CSC 252: Computer Organization Spring 2022: Lecture 11

Instructor: Yuhao Zhu

Department of Computer Science University of Rochester

Announcements

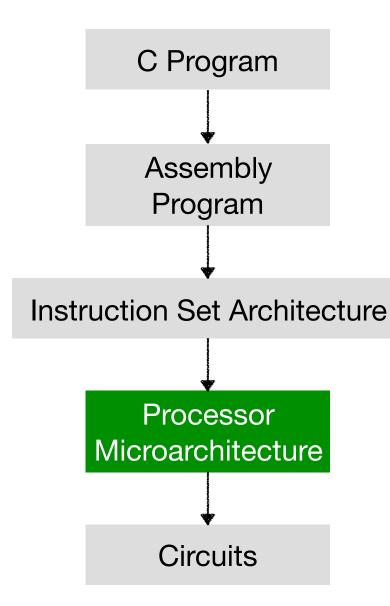
- Programming assignment 3 is out
 - Details: <u>https://www.cs.rochester.edu/courses/252/spring2022/labs/</u> assignment3.html
 - Due on March 3, 11:59 PM
 - You (may still) have 3 slip days

13	14	15	16	17	18	19
				Today		
20	21	22	23	24	25	26
27	28	Mar 1	2	³ Due	4	5
				Mid-term		
						2

Announcements

- Grades for Lab 1 are posted.
- We are processing regrading requests
- Will grade Lab 2 soon.
- Programming assignment 3 is in x86 assembly language. Seek help from TAs.
- TAs are best positioned to answer your questions about programming assignments!!!
- Programming assignments do NOT repeat the lecture materials. They ask you to synthesize what you have learned from the lectures and work out something new.

So far in 252...

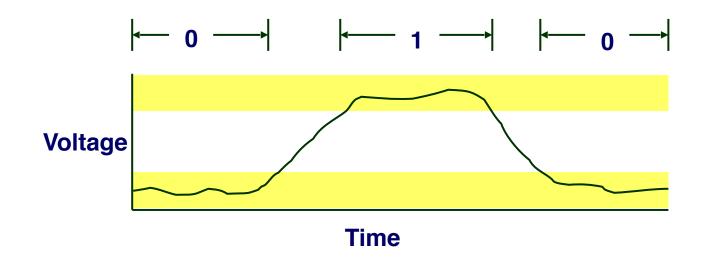


Today: Circuits Basics

- Basics
- Circuits for computations
- Circuits for storing data

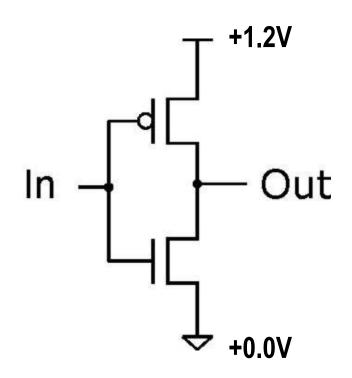
Overview of Circuit-Level Design

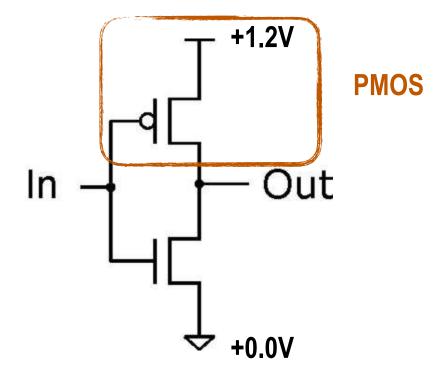
- Fundamental Hardware Requirements
 - Communication: How to get values from one place to another. Mainly three electrical **wires**.
 - Computation: transistors. Combinational logic.
 - Storage: transistors. Sequential logic.
- Circuit design is often abstracted as logic design

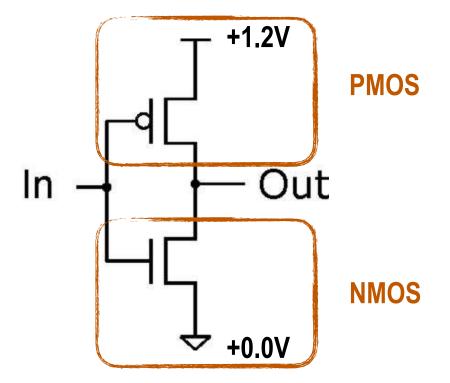


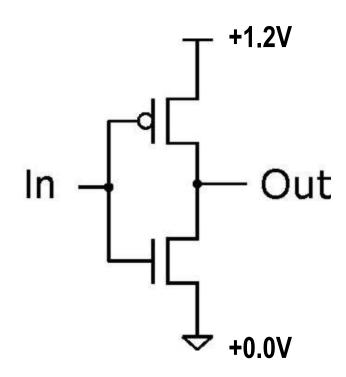
Today: Circuits Basics

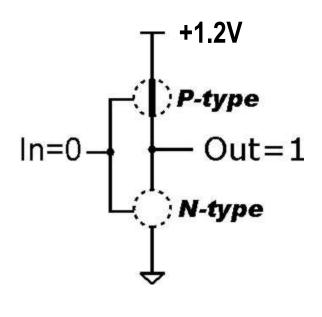
- Transistors
- Circuits for computations
- Circuits for storing data

