

Communication (I)

Kai Huang



Amazon testing drones for deliveries

- The drones, called Octocopters, could deliver packages weighing up to 2.3kg to customers within 30 minutes of them placing the order (Jeff Bezos)



<http://www.bbc.co.uk/news/technology-25180906>



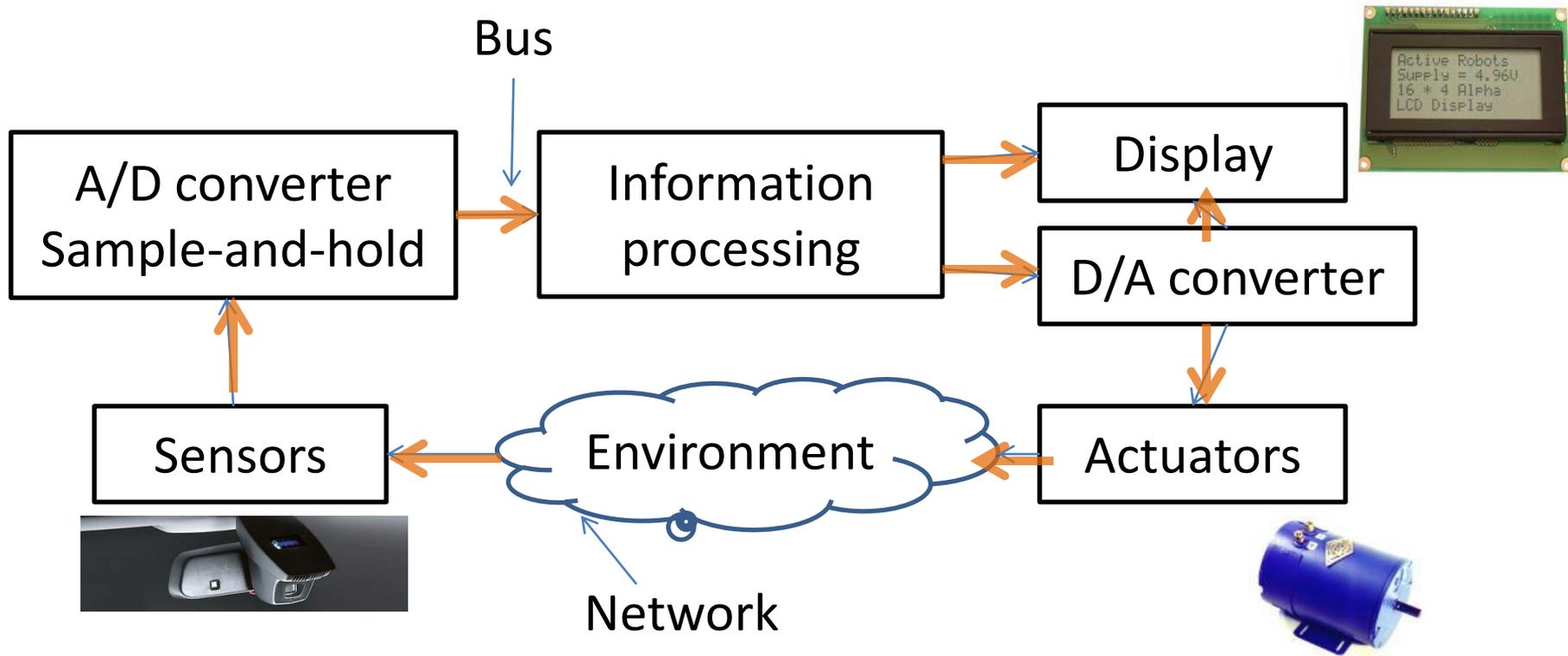
Outline

- Bus basics
- Multiple Master Bus
 - Centralized arbitration
 - Distributed arbitration
- Network-on-Chip
- Examples



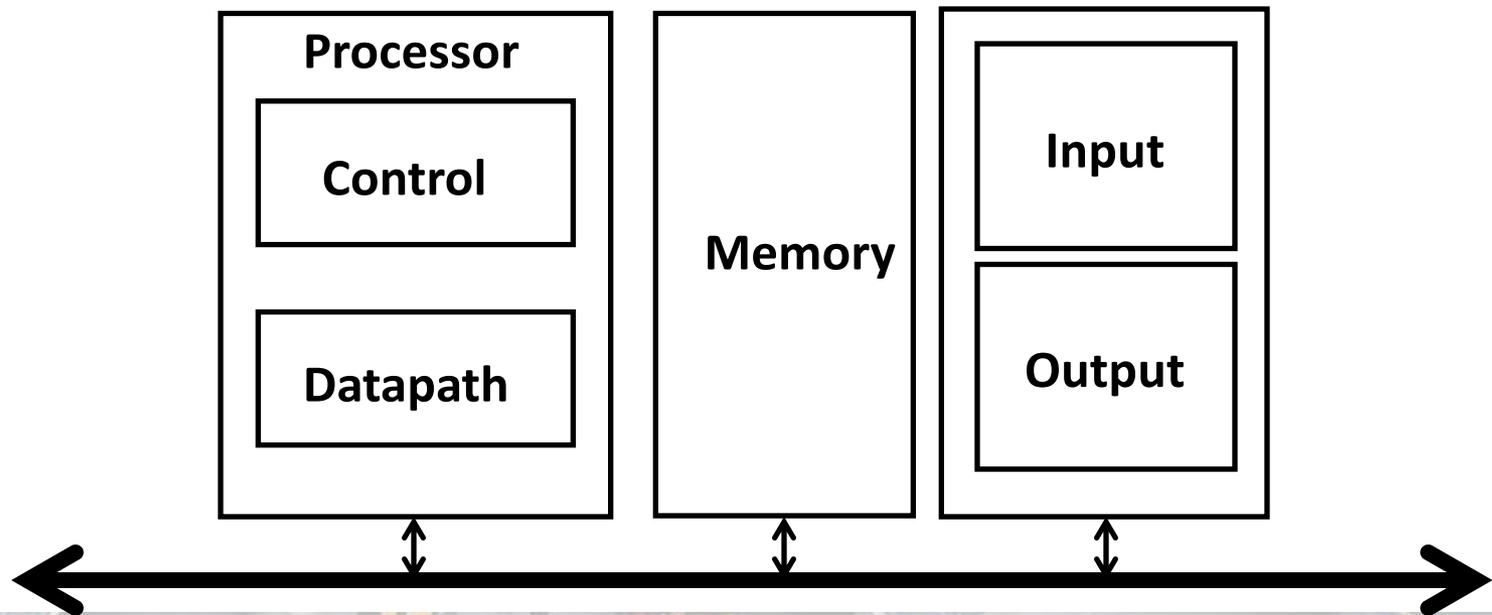
Communication

- Is everywhere ...



What is Bus?

- A bus is a shared communication link, connecting I/O to processor and memory
- It uses one set of wires to connect multiple subsystems



Why Bus ?

■ Pros

○ Versatility:

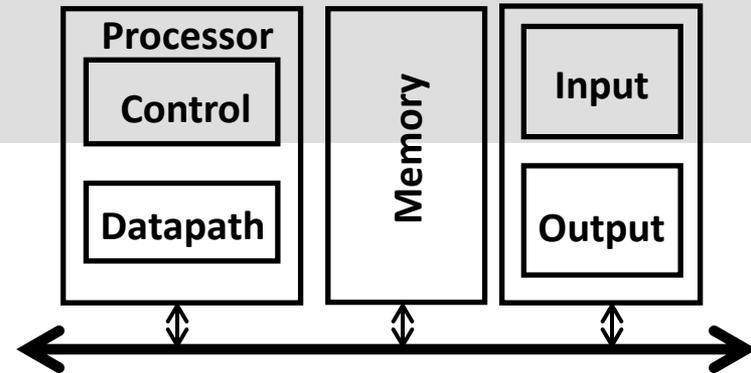
- New devices can be added easily
- Peripherals can be moved between computer systems that use the same bus standard

○ Low Cost:

