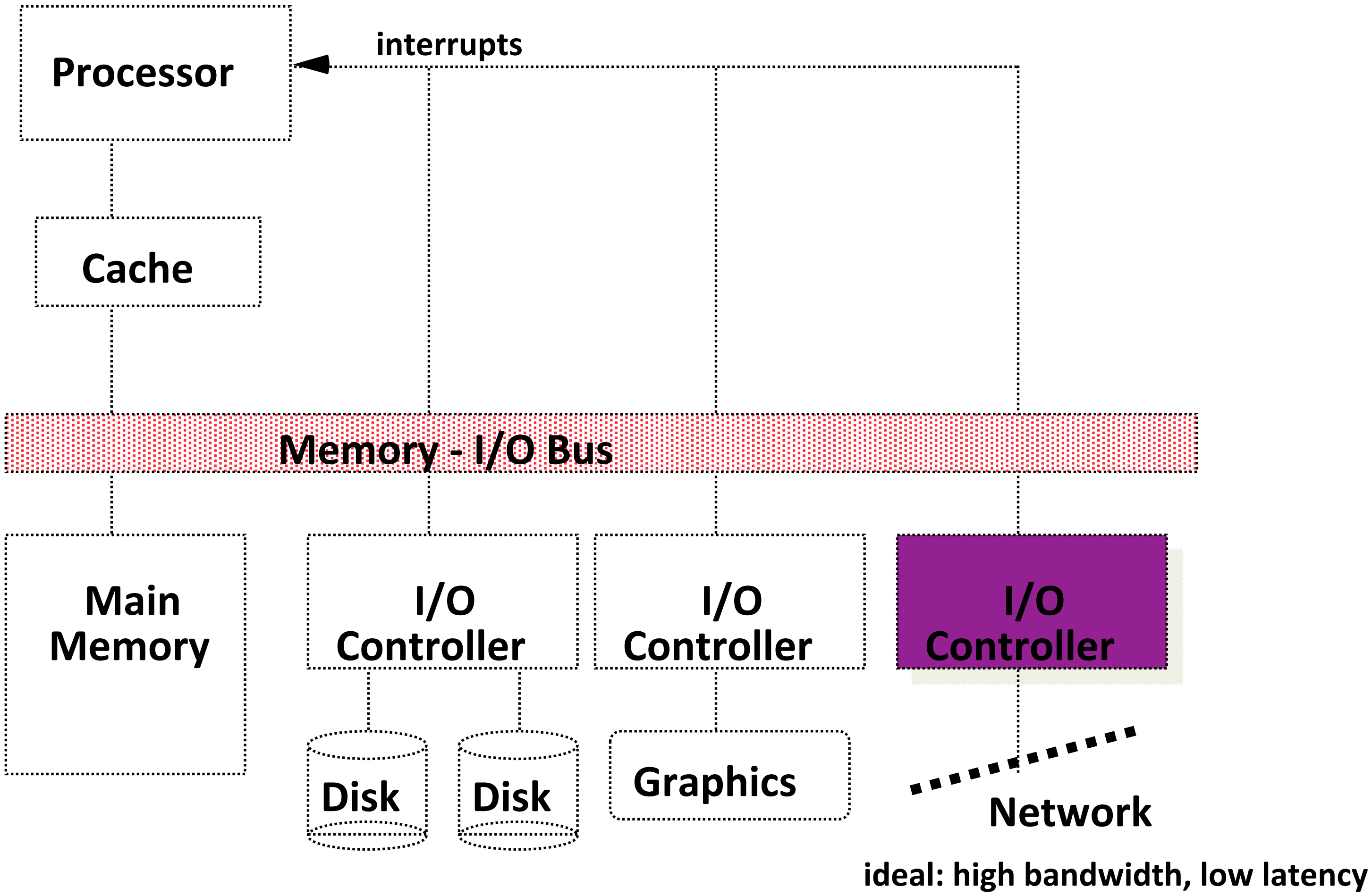
Lecture 10
Network & Interconnect

Assoc. Prof Dr. Norafida Ithnin

Advanced Computer System & Architecture

MCC 2313

Network Interface

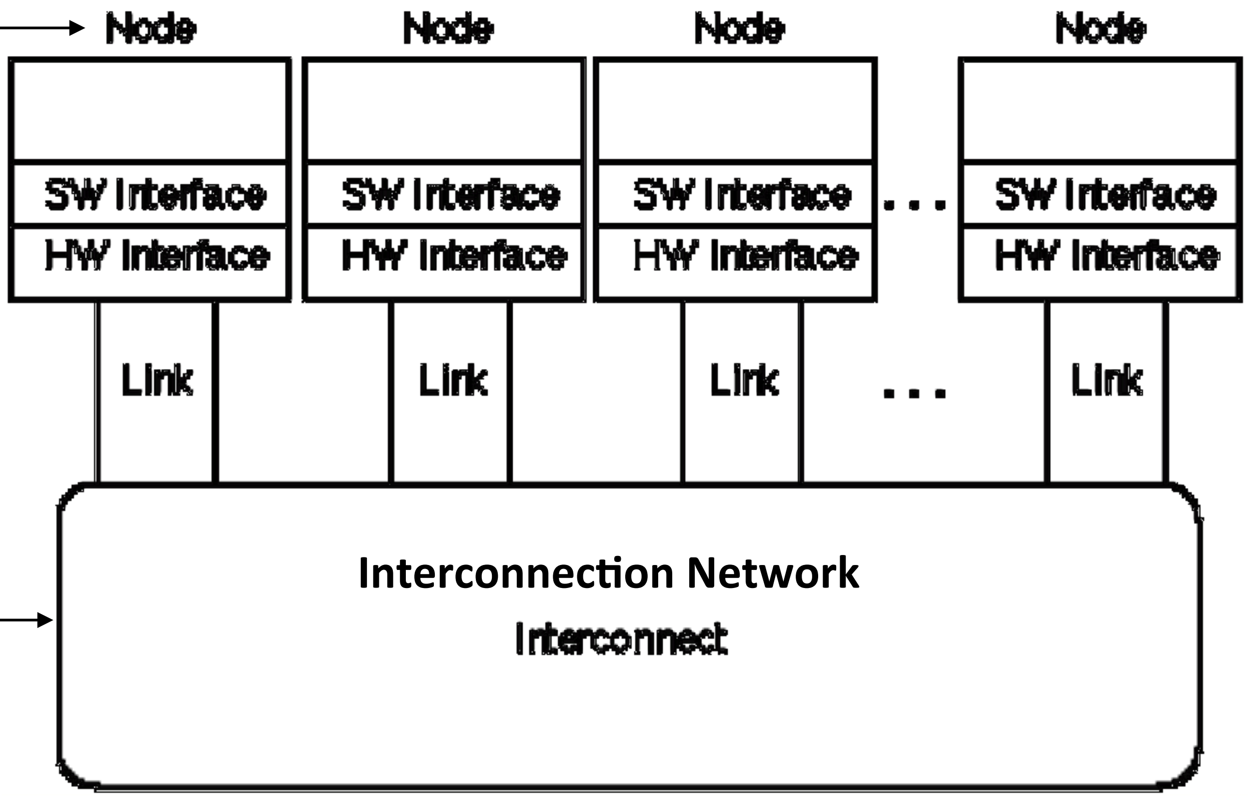


Networks

- Goal: Communication between computers
- Eventual Goal: Treat a collection of computers as one large computer: distributed resource sharing
- An interconnection network is used to allow computers, called nodes, to communicate with one another.
  - Massively parallel processor (MPP) network (e.g., CM5)
    - Thousands of nodes, less than 25 meters apart
  - Local area network (LAN) (e.g., Ethernet)
    - Hundreds of computers, up to a few kilometers apart
  - Wide area network(WAN) (e.g., ATM)
    - Several thousands of computers, several thousand kilometers apart