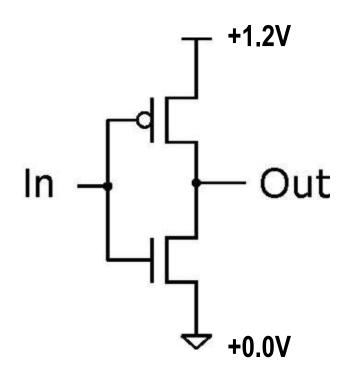


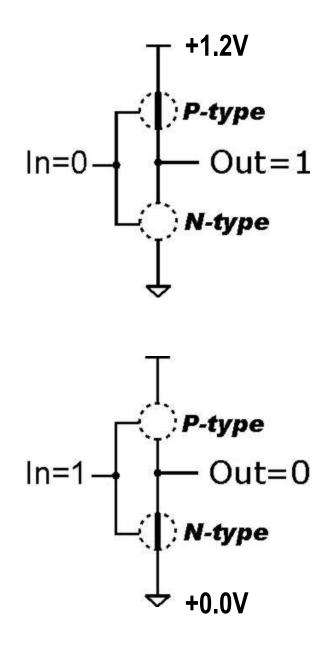


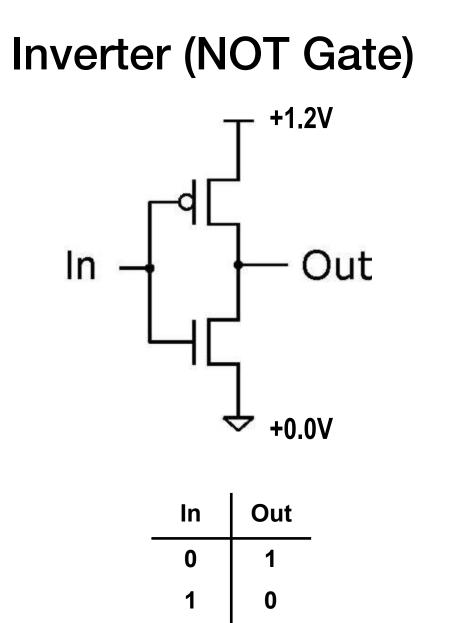


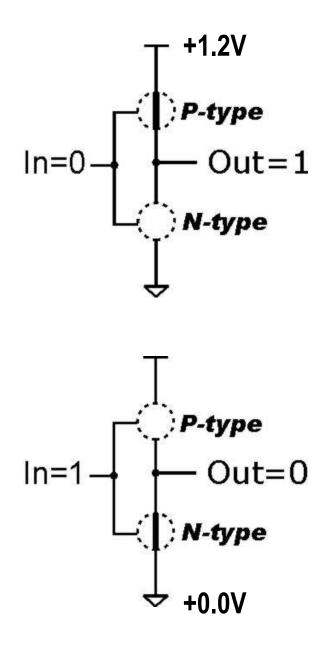




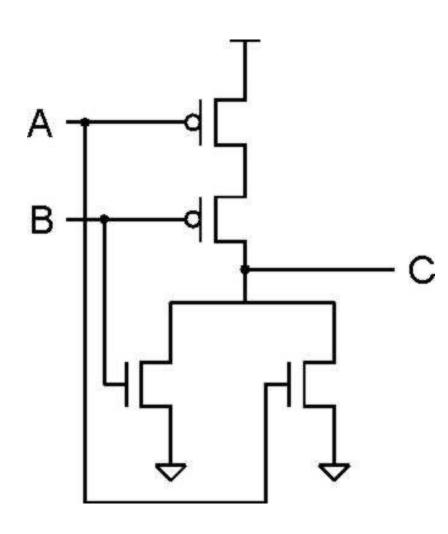


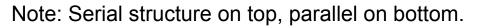


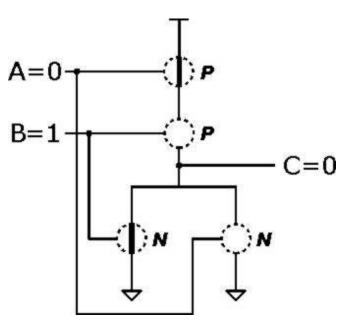




NOR Gate (NOT + OR)

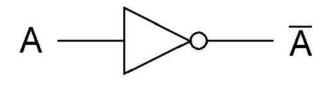




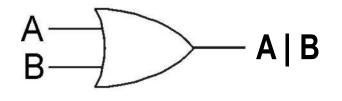


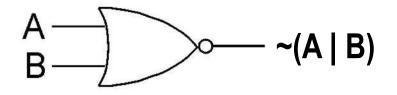
Α	В	С
0	0	1
0	1	0
1	0	0
1	1	0

Basic Logic Gates



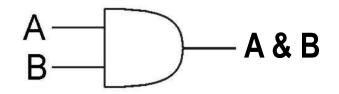


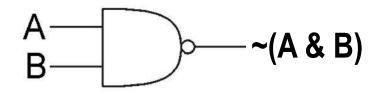






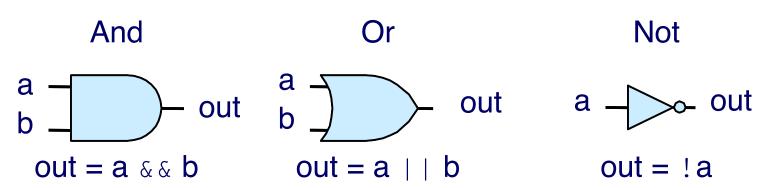
NOR



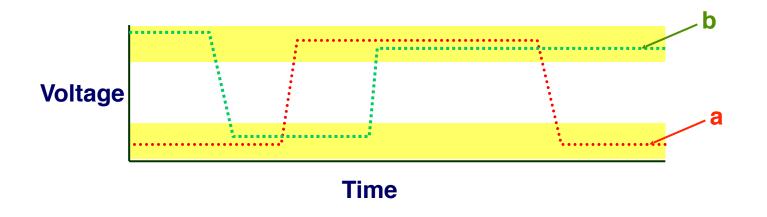


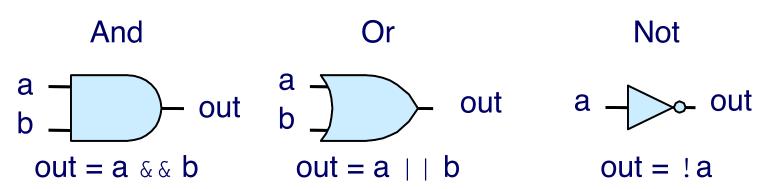
AND

NAND

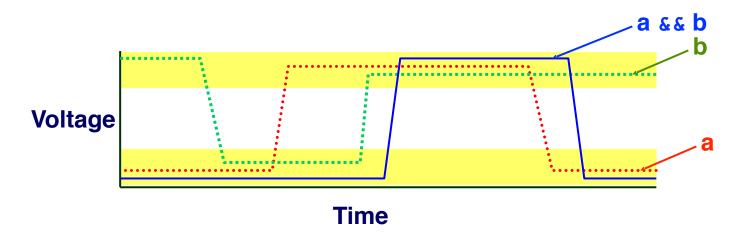


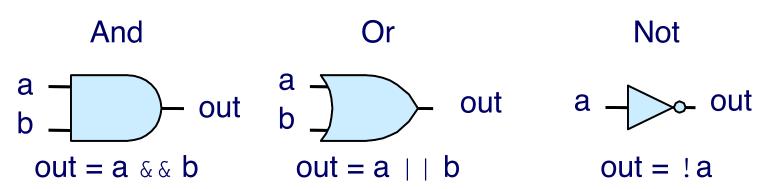
- Outputs are Boolean functions of inputs
- Respond continuously to changes in inputs with some small delay
- Different gates have different delays (b/c different transistor combinations)



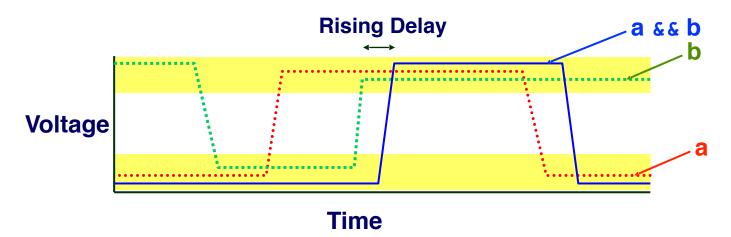


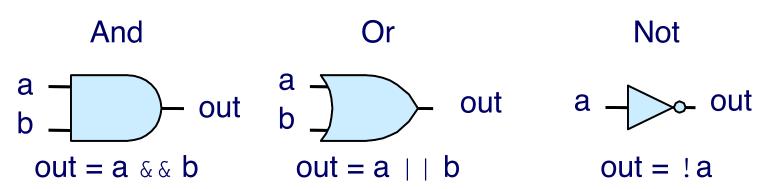
- Outputs are Boolean functions of inputs
- Respond continuously to changes in inputs with some small delay
- Different gates have different delays (b/c different transistor combinations)