- A single set of wires is shared in multiple ways

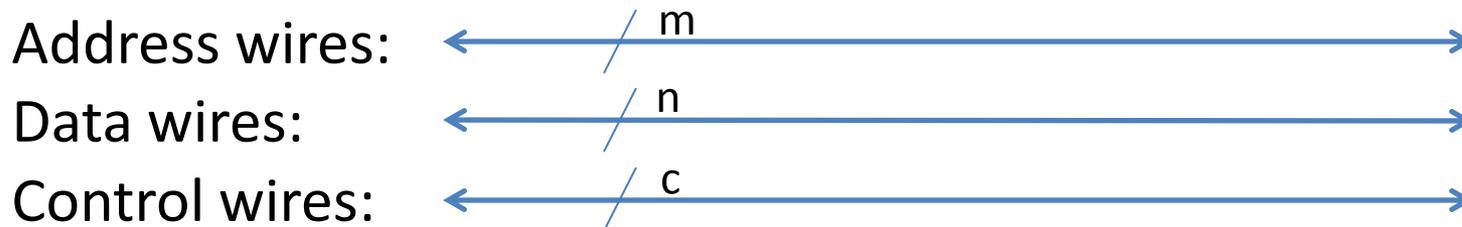
■ Cons

- It creates a communication bottleneck

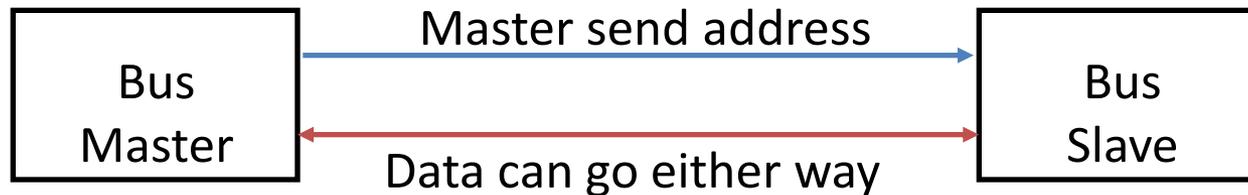


Generic Bus Structure

- Set of wires
 - Control wires:
 - Signal requests and acknowledgments
 - Indicate what type of information is on the data lines
 - Data/ address Wires:
 - carry information between the source and the destination:
 - Complex commands
- A bus transaction includes two parts:
 - Sending the address
 - Receiving or sending the data

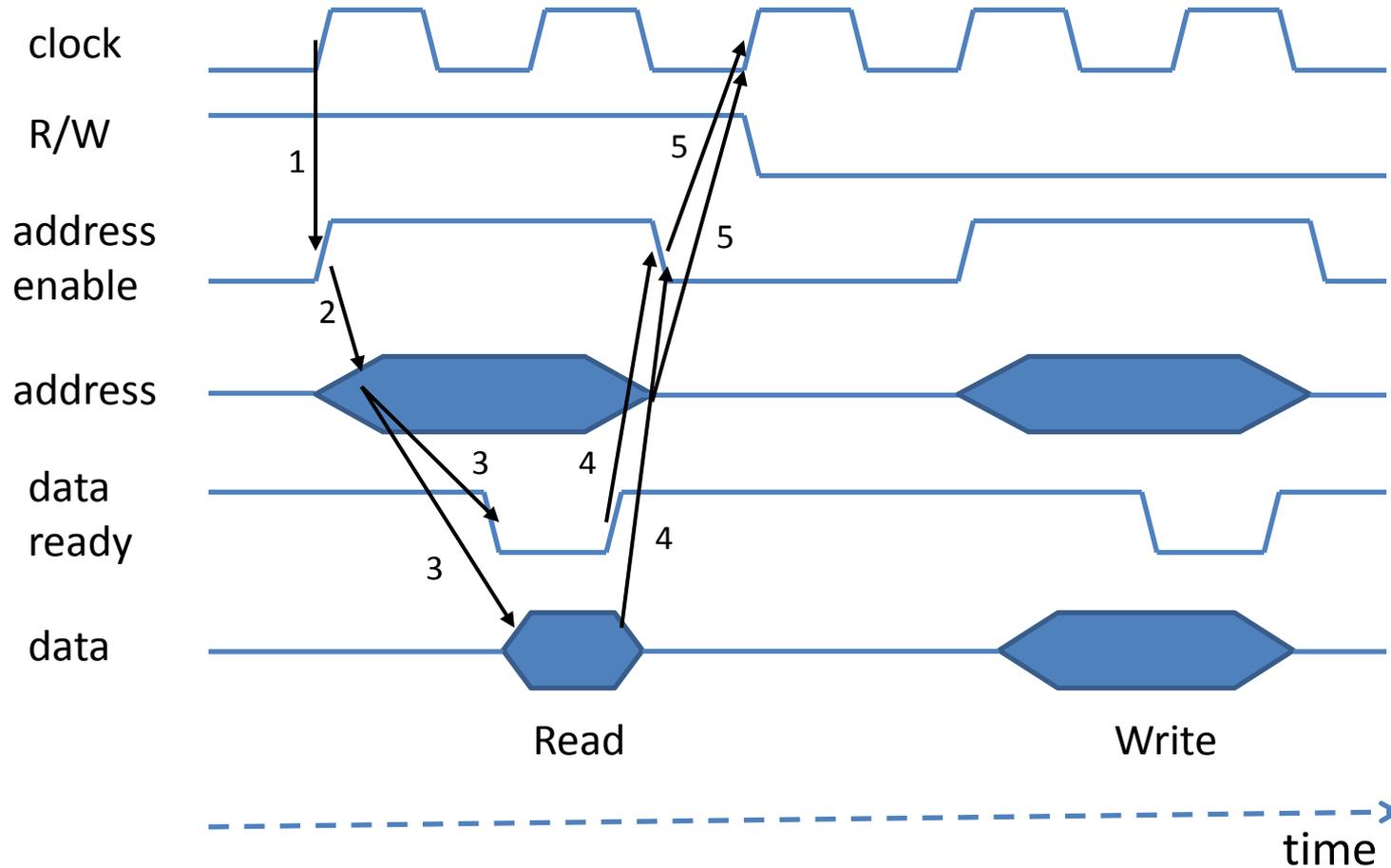


Master versus Slave



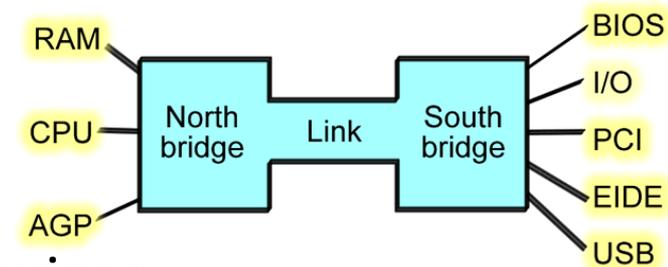
- A bus transaction includes two parts:
 - Sending the address
 - Receiving or sending the data
- Master is the one who starts the bus transaction by:
 - Sending the address
- Slave is the one who responds to the address by:
 - Sending data to the master if the master ask for data
 - Receiving data from the master if the master wants to send data