Network Overview

a.k.a.
end systems,
hosts

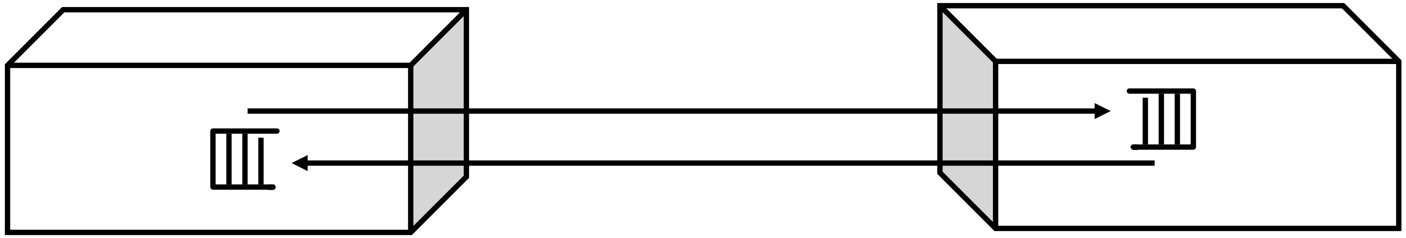


a.k.a.
network,
communication
subnet

Network Metrics

- Facets people talk a lot about:
  - direct vs. indirect
  - topology (e.g., bus, ring, mesh)
  - routing algorithms (how is message passed)
  - wiring (e.g., choice of media - copper, coax, fiber)
- What really matters:
  - latency
  - bandwidth
  - cost
  - reliability

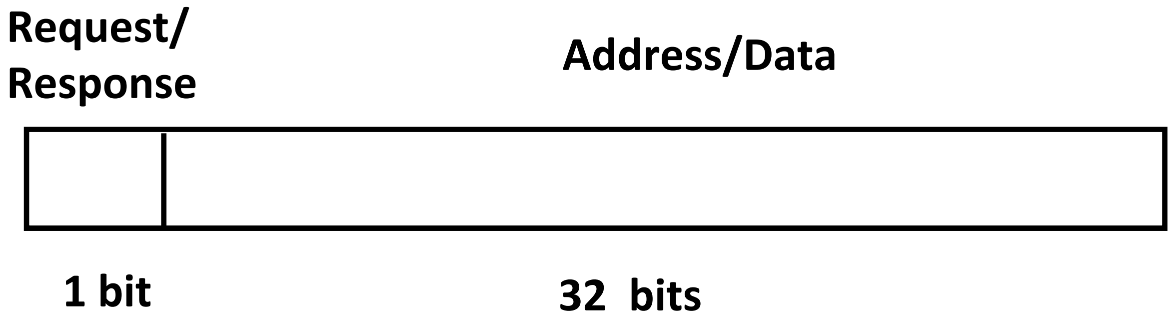
A Simple Network



- Starting Point: Send bits between 2 computers
- Queue (FIFO) on each end
- Information sent called a "message"
- Can send info both ways ("Full Duplex")
- Rules for communication called a protocol
  - Request : Send address of desired data
  - Response : Send requested data
  - Packet : Name for standard group of bits making up message

A Simple Example

- What is the format of the message?
  - Fixed length? Number bytes?



0: Please send data from Address
1: Packet contains data corresponding to request