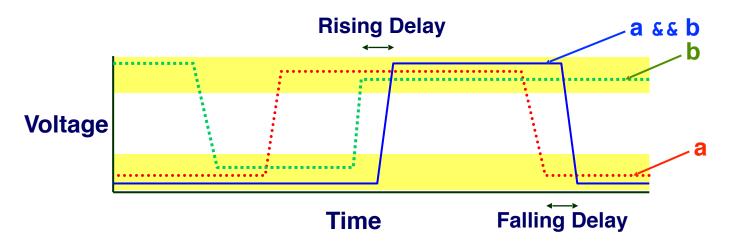


- Outputs are Boolean functions of inputs
- Respond continuously to changes in inputs with some small delay
- Different gates have different delays (b/c different transistor combinations)

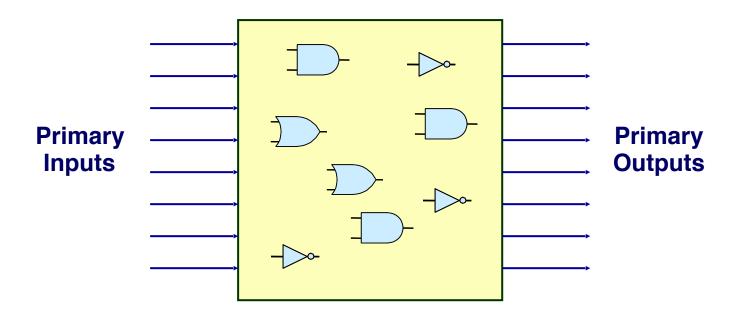




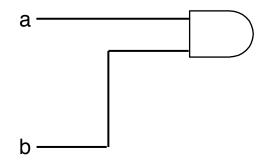
- Outputs are Boolean functions of inputs
- Respond continuously to changes in inputs with some small delay
- Different gates have different delays (b/c different transistor combinations)

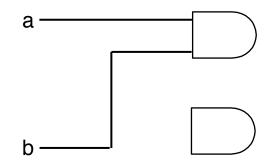


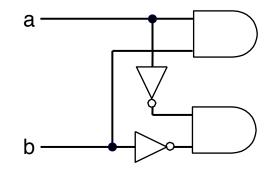
Combinational Circuits

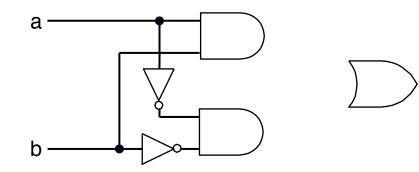


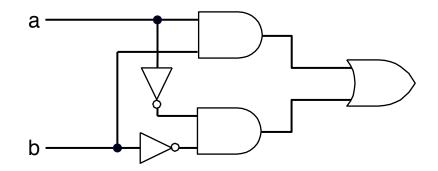
- A Network of Logic Gates
 - Continuously responds to changes on primary inputs
 - Primary outputs become (after some delay) Boolean functions of primary inputs

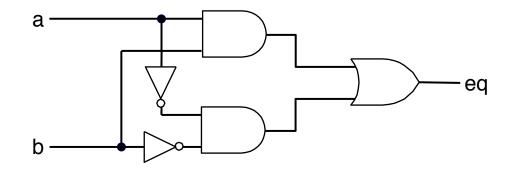


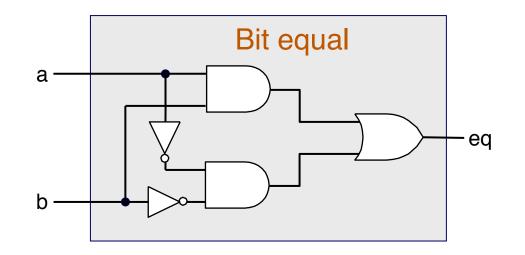


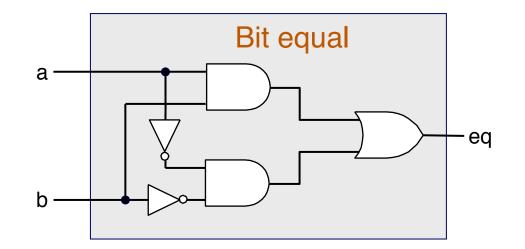




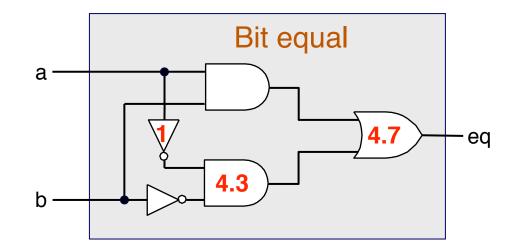




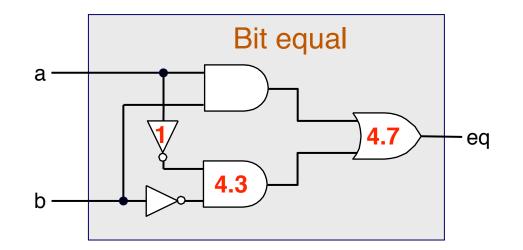




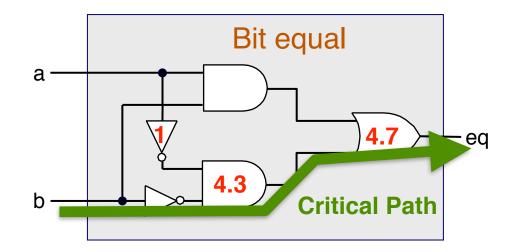
- What's the delay of this bit equal circuit?
 - Assuming 1-input NOT takes 1 unit of time, 2-input AND takes 4.3, and 2-input OR takes 4.7



- What's the delay of this bit equal circuit?
 - Assuming 1-input NOT takes 1 unit of time, 2-input AND takes 4.3, and 2-input OR takes 4.7

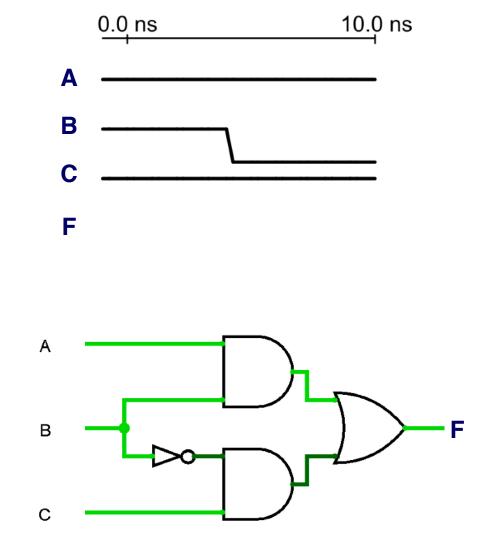


- What's the delay of this bit equal circuit?
 - Assuming 1-input NOT takes 1 unit of time, 2-input AND takes 4.3, and 2-input OR takes 4.7
- The delay of a circuit is determined by its "critical path"
 - The path between an input and the output that the maximum delay
 - Estimating the critical path delay is called static timing analysis