Typical Bus Access Protocol

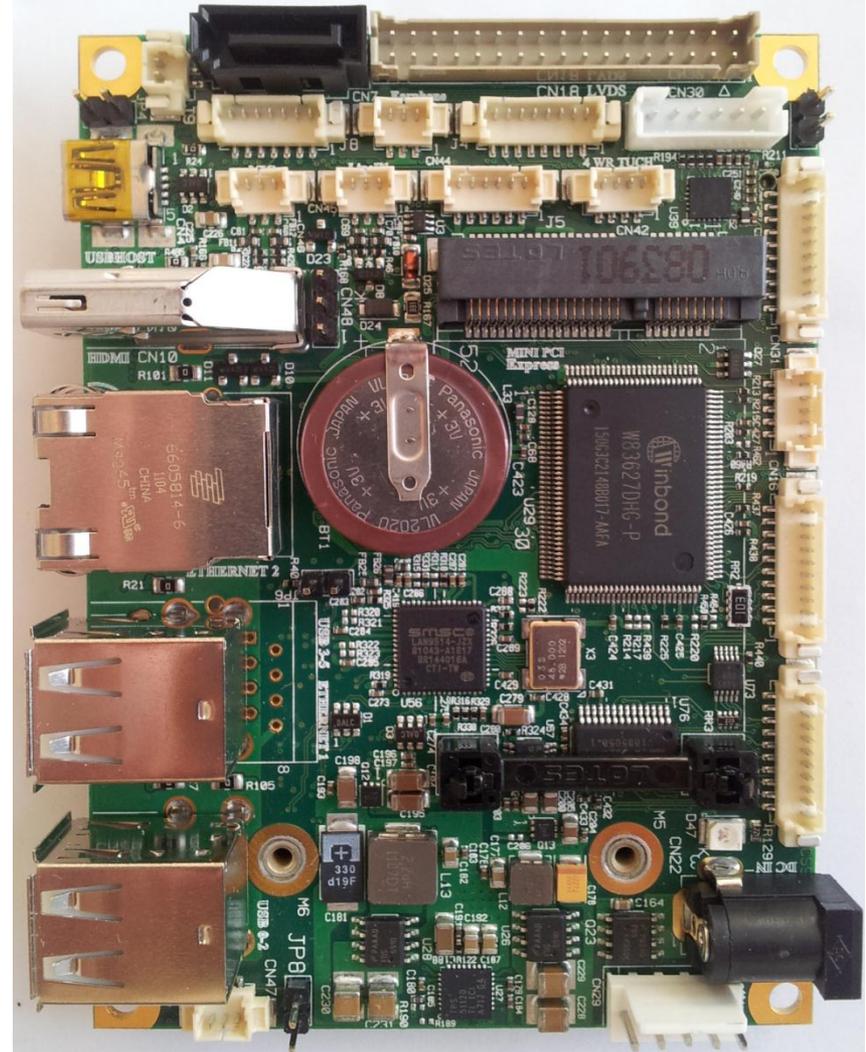
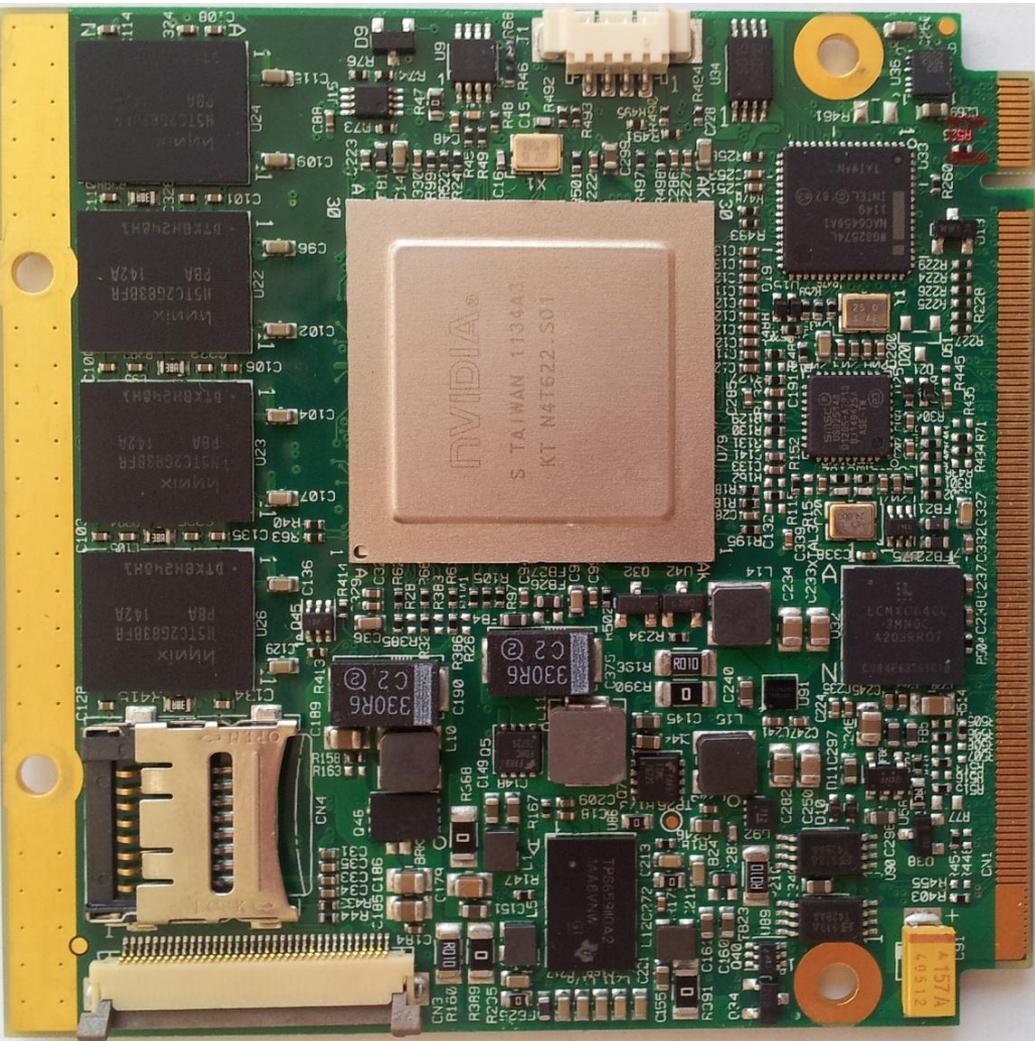


Types of Buses

- Processor-Memory Bus (design specific)
 - Short and high speed
 - Only need to match the memory system
 - Maximize memory-to-processor bandwidth
 - Connects directly to the processor
- I/O Bus (industry standard)
 - Usually is lengthy and slower
 - Need to match a wide range of I/O devices
 - Connects to the processor-memory bus or backplane bus
- Backplane Bus (industry standard)
 - Backplane: an interconnection structure within the chassis
 - Allow processors, memory, and I/O devices to coexist
 - Cost advantage: one single bus for all components