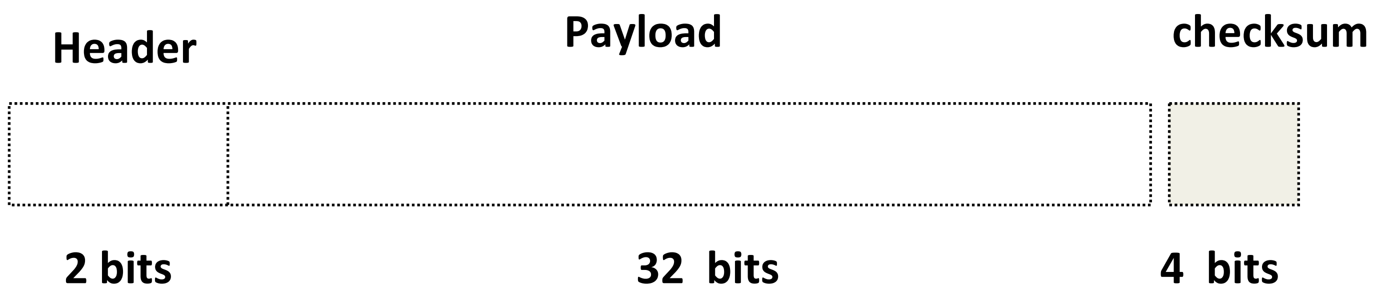
- **Header/Trailer: information to deliver a message**
- **Payload: data in message (1 word above)**

Questions About Simple Example

- What if more than 2 computers want to communicate?
  - Need computer address field (destination) in header
- What if packet is garbled in transit?
  - Add error detection field in packet (e.g.,ckecksum)
- What if packet is lost?
  - More elaborate protocols to detect loss
- What if multiple processes/machine?
  - Queue per process - need to indicate which process
- What if want larger or variable-length packet?
  - Some messages may be thousands of bytes
- Questions such as these lead to more complex protocols and packet formats

A Simple Example Revisted

- A more complex packet format might include a longer header and a checksum



00: Request—Please send data from Address
01: Reply—Packet contains data corresponding to request
10: Acknowledge request
11: Acknowledge reply

Software to Send and Receive

- **SW Send steps**
  - **1: Application copies data to OS buffer**
  - **2: OS calculates checksum, starts timer**
  - **3: OS sends data to network interface HW and says start**
- **SW Receive steps**
  - **3: OS copies data from network interface HW to OS buffer**
  - **2: OS calculates checksum, if matches send ACK; if not, *deletes message* (sender resends when timer expires)**
  - **1: If OK, OS copies data to user address space and signals application to continue**
- **Sequence of steps for SW: protocol**
  - **Example similar to UDP/IP protocol in UNIX**

