CS-534: Packet Switch Architecture   Department of Computer Science
Spring 2013   © copyright: University of Crete, Greece

Exercise Set 2:
Transmission Rate and Throughput, Turn-Around Overhead

Assigned: Mon. 4 Mar. 2013 (week 2) -- Due: Wed. 20 Mar. 2013 (week 4)

2.0   Conventions: Is "Mega" 1,000,000 or 1,048,576 ?

In this exercise set (and in general), 1 MHz = 1,000,000 Hz and 1 GHz = 1,000 MHz; also, 1 ms = (1/1000) s, 1us = (1/1000) ms, 1 ns = (1/1000) us.

On the other hand, as customary in computer science and engineering, we will assume that 1 KByte = 1024 bytes, and 1 MByte = 1024 KBytes = 1,048,576 bytes.

The situation is more complicated with Mbits and Gbits. When talking e.g. of a RAM, its size is usually a power of 2, so a 1 Mbit RAM contains usually 1,048,576 bits. On the other hand, when talking of a network link that is driven by a circuit with a 100 MHz clock and transmits 1 bit during each clock period, we call the link throughput "100 Mbits/s" for convenience; in this case, 1 Mbit = 1,000,000 bits. So, the meaning of 1 Mbit or 1 Gbit is *context dependent*! In this exercise set, you will use the definition 1 Mbit = 1,000,000 bits, and 1 Gbit = 1,000 Mbits.

By the way, I recommend that you stick to the common conventions that Mb (or Gb) is Mbit (or Gbit), while MB (or GB) is MByte (or GByte); also, remember that Mbps (or Gbps) is Mbits/s (or Gbits/s). Concerning "kilo", notice that 1000 is small "k", as in km or kHz, while 1024 is capital "K", as in KBytes. A good reference for the above is the Peterson/Davie "Computer Networks" book, page 26.

2.1   SONET Rates

SONET (Synchronous Optical Network) has its roots in digital telephony. An (uncompressed) digital telephone connection transmits (in each direction) 8000 voice samples per second, where each sample is one byte. Thus, each SONET frame takes exactly 0.125 ms (1/8000 s) to be transmitted, so that, when carrying multiplexed digital telephony circuits, each frame contains (in its payload) one byte per telephony circuit (the SONET payload may alternatively contain other things, e.g. IP packets, or ATM cells).

The basic SONET rate is OC-1, where each frame consists of 810 bytes of 8 bits each. An OC-*N* frame consists of a number *N* of OC-1 frames that are *byte-interleaved* with each other.

Calculate the "fast clock" frequency (as defined in class) for the (bit-serial) transmission of OC-3, OC-12, OC-48, OC-192, and OC-768, which are the most popular SONET rates of the last and coming few years.

2.2   SONET, ATM, and UTOPIA Throughput

The 810 bytes in each SONET OC-1 frame are considered to be arranged in 9 "rows" of 90 columns each. Three (3) out of the 90 bytes in each row of the frame are overhead, (framing and many other bits), and the remaining 87 bytes per row contain (mostly) payload.

(a) Calculate the "payload throughput" of OC-12, OC-48, OC-192, and OC-768 in Mbits/s (how many millions of *payload* bits are carried per second). Compare and contrast this to the transmission rate that you calculated in exercise 2.1 above, which concerned all bits --payload and overhead (you calculated that rate in MHz, but it corresponded to an equal number of Mbps).

(b) Consider now a SONET link that carries ATM cells in the payload of its frames. Each ATM cell occupies 53 of the frame's payload bytes, and consists of a 5-byte (ATM) header plus 48 bytes of

(ATM) payload. The ATM cells are packed, one immediately after the other, within the SONET frame payload; ATM cells are allowed to cross SONET frame boundaries --the SONET frame's payload size is not, in general, a multiple of 53. Calculate the "ATM payload throughput" of OC-12, OC-48, OC-192, and OC-768 in MBytes/s and in Mbps (how many MB or millions bits are carried in the payloads of the ATM cells per second). Compare and contrast this to your answer in (a).

(c) UTOPIA-1 is an 8-bit-wide parallel link for carrying OC-3 traffic; UTOPIA-2 is a 16-bit-wide parallel link for carrying OC-12 traffic. Besides the 8 or 16 "data" wires of a UTOPIA link, there are a couple of other, parallel "control" wires. Using the control wires, UTOPIA links have the capability to encode, among others, an "idle" state, which means "no valid word is being carried during the current clock cycle". Thus, a UTOPIA link may easily run (and often does run) at a rate that is quite higher than the actual data throughput being carried. This is done by running the UTOPIA link with a clock whose frequency is higher than the minimum frequency required for the link to carry the needed throughput; if the clock is, say, 20% faster than the minimum required, then the UTOPIA link will be idle during 20% of its clock cycles.

Calculate the minimum UTOPIA clock frequency required in the following cases: *(c1)* UTOPIA-1 carrying the entire contents (overhead and payload) of an OC-3 link (this is an imaginary situation); *(c2)* UTOPIA-1 carrying the payload but not the overhead of an OC-3 link (this is the usual situation); *(c3)* UTOPIA-2 carrying the entire contents of an OC-12 link; *(c4)* UTOPIA-2 carrying the payload only of an OC-12 link.

2.3 Gigabit Ethernet Throughput

A popular physical-layer technology for carrying Gigabit Ethernet is the "8B/10B encoding" (same as the encoding used in Fibre Channel) (even though capital, the B's in this name mean "bit" and not "byte"!). This scheme encodes each 8-bit byte (plus a few control symbols) using a 10-bit code word.