SECOCQ7-Xboard



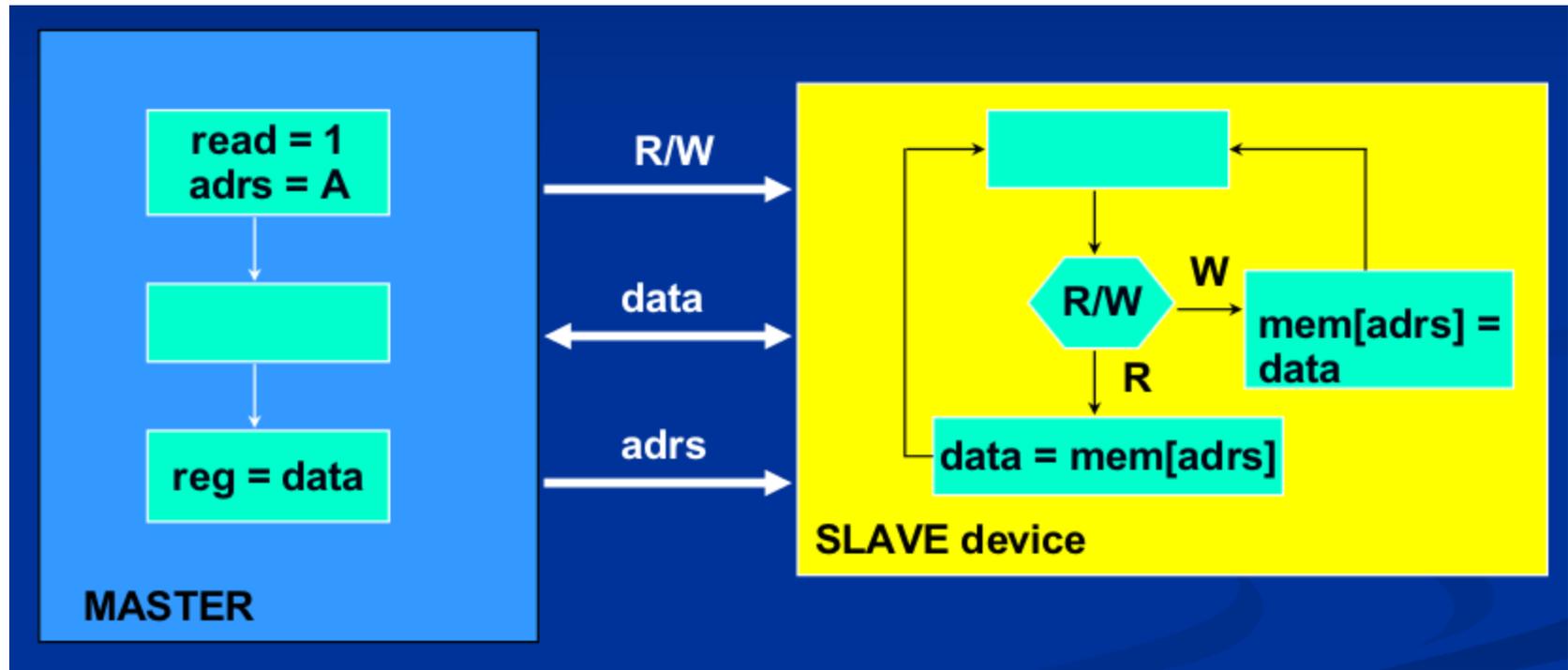
Synchronous and Asynchronous Bus

- Synchronous Bus:
 - Includes a clock in the control lines
 - A fixed protocol for communication that is relative to the clock
 - Advantage: involves very little logic and can run very fast
 - Disadvantages:
 - Every device on the bus must run at the same clock rate
 - To avoid clock skew, they cannot be long if they are fast
- Asynchronous Bus:
 - It is not clocked
 - It can accommodate a wide range of devices
 - It can be lengthened without worrying about clock skew
 - It requires a handshaking protocol



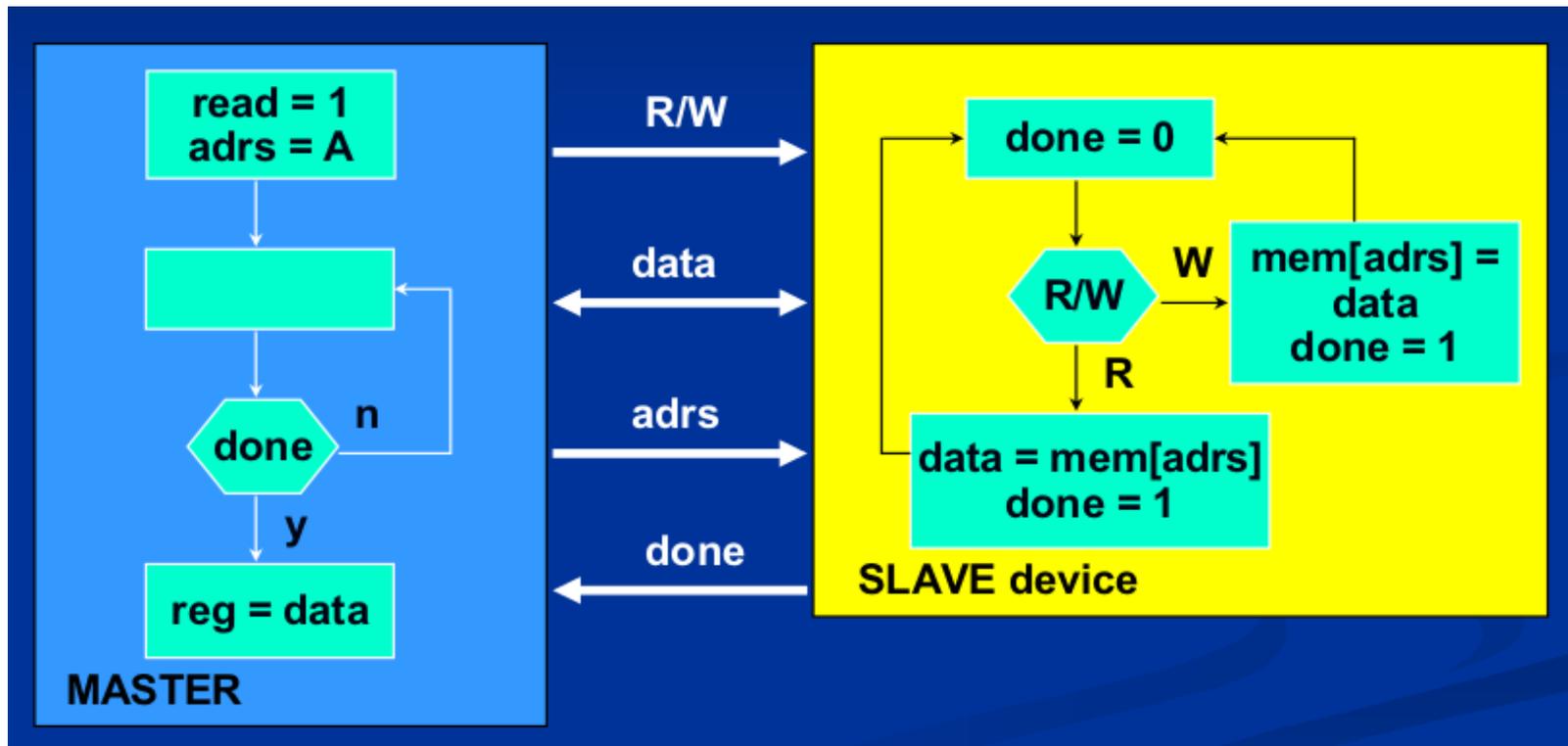
Generic Fixed-delay Access Protocol

- Synchronous Bus:



Generic Variable-delay Access Protocol

- Asynchronous Bus:



Outline

- Bus basics
- **Multiple Master Bus**
 - Centralized arbitration
 - Distributed arbitration
- Network-on-Chip
- Examples



Multiple Masters Simultaneously Access

- Arbiter is used
 - Controls access to the shared bus
 - Uses arbitration policy to select master to grant access to bus
- Bus arbitration schemes usually try to balance two factors:
 - **Bus priority**: The highest priority device should be serviced first
 - **Fairness**: Even the lowest priority device should never be completely locked out from the bus
- Arbitration policy
 - Random Access
 - Centralized Arbitration
 - Round Robin policy
 - Priority policy
 - TDMA policy
 - Distributed Arbitration
 - Carrier Sense Multiple Access (CSMA) / Collision Detection (CD) / Collision Avoidance (CA) / Collision Resolution (CR)

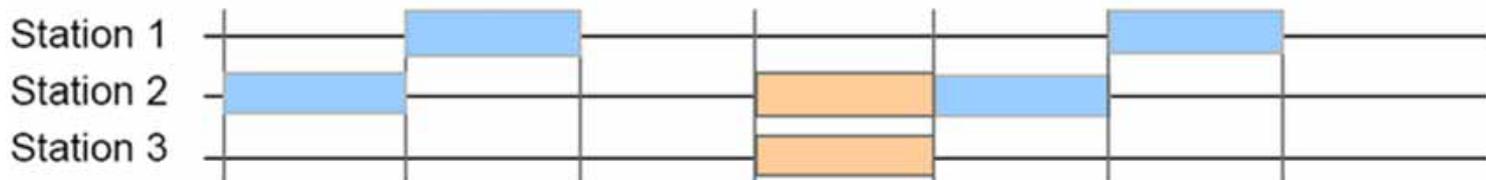


Random Access

- Random access to communication medium
 - no access control; requires low medium utilization