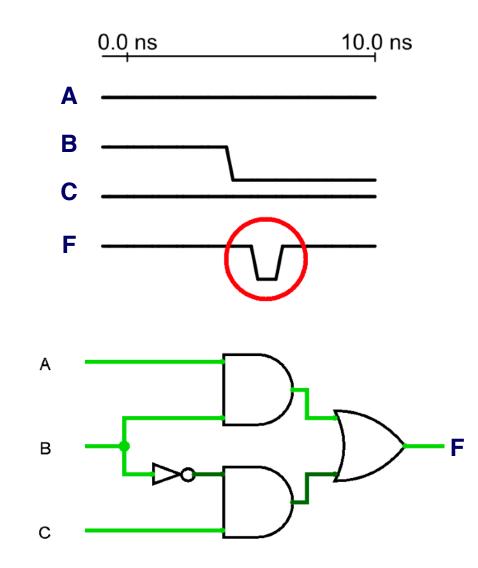


- What's the delay of this bit equal circuit?
 - Assuming 1-input NOT takes 1 unit of time, 2-input AND takes 4.3, and 2-input OR takes 4.7
- The delay of a circuit is determined by its "critical path"
 - The path between an input and the output that the maximum delay
 - Estimating the critical path delay is called static timing analysis

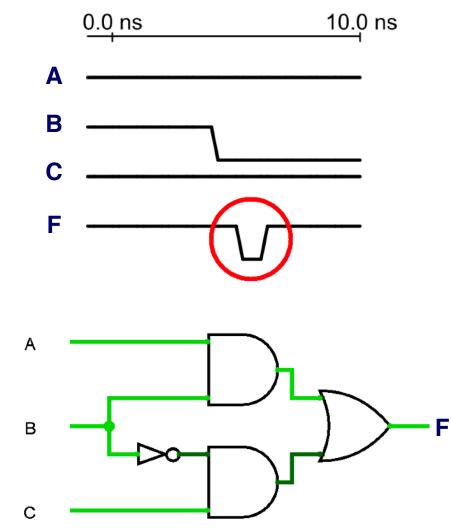
Glitch/Hazard



Glitch/Hazard

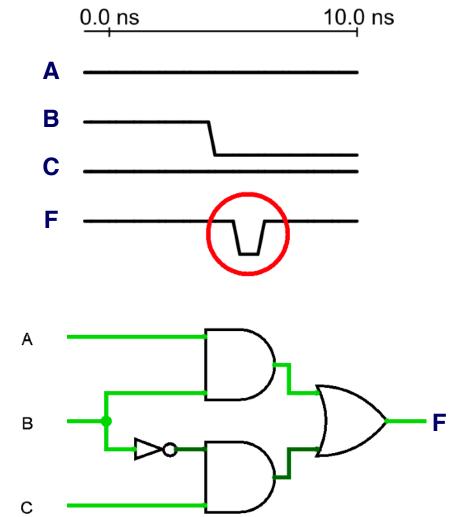


Glitch/Hazard



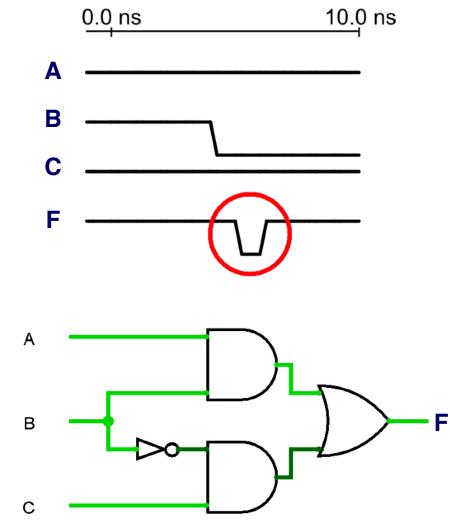
• A glitch is an unnecessary signal transition without functionality.

Glitch/Hazard



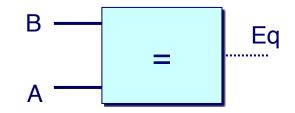
- A glitch is an unnecessary signal transition without functionality.
- Why is it bad? When transistors switch they consume power, but the power consumed during a glitch is a waste.

Glitch/Hazard

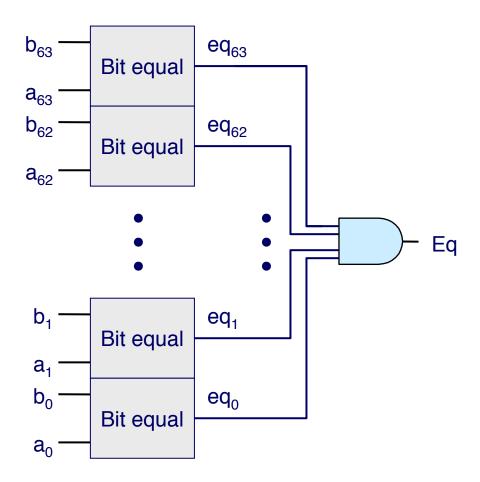


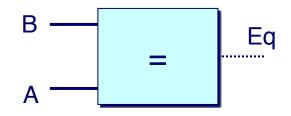
- A glitch is an unnecessary signal transition without functionality.
- Why is it bad? When transistors switch they consume power, but the power consumed during a glitch is a waste.
- Without care, glitch power dissipation is 20%-70% of total power dissipation.

64-bit Equality



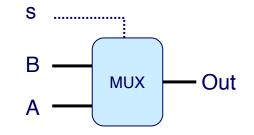
64-bit Equality





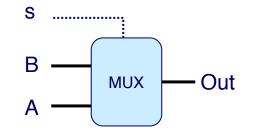
Bit-Level Multiplexor (MUX)

- Control signal s
- Data signals A and B
- Output A when s=1, B when s=0



Bit-Level Multiplexor (MUX)

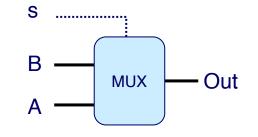
- Control signal s
- Data signals A and B
- Output A when s=1, B when s=0

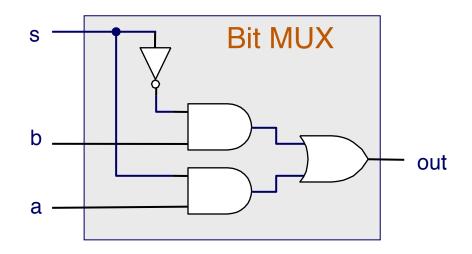


bool out = (s&&a) || (!s&&b)

Bit-Level Multiplexor (MUX)