(a) The "useful user bit" rate of Gigabit Ethernet is 1 Gbps. This refers to unencoded user data. When Gigabit Ethernet is carried over an 8B/10B physical-layer, calculate the "coded bit" rate that will appear on the wire/fiber (i.e. how many bits per second of encoded data must be transmitted on the wire/fiber).

(b) Packets ("frames") carried over poin-to-point Gigabit Ethernet links must be separated from each other by a 12-byte (minimum) "interframe gap", and each packet must be preceeded by an 8-byte "preamble". The shortest Gigabit Ethernet packet has a size of 64 bytes. Calculate the maximum possible "packet throughput" of the above link, when carrying minimum-sized packets, in Mpps (million packets per second), MBytes/s, and Mbps (i.e. the throughput of (unencoded) bits inside packets only, not bits in interframe gaps and preambles.

(c) The longest Gigabit Ethernet packet has a size of 1518 bytes. Answer the same question as (b) for maximum-sized packets.

2.4 Cut-through when Port Rates differ

Consider a switch with some OC-12 links and some gigabit Ethernet links. The OC-12 (actually "OC-12c") links carry ATM traffic, with a peak throughput equal to what you calculated in exercise 2.2 above. The gigabit Ethernet links carry packets at the rate indicated in exercise 2.3 above. We wish the switch to provide cut-through, but, because of rate mismatches among its ports, cut-through transmission cannot always start "right away"; we wish to calculate the worst such required delay.

Consider the situation where a packet destined to a gigabit Ethernet output arrives through an OC-12 input. The maximum-sized Ethernet packet is 1518 bytes long, and, when arriving through an ATM port, it arrives segmented into 32 ATM cells (1518 bytes / 48 payload-bytes/cell = 31.63 cells). The first 1518 bytes in the (48-byte) payloads of the 32 cells are precisely the 1518 bytes that constitute the Ethernet packet. Assume that we are guaranteed that these 32 cells will always arrive back-to-back, without any idle cells in-between them (when ATM traffic is carried over SONET links, cells are always transmitted back-to-back, without any "spacing" between them other than the SONET overhead bytes; however, in general, not all cells need to be valid --idle cells can be freely injected into the SONET payload in the general case, but not in our case between same-packet cells).

How soon, in nanoseconds, after the arrival of the first payload byte of the first ATM cell of the above 32-cell train can we transmit the first byte of the 8-byte gigabit Ethernet preamble? Obviously, right after the 8-byte preamble is transmitted, the 1518 bytes of the Ethernet packet must be transmitted, back-to-back, at the gigabit Ethernet rate, without any "hiccups". Our switch circuits are able to transmit any given byte within 100 ns at the earliest, after that byte is received at an input port.

2.5 Synchronous DRAM (SDRAM) Throughput

Synchronous DRAM (SDRAM) chips have an interface that, after the initial delay of accessing the first bit(s) at a new (random) address, can supply additional consecutive bits at a very high rate (clock frequency). Assume in this exercise a 400 MHz clock and "DDR" (double-data-rate) signaling, yielding 800 Mbps per pin for the data pins. Consider a memory system made of such SDRAM chips, that has a 64-bit-wide data bus.

(a) What would be the (theoretical) peak throughput of this memory system, in Gbps and in GBytes/s, if it could be supplying continuously consecutive bits and never access new (random) addresses?

(b) Assume that every access to a new location (address) costs 16 clock cycles (of the 400 MHz clock) for the first 128-bit word, plus 1 additional clock cycle for each additional consecutive 128-bit word (when using DDR, a 64-bit bus delivers 128 bits per clock cycle). Calculate the memory throughput in the following situations: (b1) continuous accesses to 16-Byte words that are located at a different address each; (b2) continuous accesses to 64-Byte blocks (i.e. blocks of 4 consecutive words each), where each block is located at a different address; (b3) continuous accesses to 256-Byte blocks.

2.6 Throughput - Time Window - Buffer Space

(a) Consider a 16x16 switch with 10 Gb/s links, that we want to be able to withstand "heavy" output contention for periods of up to 200 ms (milli-seconds). Define "heavy" output contention as the situation where up to 11 of the 16 input links may be injecting, at peak rate, traffic destined to a single output link; assuming that that output link ejects traffic at its peak capacity, the overload corresponds to 10 times a link capacity. What buffer size is needed inside the switch to be able to withstand such stress periods without losing any packet?

(b) A source of traffic transmits a "regular" stream of data at an average rate of 10 Mb/s. However, at some points in time, this source has a "new page" of data, of size 5000 Bytes, that it wants to transmit "at once", in addition to the normal stream data (assume 8b/10b encoding, hence the page size is 50,000 bits). Such transmissions occur at the source's peak rate, which is 510 Mb/s, i.e. 500 Mb/s over the normal stream rate of 10 Mb/s. What is the (time) duration of each such burst of traffic that the source injects into the network?

(c) A 64 KByte buffer is placed "in the middle" of a 1 Gb/s network link, and is used to smooth out variations between incoming and outgoing link throughput without any packet losses; assume that the link is loaded at about 50%. If we measure the incoming link throughput and the outgoing link throughput, both of them averaged over a one-second time window, by up to how much percent can these two measurements differ from each other?

(d) Assume that the average occupancy of the above buffer (c) is 8 KBytes. What is the average delay undergone by the packets going through this buffer?

2.7 Turn-around Overhead

In a shared-medium-based LAN (ehternet style), we wish to keep the turn-around overhead (as defined in class) down to 5 % of the useful transmission time when the packet size is 80 bytes. What is the maximum allowable distance between transmitters in order to achieve this when the transmission rate is *(a)* 10 Mbps; *(b)* 100 Mbps; *(c)* 1 Gbps; and *(d)* 10 Gbps. Transmission is bit-serial, and the speed of light in the LAN medium is 200 Mm/s.