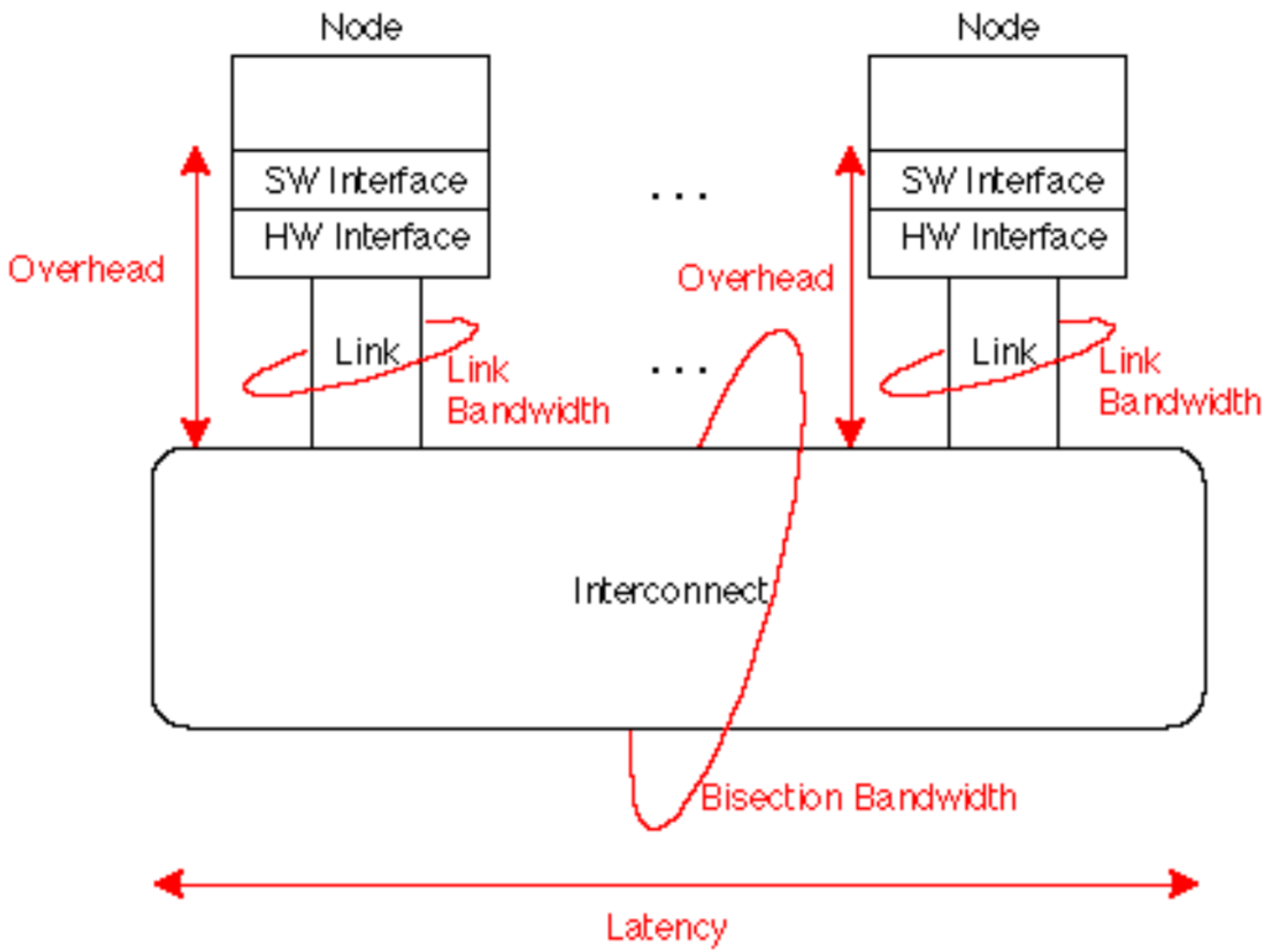
Network Performance Metrics

- Several metrics are used for network performance
  - Bandwidth : Maximum rate at which the network can propagate information, once the message enters the network (Mbits/sec)
  - Transmission time : Time for message to pass through network

    transmission time = (message size)/bandwidth
  - Time of flight : Time for first bit of message to arrive at receiver
  - Transport latency : Time message spends in network

    transport latency = transmission time + time of flight
  - Sender overhead : Time for processor to inject a message into the interconnection network
  - Receiver overhead : Time for processor to pull the message from the interconnection network
  - Total latency = sender overhead + trans. latency + receiver overhead