- Control signal s
- Data signals A and B
- Output A when s=1, B when s=0

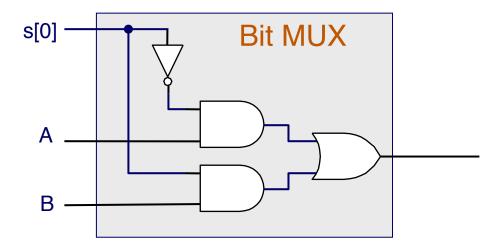




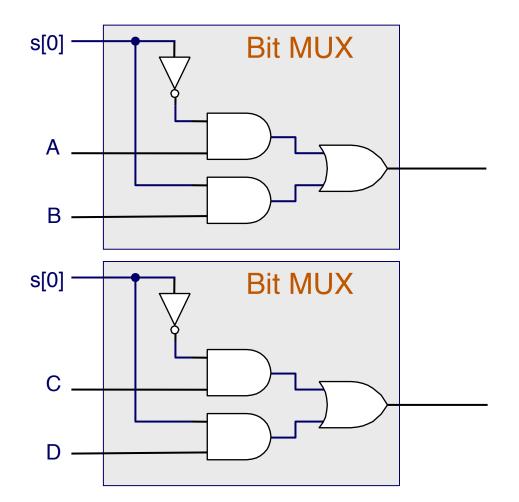
bool out = (s&&a) || (!s&&b)

- Control signal s; Data signals A, B, C, and D
- Output: A when s = 00, B when s = 01, C when s = 10, D when s = 11

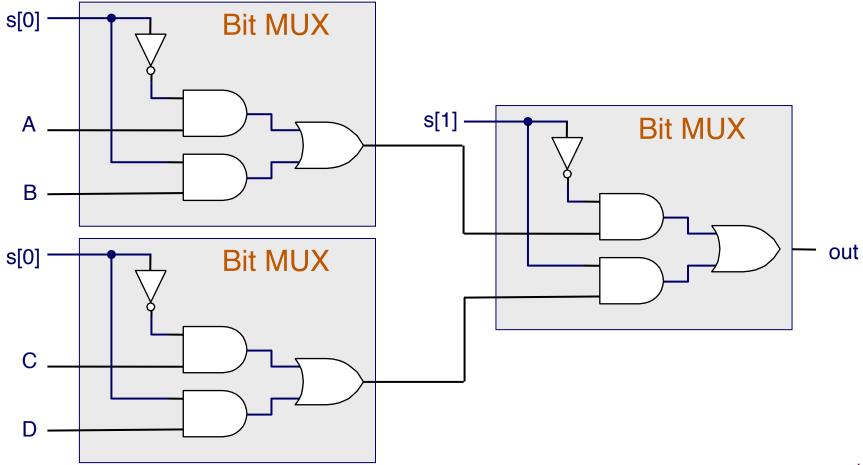
- Control signal s; Data signals A, B, C, and D
- Output: A when s = 00, B when s = 01, C when s = 10, D when s = 11



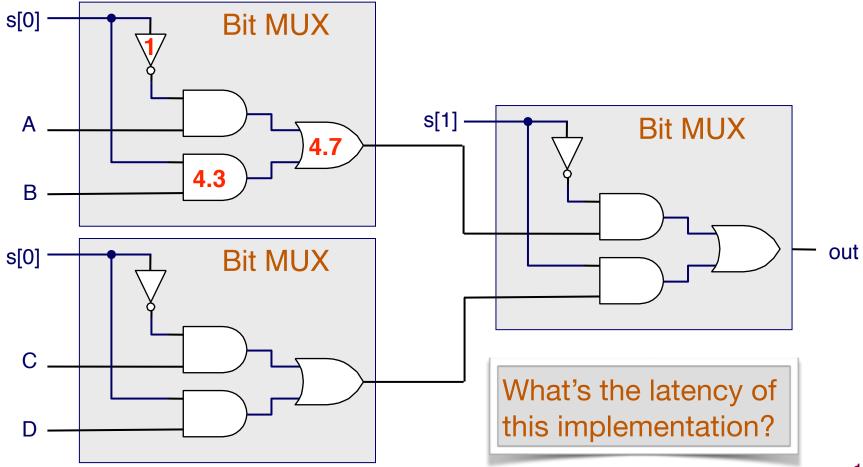
- Control signal s; Data signals A, B, C, and D
- Output: A when s = 00, B when s = 01, C when s = 10, D when s = 11



- Control signal s; Data signals A, B, C, and D
- Output: A when s = 00, B when s = 01, C when s = 10, D when s = 11



- Control signal s; Data signals A, B, C, and D
- Output: A when s = 00, B when s = 01, C when s = 10, D when s = 11



• The number of inputs of a gate (fan-in) and the number of outputs of a gate (fan-out) will affect the gate delay.

- The number of inputs of a gate (fan-in) and the number of outputs of a gate (fan-out) will affect the gate delay.
- Think of logic gates as LEGOs, using which you generate the gate level circuit design for complex functionalities.

- The number of inputs of a gate (fan-in) and the number of outputs of a gate (fan-out) will affect the gate delay.
- Think of logic gates as LEGOs, using which you generate the gate level circuit design for complex functionalities.
- A standard cell library is a collection of well defined and appropriately characterized logic gates (delay, operating voltage, etc.) that can be used to implement a digital design.

- The number of inputs of a gate (fan-in) and the number of outputs of a gate (fan-out) will affect the gate delay.
- Think of logic gates as LEGOs, using which you generate the gate level circuit design for complex functionalities.
- A standard cell library is a collection of well defined and appropriately characterized logic gates (delay, operating voltage, etc.) that can be used to implement a digital design.
- The *logic synthesis tool* will automatically generate the "best" gate-level implementation of a piece of logic.

- The number of inputs of a gate (fan-in) and the number of outputs of a gate (fan-out) will affect the gate delay.
- Think of logic gates as LEGOs, using which you generate the gate level circuit design for complex functionalities.
- A standard cell library is a collection of well defined and appropriately characterized logic gates (delay, operating voltage, etc.) that can be used to implement a digital design.
- The *logic synthesis tool* will automatically generate the "best" gate-level implementation of a piece of logic.
- Take a Logic Design or Very Large Scale Integrated-Circuit (VLSI) course if you want to know more about circuit design.
 - Logic design uses the gate-level abstractions
 - VLSI tells you how the gates are implemented at transistor-level

Recall: Full (1-bit) Adder

Add two bits and carry-in, produce one-bit sum and carry-out.

Α	В	C _{in}	S	\mathbf{C}_{ou}
				t
0	0	0	0	t 0
0	0	1	1	0
0	1	0	1	0
0	1	1	0	1
1	0	0	1	0
1	0	1	0	1
1	1	0	0	1
1	1	1	1	1
			I	

Recall: Full (1-bit) Adder

Add two bits and carry-in, produce one-bit sum and carry-out.