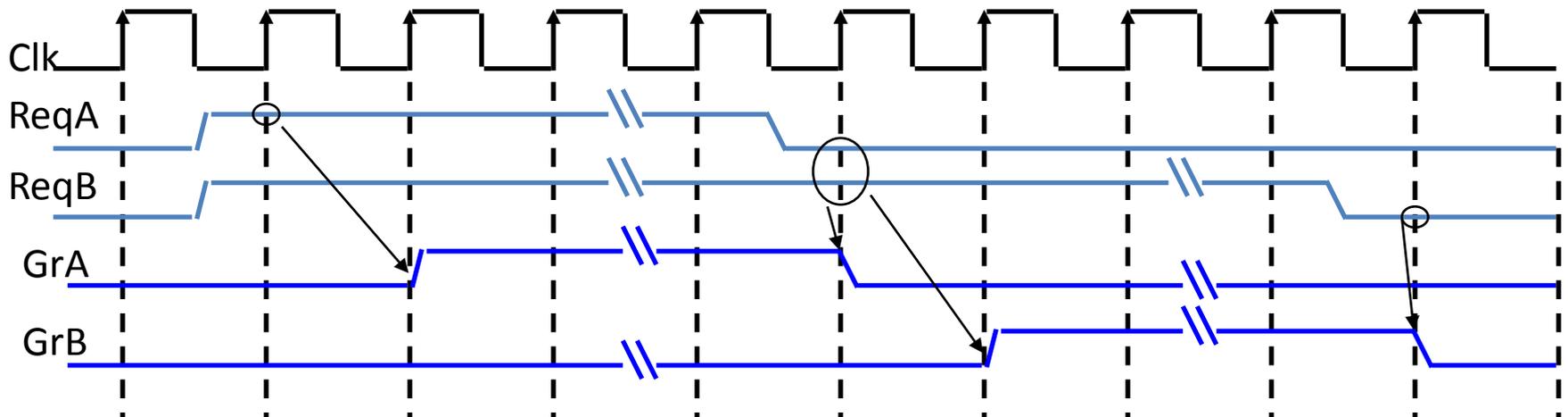
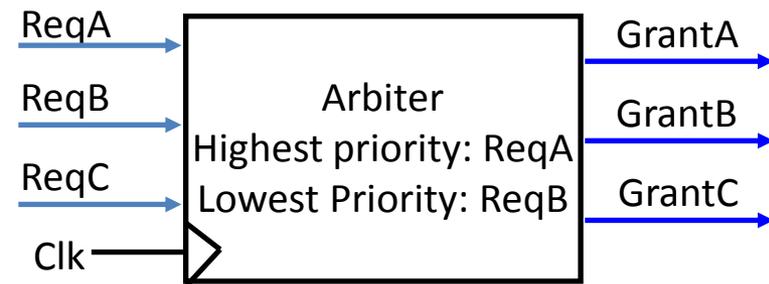
- improved variant: slotted random access



- What is the optimal sending rate p in case of n stations?
 - probability that a slot is not taken by others: $(1 - p)^{n-1}$
 - probability that a station transmits successfully: $P = p \cdot (1 - p)^{n-1}$
 - determine maximum with respect to p : $dP/dp = 0 \rightarrow p = 1/n$

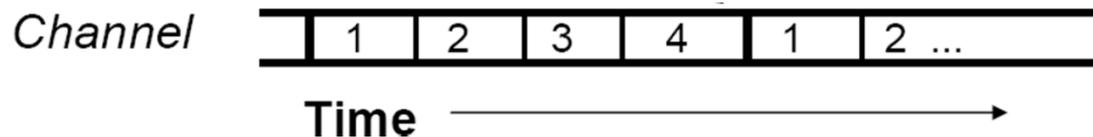
Centralized Arbitration with a Bus Arbiter

- Two important control signals per MASTER, i.e.,
 - bus_request – from MASTER to ARBITER
 - bus_granted – from ARBITER to MASTER
- Minimal change is required if new components are added to the system

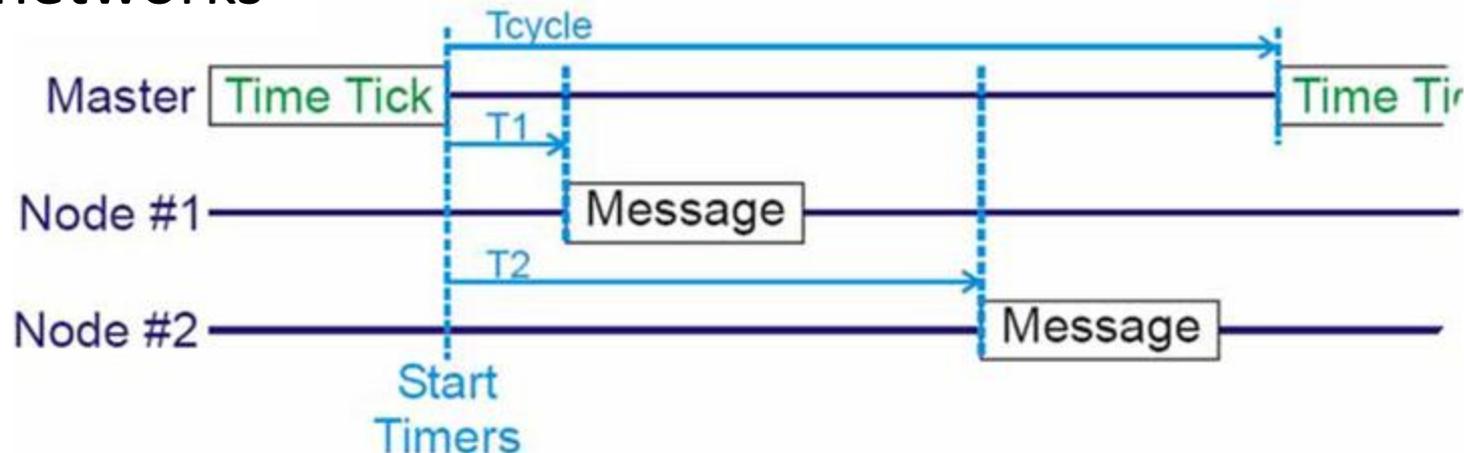


Time Division Multiple Access (TDMA)

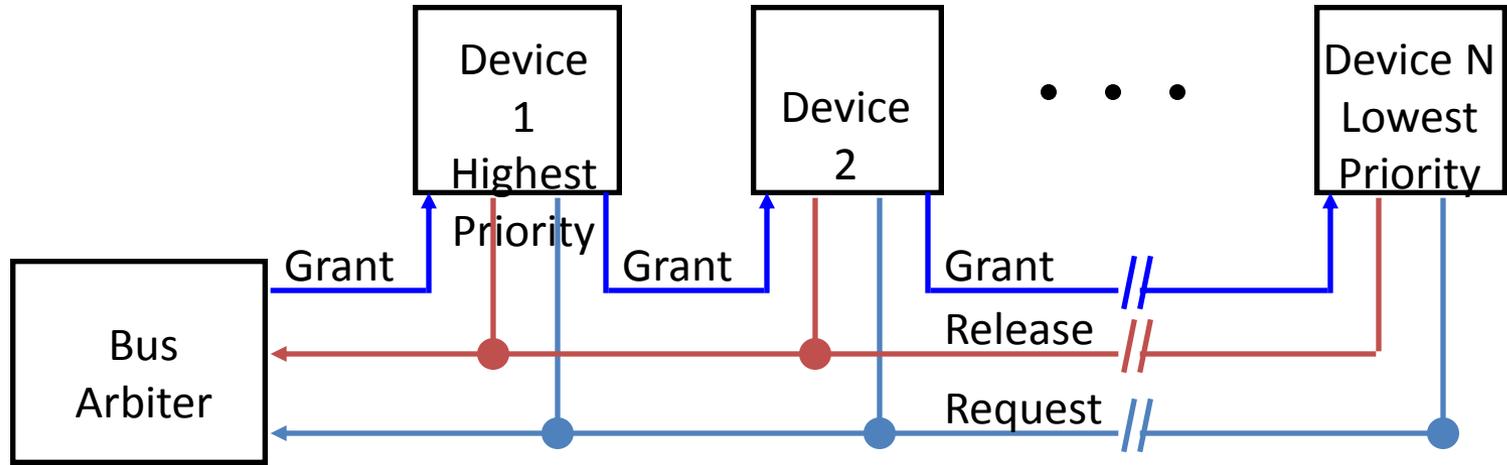
- Communication in statically allocated time slots
- Synchronization among all nodes necessary:
 - periodic repetition of communication frame or



- master node sends out a synchronization frame
- Examples: TTP, static portion of FlexRay, satellite networks