Network Performance Measures

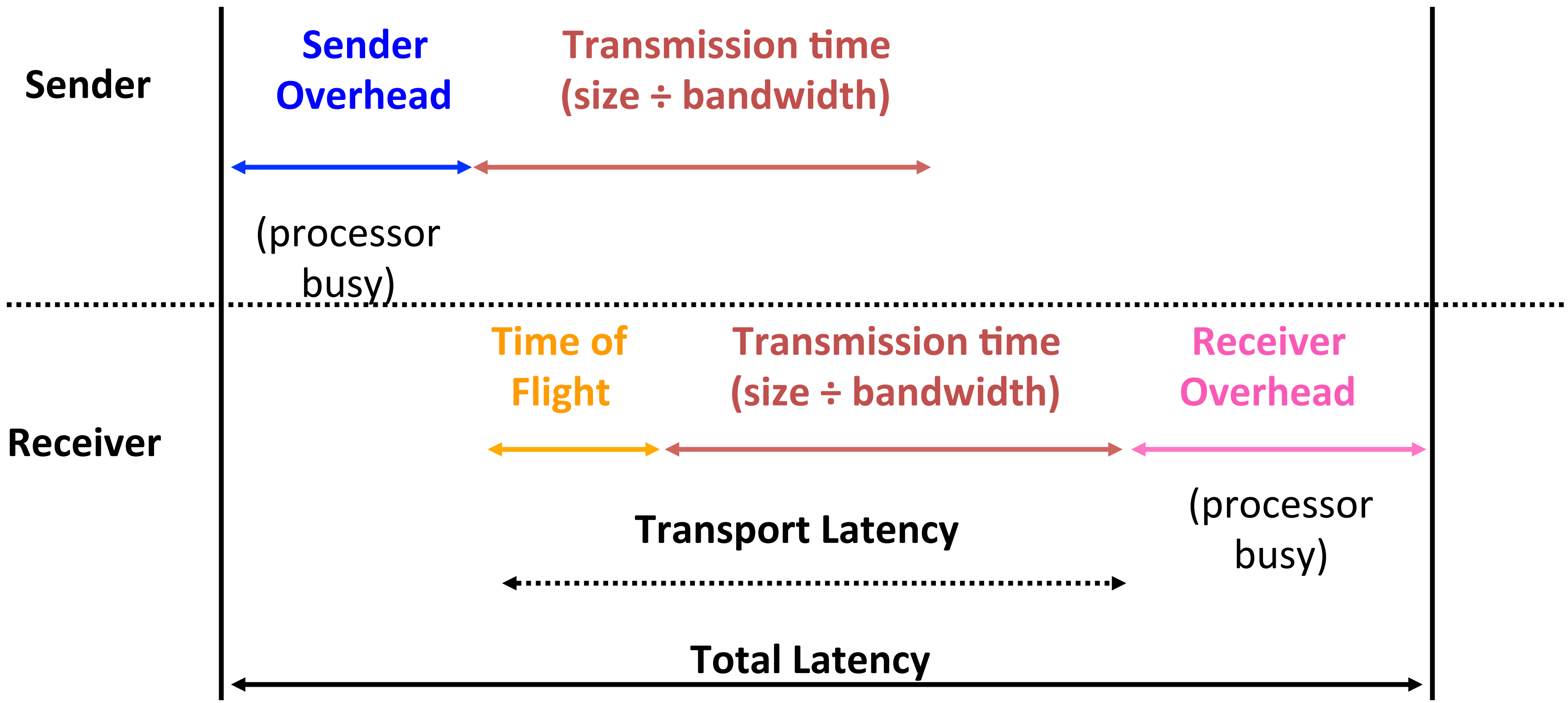


Overhead: latency of interface vs. Latency: network

What is latency in general?

- Latency: How long a message takes to travel from one end of the network to another
- .Speed of light
- propagation delay
- vacuum cable fiber
- 3×10^8 m/sec 2.3×10^8 m/sec 2.0×10^8 m/sec

Universal Performance Metrics



Total Latency = Sender Overhead + Time of Flight + Message Size ÷ BW + Receiver Overhead

Includes header/trailer in BW calculation?

Example Performance Metrics

Interconnect	*MPP*	*LAN*	*WAN*
Example	CM-5	Ethernet	ATM
Link Bandwidth	20 MB/s	10 MB/s	10 MB/s
Transport Latency	5 μsec	15 μsec	50 to 10,000 μs
HW Overhead to/from	0.5/0.5 μs	6/6 μs	6/6 μs
SW Overhead to/from	1.6/12.4 μs	200/241 μs	207/360 μs

(TCP/IP on LAN/WAN)