 $S = (~A \& ~B \& C_{in})$ (~A & B & ~C_{in}) (A & ~B & ~C_{in}) (A & B & C_{in}) $C_{ou} = (~A \& B \& C_{in})$ (A & ~B & C_{in}) (A & B & ~C_{in}) (A & B & C_{in})

	A	В	C _{in}	S	\mathbf{C}_{ou}
					t
-	0	0	0	0	0
	0	0	1	1	0
	0	1	0	1	0
	0	1	1	0	1
	1	0	0	1	0
	1	0	1	0	1
	1	1	0	0	1
	1	1	1	1	1
			I		

Recall: 1-bit Full Adder

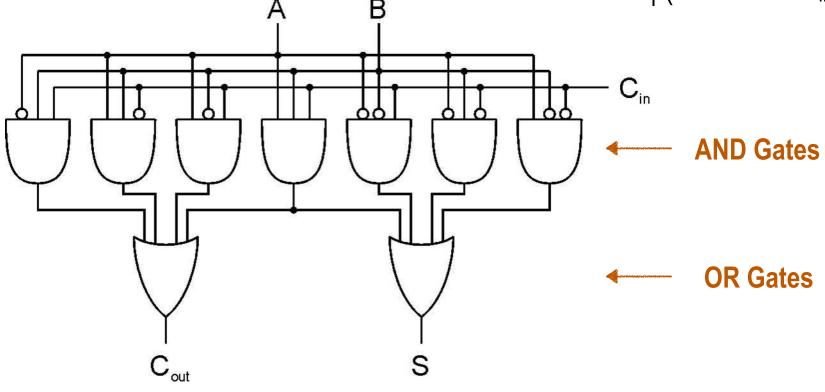
Add two bits and carry-in, produce one-bit sum and carry-out.

C_{ou} = (~A & B & C_{in}) | (A & ~B & C_{in}) | (A & B & ~C_{in}) | (A & B & C_{in})

Recall: 1-bit Full Adder

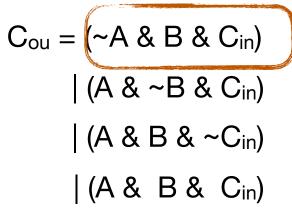
Add two bits and carry-in, produce one-bit sum and carry-out.

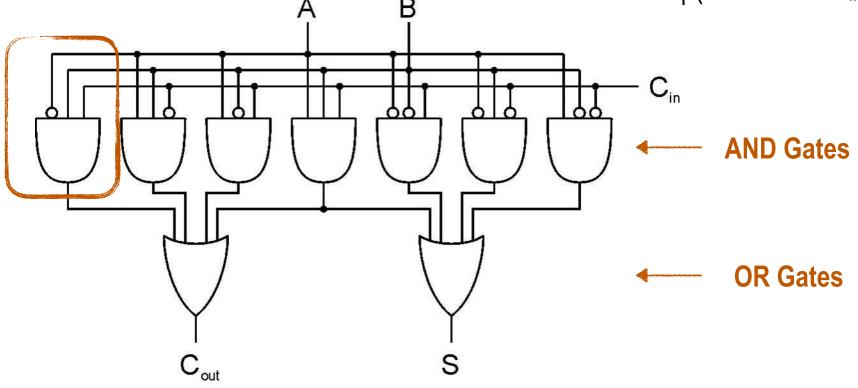
C_{ou} = (~A & B & C_{in}) | (A & ~B & C_{in}) | (A & B & ~C_{in}) | (A & B & C_{in})



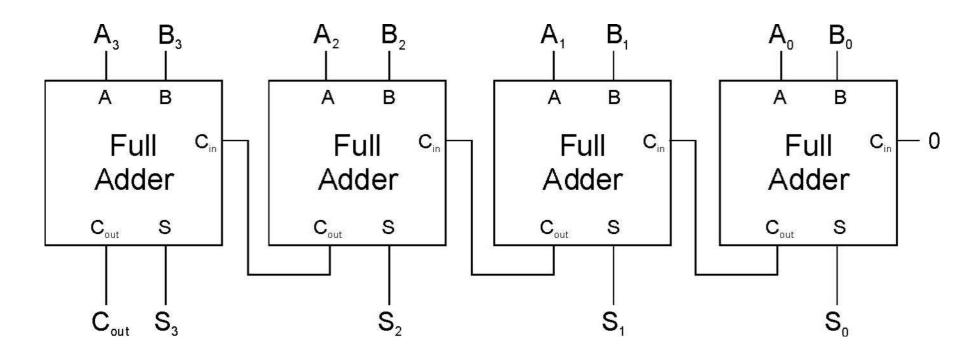
Recall: 1-bit Full Adder

Add two bits and carry-in, produce one-bit sum and carry-out.



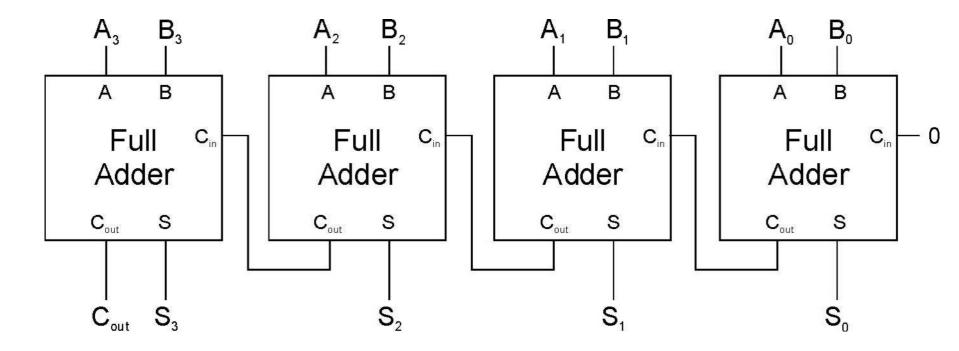


Recall: Four-bit Adder



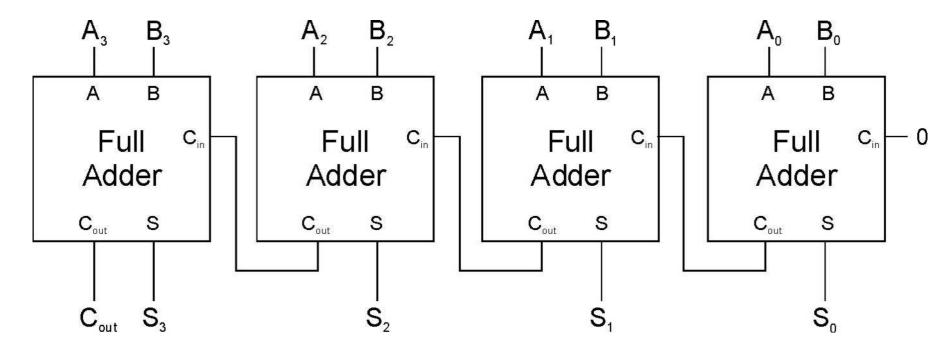
Recall: Four-bit Adder

- Ripple-carry Adder
 - Simple, but performance linear to bit width

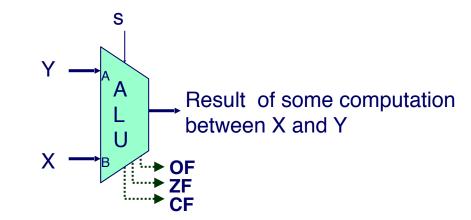


Recall: Four-bit Adder

- Ripple-carry Adder
 - Simple, but performance linear to bit width
- Carry look-ahead adder (CLA)
 - Generate all carriers simultaneously