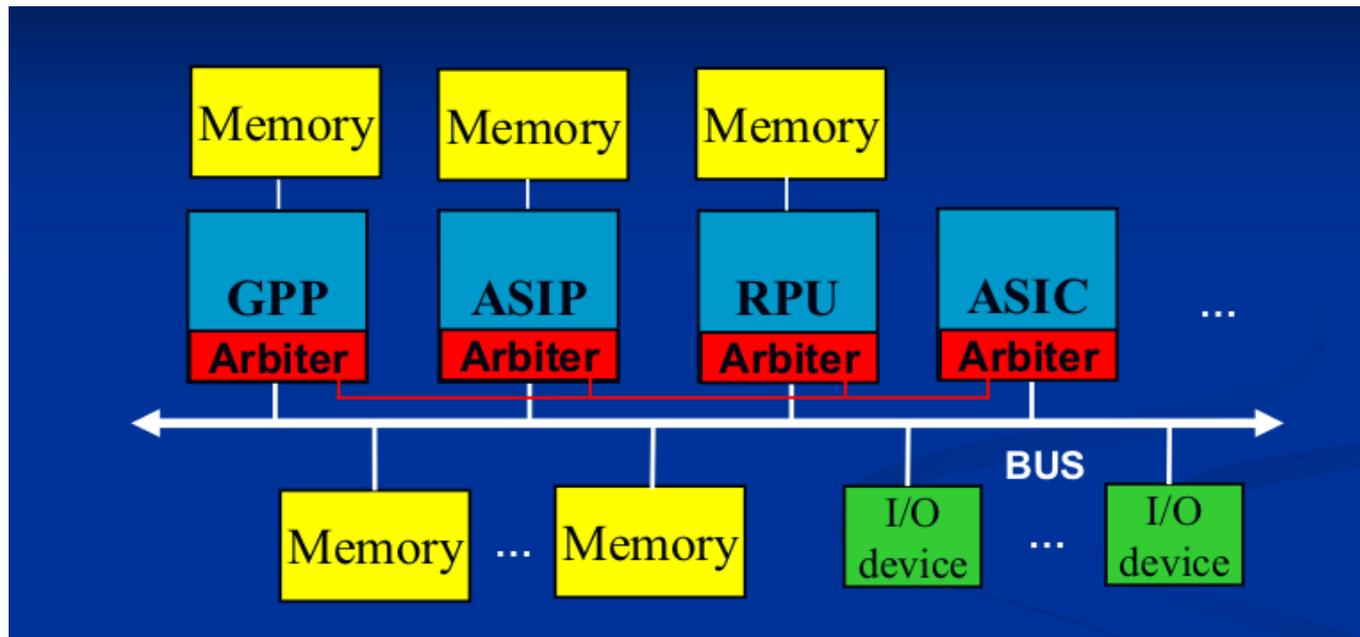
The Daisy Chain Bus Arbitrations Scheme



- Priority based
- Pros: simple
- Cons:
 - Cannot assure fairness:
 - A low-priority device may be locked out indefinitely
 - The use of the daisy chain grant signal also limits the bus speed

Distributed Arbitration

- Requires **fewer signals** compared to the centralized approach
- More hardware duplication, more logic/area, less scalable



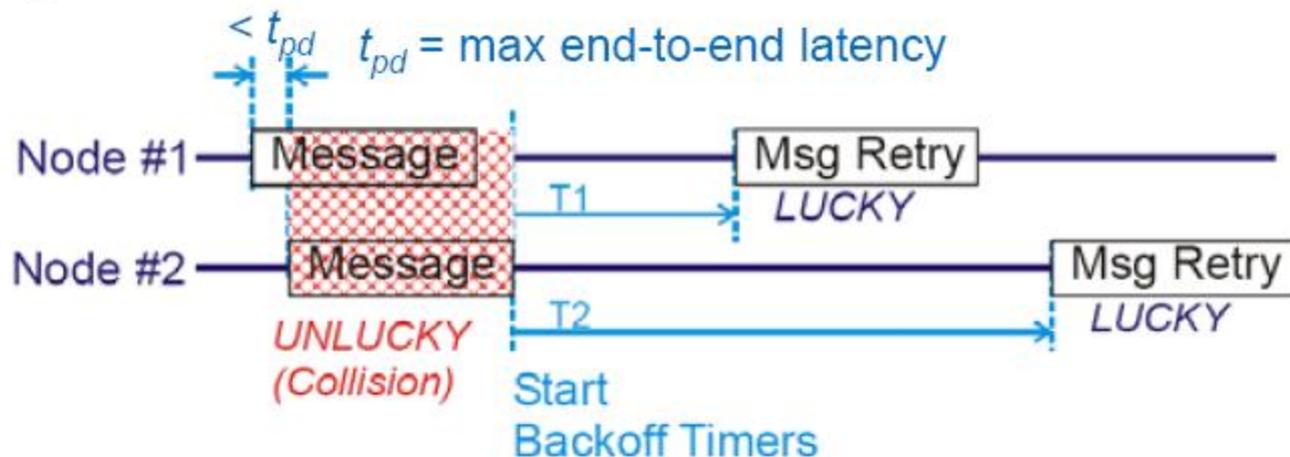
Carrier Sense Multiple Access (CSMA)

- A probabilistic media access control (MAC) protocol
- **Carrier sense** means that a transmitter uses feedback from a receiver to determine whether another transmission is in progress before initiating a transmission
- Sense before transmit / listen before talk
 - Transmitter tries to detect the presence of a **carrier wave** from another station before attempting to transmit



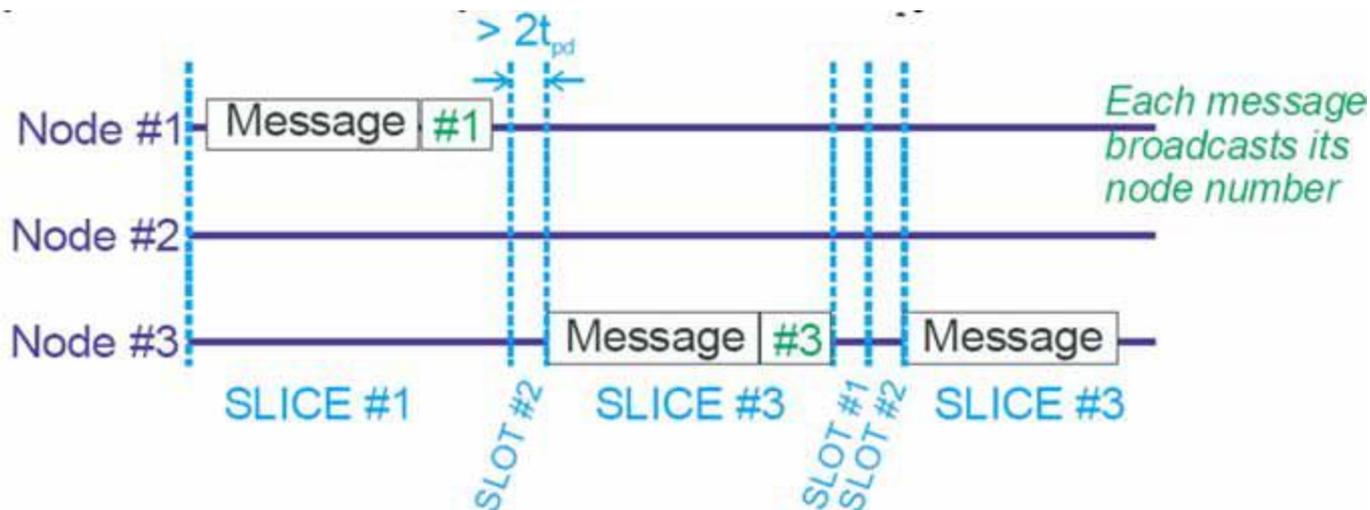
CSMA/CD

- Carrier Sense Multiple Access / Collision Detection
 - Try to avoid and detect collisions:
 - before starting to transmit, MASTER checks whether the bus is idle
 - if a collision is detected (several MASTERS started almost simultaneously), wait for some time (back-off timer)
 - Repeated collisions result in increasing back-off times
 - Examples: Ethernet, IEEE 802.3



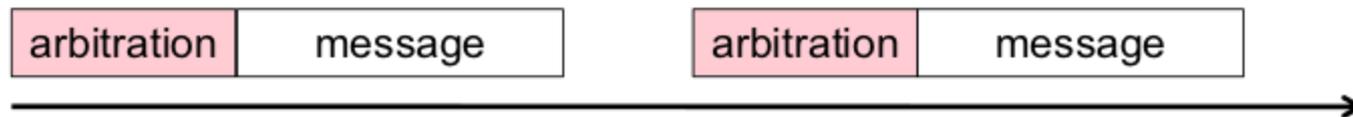
CSMA/CA – Flexible TDMA (FTDMA)