Software overhead dominates in LAN, WAN

Total Latency Example

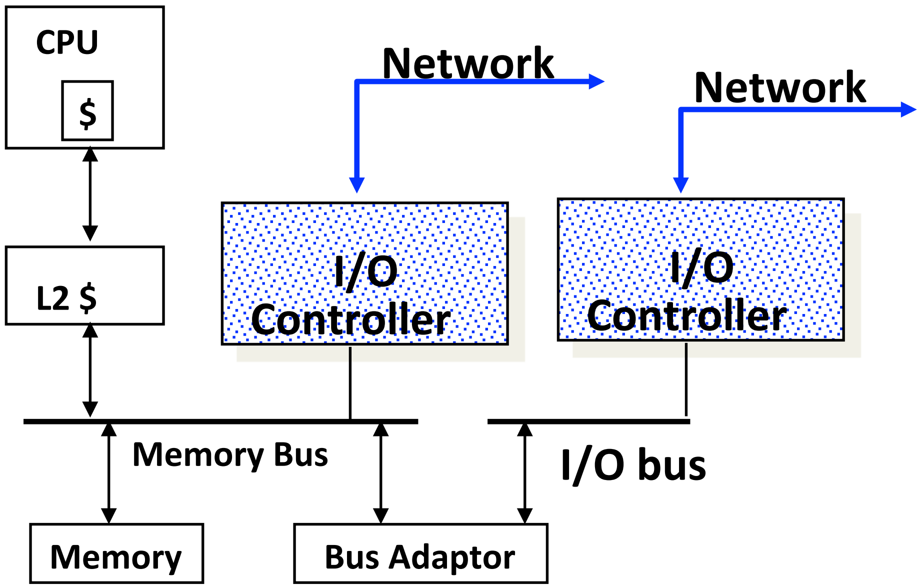
- 10 Mbit/sec., sending overhead of 230 µsec & receiving overhead of 270 µsec with 50% of speed of light
- A 1000 byte message (including the header), allows 1000 bytes in a single message.
- 2 situations: distance 100 m vs. 1000 km
- Speed of light = 299,792.5 km/sec
- $\text{Latency}_{0.1km}$ = 230 + 0.1km / (50% x 299,792.5) + 1000 x 8 / 10 + 270
- $\text{Latency}_{0.1km}$ = 230 + 0.67+ 800 + 270 = 1301 µsec
- Latency_{1000km} = 230 + 1000 km / (50% x 299,792.5) + 1000 x 8 / 10 + 270
- Latency_{1000km} = 230 + 6671+ 800 + 270 = 7971 µsec
- Long time of flight => complex WAN protocol

Simplified Latency Model

- Total Latency - Overhead + Message Size / BW
- Overhead = Sender Overhead + Time of Flight + Receiver Overhead
- Example: show what happens as vary
  - Overhead: 1, 25, 500 µsec
  - BW: 10,100, 1000 Mbit/sec (factors of 10)
  - Message Size: 16 Bytes to 4 MB (factors of 4)
- If overhead 500 µsec, how big a message > 10 Mb/s?

HW Interface Issues

- Where to connect network to computer?
  - Cache consistent to avoid flushes? (=> memory bus)
  - Latency and bandwidth? (=> memory bus)
  - Standard interface card? (=> I/O bus)
  - MPP => memory bus; LAN, WAN => I/O bus