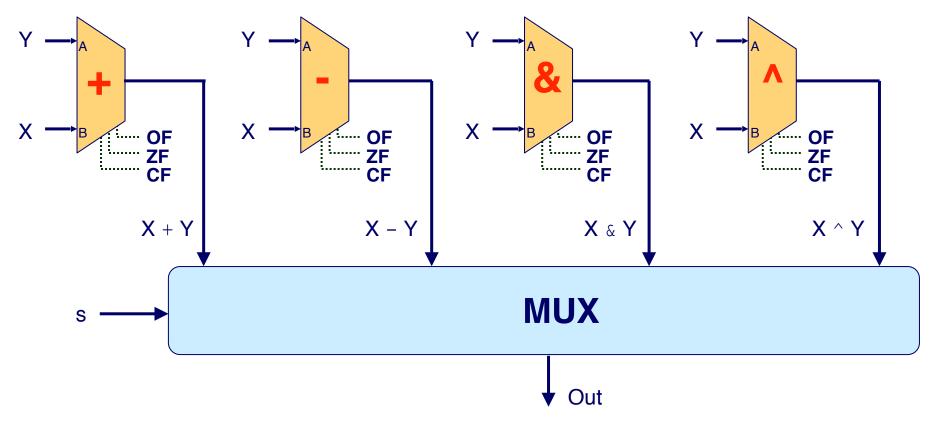
Arithmetic Logic Unit



- An ALU performs multiple kinds of computations.
- The actual computation depends on the selection signal s.
- Also sets the condition codes (status flags)
- For instance:
 - X + Y when s == 00
 - X Y when s == 01
 - X & Y when s == 10
 - X ^ Y when s == 11
- How can this ALU be implemented?

Arithmetic Logic Unit

- Implement 4 different circuits, one for each operation.
- Then use a MUX to select the results



Today: Circuits Basics

- Transistors
- Circuits for computations
- Circuits for storing data

- Assembly programs set architecture (processor) states.
 - Register File
 - Status Flags
 - Memory
 - Program Counter

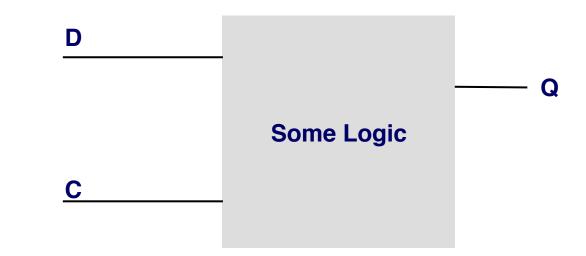
- Assembly programs set architecture (processor) states.
 - Register File
 - Status Flags
 - Memory
 - Program Counter
- Every state is essentially some bits that are stored/loaded.

- Assembly programs set architecture (processor) states.
 - Register File
 - Status Flags
 - Memory
 - Program Counter
- Every state is essentially some bits that are stored/loaded.
- Think of the program execution as an FSM.

- Assembly programs set architecture (processor) states.
 - Register File
 - Status Flags
 - Memory
 - Program Counter
- Every state is essentially some bits that are stored/loaded.
- Think of the program execution as an FSM.
- The hardware must provide mechanisms to load and store bits.

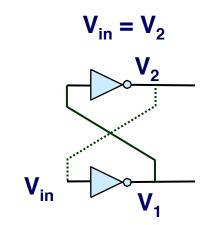
- Assembly programs set architecture (processor) states.
 - Register File
 - Status Flags
 - Memory
 - Program Counter
- Every state is essentially some bits that are stored/loaded.
- Think of the program execution as an FSM.
- The hardware must provide mechanisms to load and store bits.
- There are many different ways to store bits. They have trade-offs.

Build a 1-Bit Storage

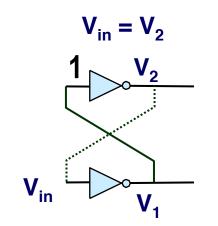


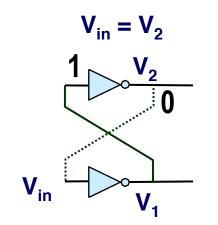
- What we would like:
 - D is the data we want to store (0 or 1)
 - C is the control signal
 - When C is 1, Q becomes D (i.e., storing the data)
 - When C is 0, Q doesn't change with D (data stored)

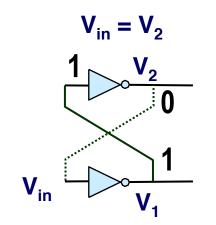
Bitstable Element

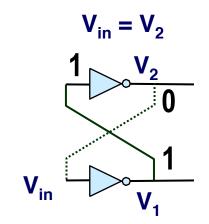


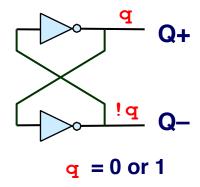
Bitstable Element

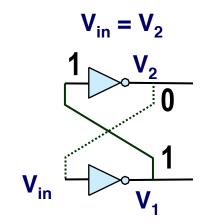




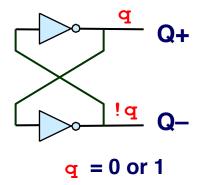




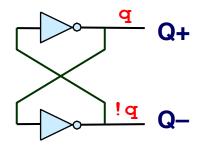




Bistable Element

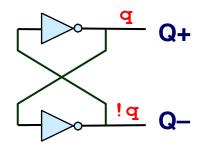


Q+ *continuously* outputs q.