- Carrier Sense Multiple Access / Collision Avoidance
- Operation:
 - Reserve s slots for n nodes; note: slots are normally idle – they are (short) time intervals, not signals; if slot is used it becomes a slice.
 - nodes keep track of global communication state by sensing
 - nodes start transmitting a message only during the assigned slot
 - If $s=n$, no collisions; if $s < n$, statistical collision avoidance
- Examples: 802.11, part of FlexRay

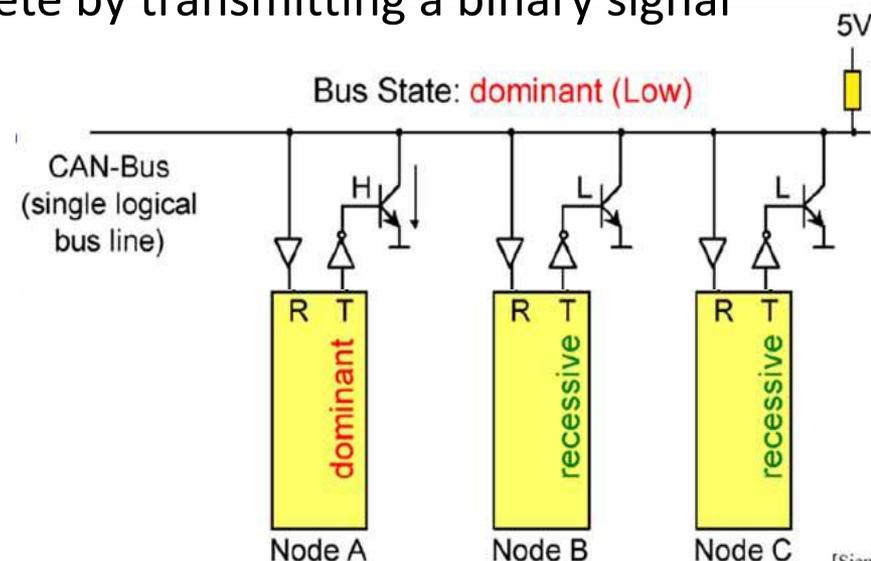


CSMA/CR

- Carrier Sense Multiple Access / Collision Resolution
- Operation:
 - Before any message transmission, there is a global arbitration



- Each node is assigned a unique identification number
- All nodes wishing to transmit compete by transmitting a binary signal based on their identification value
- A node drops out the competition if it detects a dominant state while transmitting a passive state
- Thus, the node with the lowest identification value wins
- Example: CAN Bus

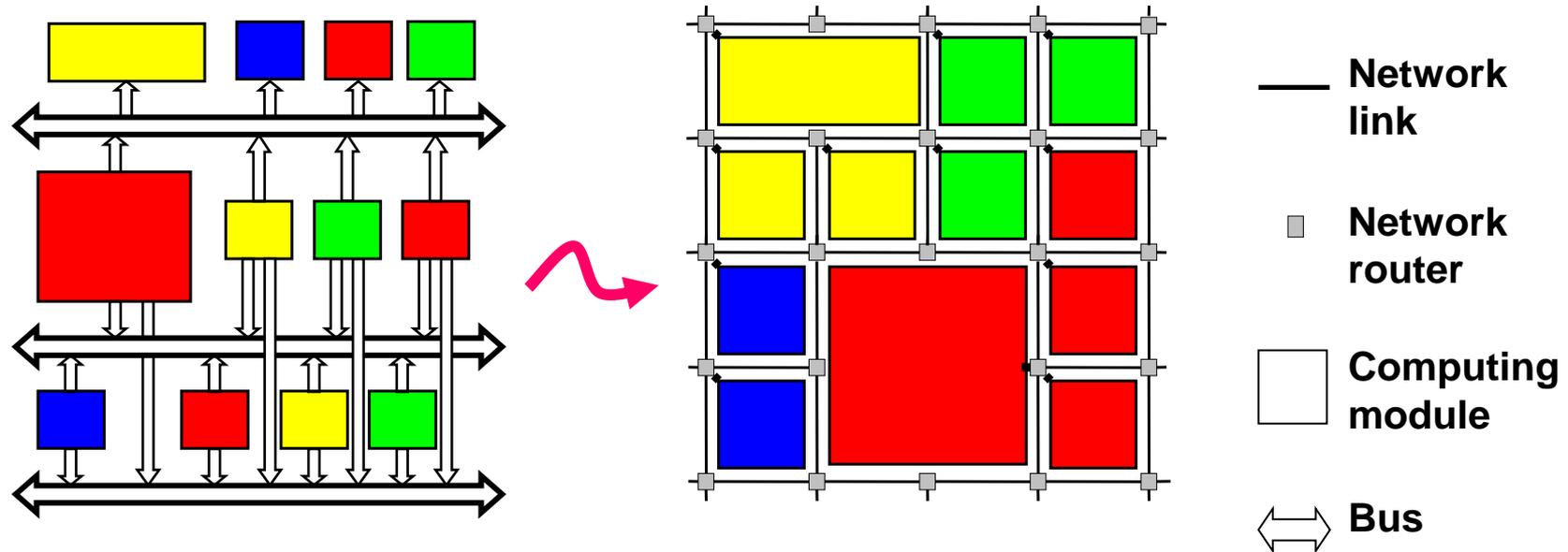


Outline

- Bus basics
- Multiple Master Bus
 - Centralized arbitration
 - Distributed arbitration
- **Network-on-Chip**
- Examples



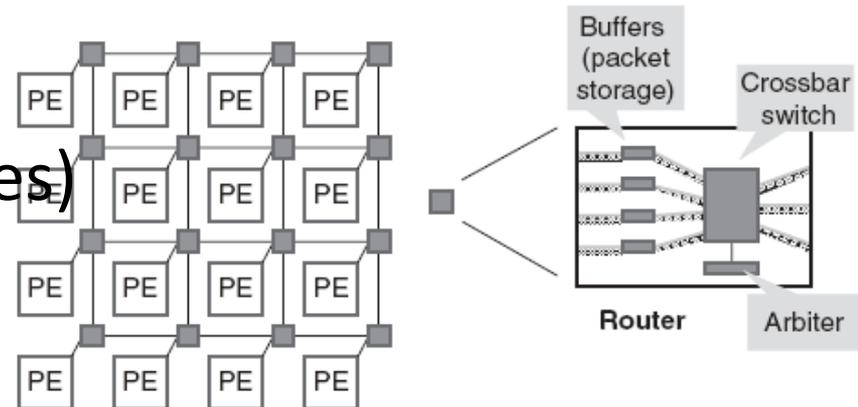
Evolution or Paradigm Shift?



- Architectural paradigm shift
 - Replace wire spaghetti by an intelligent network infrastructure
- Design paradigm shift
 - Busses and signals replaced by packets
- Organizational paradigm shift
 - Create a new discipline, a new infrastructure responsibility

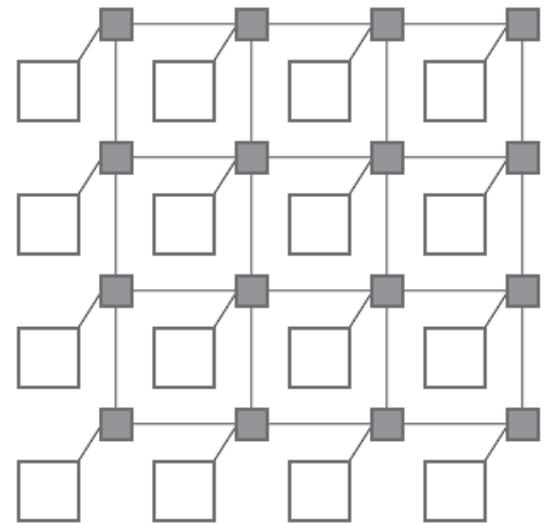
Network-On-Chip: Introduction

- Network-on-chip (NoC) is a packet switched onchip communication network
- NoCs use packets to route data from the source to the destination PE via a network fabric that consists of
 - network interfaces (NI)
 - switches (routers)
 - interconnection links (wires)