ideal: high bandwidth, low latency, standard interface

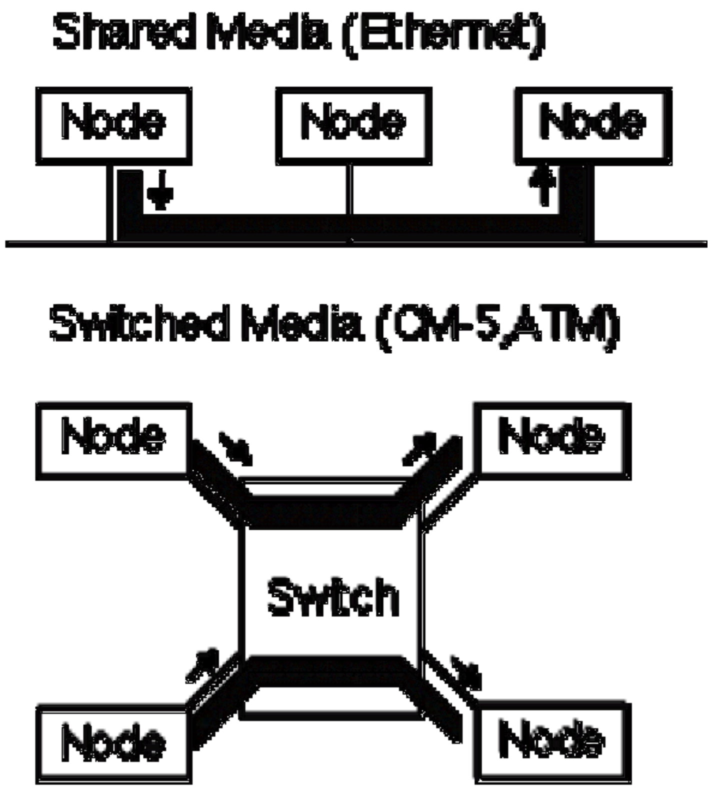
UTM
UNIVERSITI TEKNOLOGI MALAYSIA

Implementation Issues

Interconnect	*MPP*	*LAN*	*WAN*
Example	CM-5	Ethernet	ATM
Maximum length	25 m	500 m;	copper: 100 m
	between nodes		
	optical: 2 km—25 km		
Number data lines	4	1	1
Clock Rate	40 MHz	10 MHz	155.5 MHz
Shared vs. Switch	Switch	Shared	Switch
Maximum number of nodes	2048	254	> 10,000
Media Material	Copper	Twisted pair copper wire or coaxial	Twisted pair copper wire or optical fiber
	cable		

Implementation Issues

- Advantages of Serial vs. Parallel lines:
  - No synchronizing signals
  - Higher clock rate and longer distance than parallel lines
  - 60 MHz x 256 bits x 0.5 m vs. 155 MHz x 1 bit x 100 m
- Switched vs. Shared Media
  - Switched : many messages at same time
  - Shared : one message at a time



Connecting to the Computer

- Should network interface to memory bus or I/O bus? Why?
  - MPPs plug into memory bus
  - LANs and WANs plug into I/O bus
- How is the receiver notified of a message?
  - Poll network waiting for it to arrive
  - Be interrupted when message arrives
  - Interrupts work better when fewer messages
- General guidelines
  - Avoid invoking the operating system (context switch)

Summary: Interconnections

- Communication between computers
- Packets for sending information: header + payload
- Protocols to cover normal and abnormal events
- Performance issues: overhead, latency, bandwidth
- Implementation issues: length, width, media
- Topologies: many to chose from, but SW overheads make them look the alike

Discussions

Example1

- Assume a network with a bandwidth of 1000 Mbits/second has a sending overhead of 80 microseconds and a receiving overhead of 100 microseconds. Assume two machines. One wants to send a 10000-byte message to the other (including the header), and the message format allows 10000 bytes in a single message.

Example1 cont.

- Let's compare SAN, LAN, and WAN by changing the distance between the machines. Calculate the total latency to send the message from one machine to another in a SAN assuming they are 10 meters apart. Next, perform the same calculation but assume the machines are now 500 meters apart, as in a LAN. Finally, assume they are 1000 kilometers apart, as in a WAN.

Solution

- The speed of light is 299,792.5 kilometers per second in a vacuum, and signals propagate at about 63% to 66% of the speed of light in a conductor. Since this is an estimate, in this chapter we'll round the speed of light to **300,000 km/sec**, and assume we can achieve two-thirds of that in a conductor. Hence, we can estimate time of flight. Let's plug the parameters for the short distance of a SAN into the formula:

Solution

Total Lat. = Tx overhead + Time of flight + (Message size/Bandwidth)+ Rx overhead

= 80us + [0.01km / (2/3 x 30x105km/s)] + 10000byte/(1000MB/s) + 100us

= 80us + 0.05us + 80us + 100us

= 260us