transmitting, so the efficiency is 100%
- However, it is not enough.
- So the answer is No because the green link R3-C'1 is a **bottleneck**

Putting Everything Together: Q5

- C1 generates a 3 Mb/s CBR traffic to C'1.
- Also C'2-C'20 generate 1.42 Mb/s towards the Internet.

? **What is the maximum communication speed from C1 to C'1?**



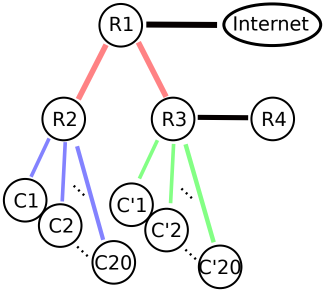
Answer 5

- The green network allows $C_g = \frac{30}{e} \simeq 11Mb/s$
- This is possible when G=1, so every slot is used, so when 30Mb/s are generated.
- The total load is $1.42 * 19 + 3 = 30$. However, only 11 Mb/s are delivered, if we assume these are shared among all the terminals, then C'1 receives only $\frac{11}{20} = 0.55Mb/s$.
- The green network is the bottleneck.

Putting Everything Together: Q6

- Same as Q5, but also R4 generates 30 Mb/s towards the Internet.

Can R4 successfully deliver its traffic?



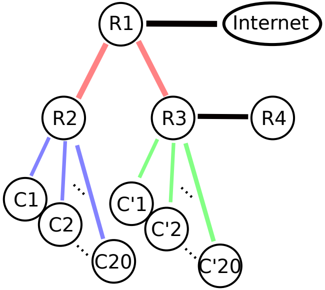
Answer 6

- The green network allows $C_g = \frac{30}{e} \simeq 11Mb/s$
- So on link R3-R1 there are 11 + 30 Mb/s
- However we know that the capacity of R3-R1 is only 40 Mb/s.
- So no, R4 will lose some frames, as the red network is another bottleneck

Putting Everything Together: Q7

- Same as Q6, but C'2-C'20 generate 3 Mb/s instead of 1.42.

Can R4 successfully deliver its traffic?



Answer 7

- The green network allows $C_g = \frac{30}{e} \simeq 11Mb/s$
- This is possible when G=1, so every slot is used, so when 30Mb/s are generated.
- The total load is $3 * 19 + 3 = 60$, so G=2 and the efficiency drops to $2e^{-2} = 0.27$.
- Then the outgoing traffic from the green network is $30 * 0.27 = 8Mb/s$
- Then R3-R1 has 38 Mb/s and R4 can deliver all its traffic.
- Note that the collapse of the green network helps R4 to deliver.

Conclusions

- On the one hand you have deterministic systems like TDMA, that guarantee performance and are easy to implement, but are inflexible and waste resources with low load
- On the other you have random access schemes, that are faster at low load, but may totally collapse at high load.
- They are both used, in different applications.

Conclusions (2)

- The analysis of TDMA and ALOHA make us understand another cause of unreliability of the links.
- During congestion, the datalink layer drops frames, and this is unavoidable
- Even without congestion, a random access scheme drops frames.

! **The datalink layer is inherently unreliable, it can drop frames. The upper layers must be aware of this and try to compensate.**