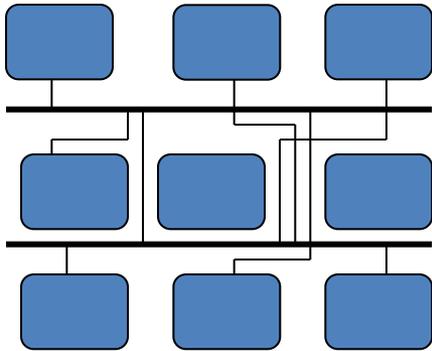


NoC Properties

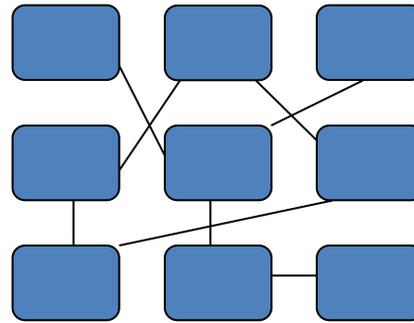
- NoCs are an attempt to scale down the concepts of large-scale networks, and apply them to the Embedded System-on-chip (SoC) domain
- NoC Properties
 - Reliable and predictable electrical and physical properties
 - Regular geometry that is scalable
 - Flexible QoS guarantees
 - Higher bandwidth
 - Reusable components
 - Buffers, arbiters, routers, protocol stack



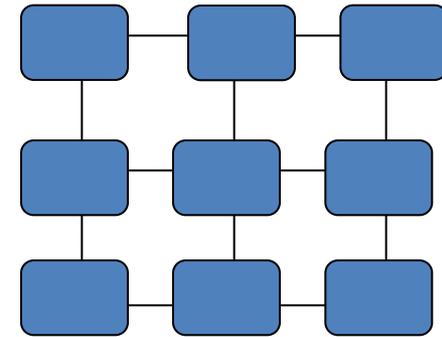
Bus vs NoCs



Bus-based architectures



Irregular architectures



Regular Architectures

■ Bus based interconnect

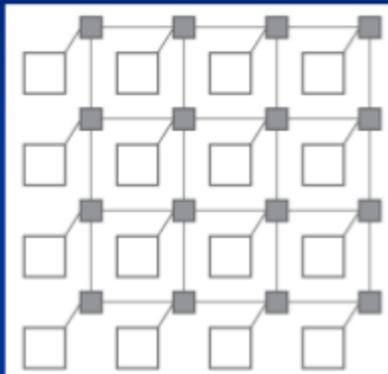
- Low cost
- Easier to Implement
- Flexible

■ Networks on Chip

- Layered Approach
- Buses replaced with Networked architectures
- Better electrical properties
- Higher bandwidth
- Energy efficiency
- Scalable

NoC Topology

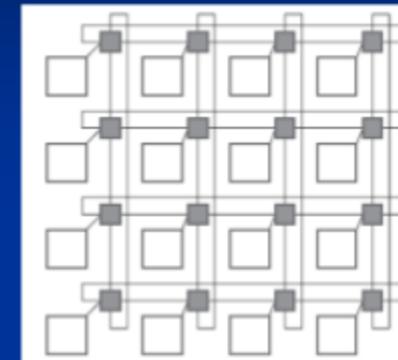
2-D Mesh



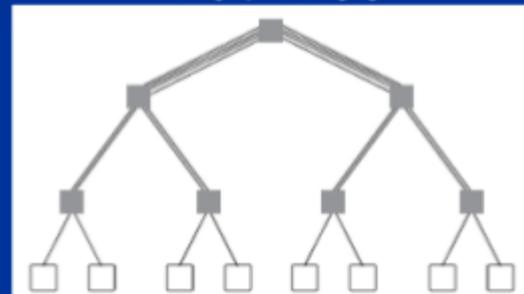
1-D Torus



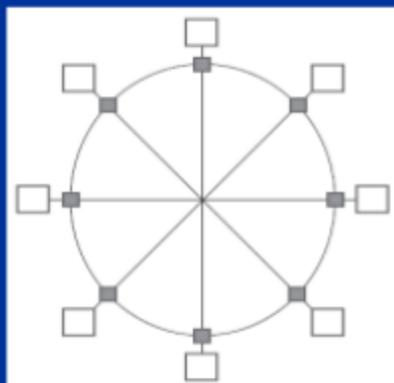
2-D Torus



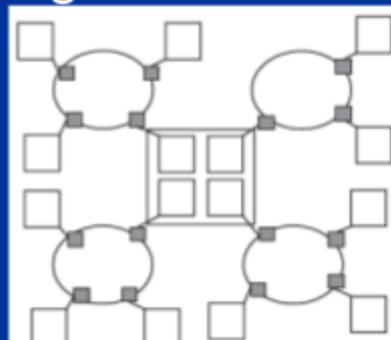
Fat Tree



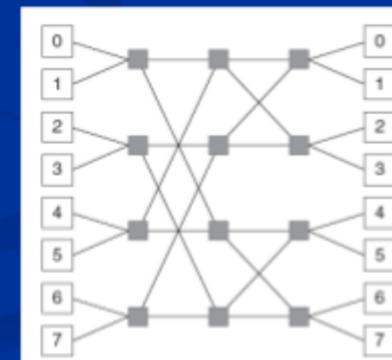
Octagon



Irregular or Ad-hoc



Butterfly



Scalability – Area and Power in NoCs

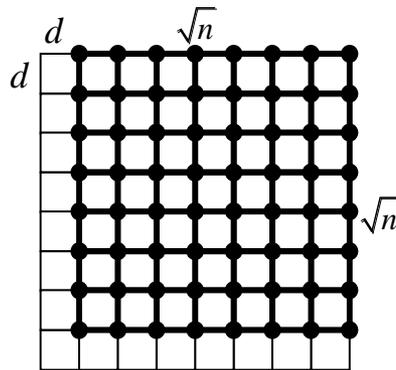
For Same Performance, compare the:

Wire-area and power:

NoC:

$$O(n)$$

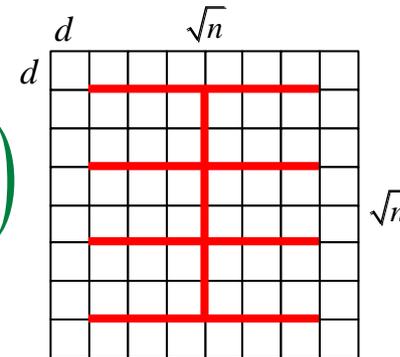
$$O(n)$$



Simple Bus:

$$O(n^3 \sqrt{n})$$

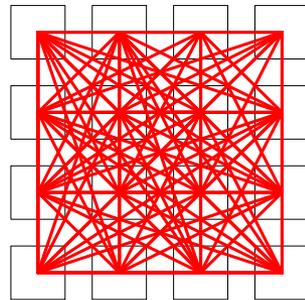
$$O(n\sqrt{n})$$



Point-to-Point:

$$O(n^2 \sqrt{n})$$

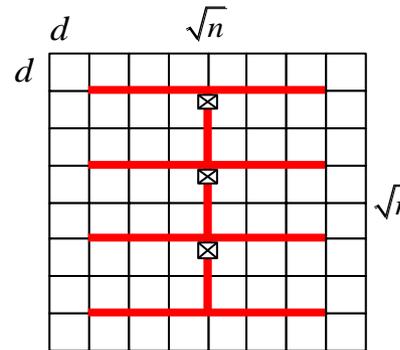
$$O(n\sqrt{n})$$



Segmented Bus:

$$O(n^2 \sqrt{n})$$

$$O(n\sqrt{n})$$



E. Bolotin et al., "Cost Considerations in Network on Chip", *Integration*, special issue on Network on Chip, October 2004



Network-On-Chip: Issues

- Very hot research topic at the moment
- NO wide adoption in Industry because
 - Power
 - complex NI and switching/routing logic blocks are power hungry
 - several times greater than for current bus-based approaches
 - Latency
 - additional delay to packetize/de-packetize data at NIs
 - flow/congestion control and fault tolerance protocol overheads
 - delays at the numerous switching stages encountered by packets
 - even circuit switching at routers has overhead
 - lags behind what can be achieved with bus-based/dedicated wiring
 - Lack of tools and benchmarks