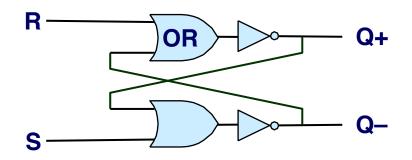


q = 0 or 1

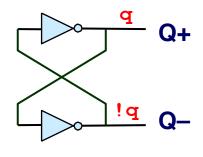
Bistable Element



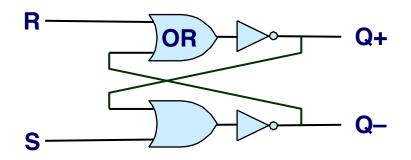
q = 0 or 1

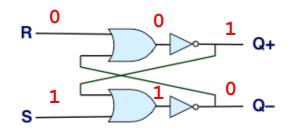


Bistable Element

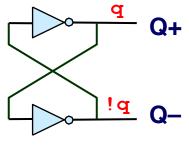


q = 0 or 1

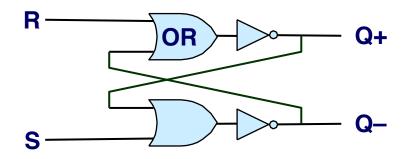




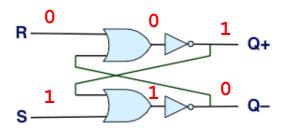
Bistable Element

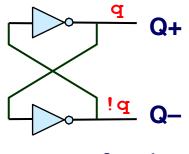


q = 0 or 1

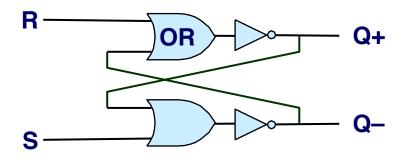


Setting Q+ to 1

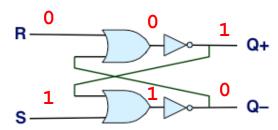


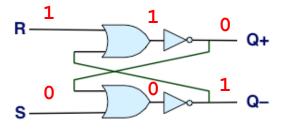


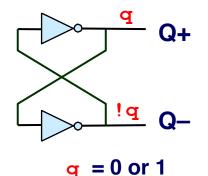


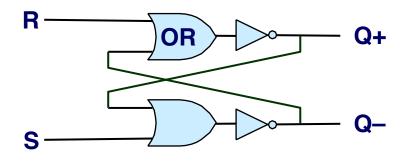


Setting Q+ to 1

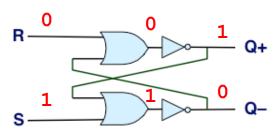




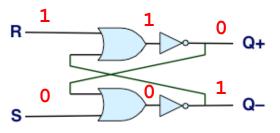


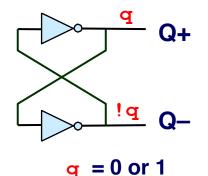


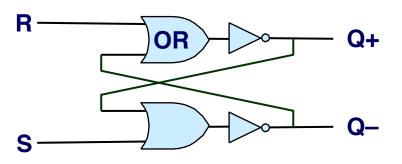
Setting Q+ to 1



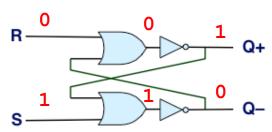
Setting Q+ to 0



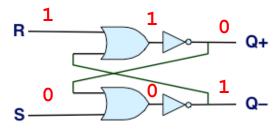


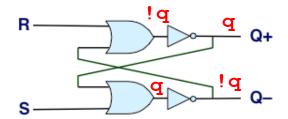


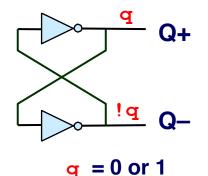
Setting Q+ to 1

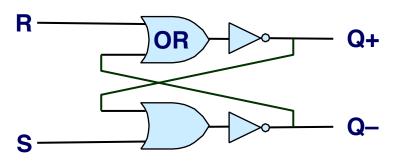


Setting Q+ to 0

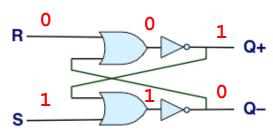




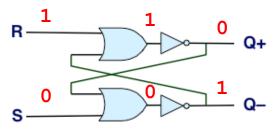


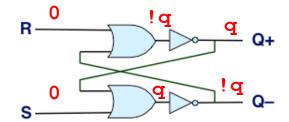


Setting Q+ to 1

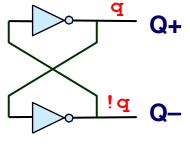


Setting Q+ to 0

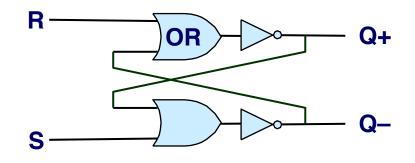




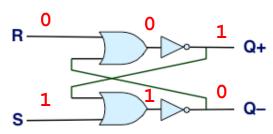
Bistable Element



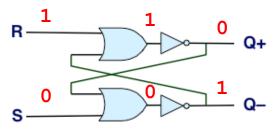
q = 0 or 1



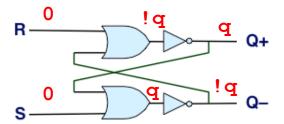
Setting Q+ to 1



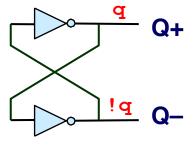
Setting Q+ to 0



Q+ value unchanged i.e., stored!

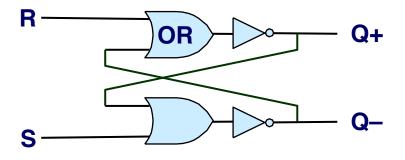


Bistable Element

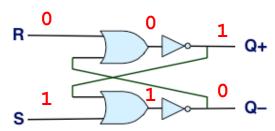


q = 0 or 1

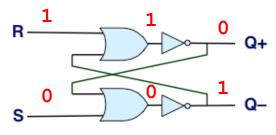




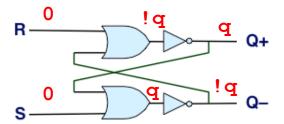
Setting Q+ to 1



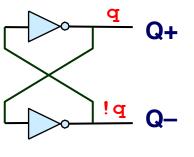
Setting Q+ to 0



Q+ value unchanged i.e., stored!

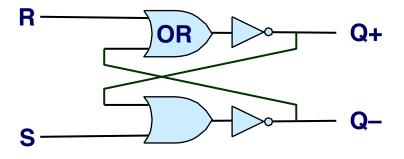


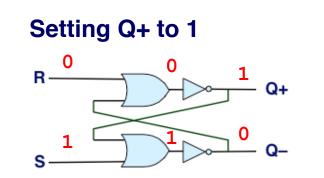
Bistable Element



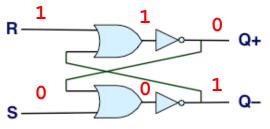
q = 0 or 1



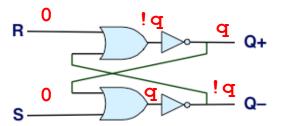




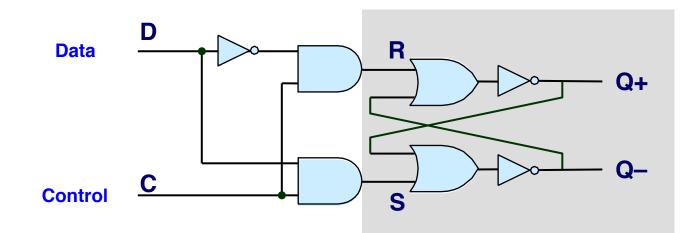




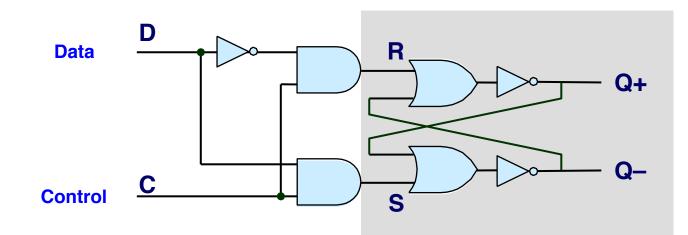




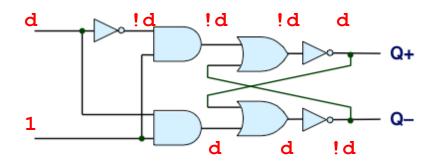
If R and S are different, Q+ is the same as S

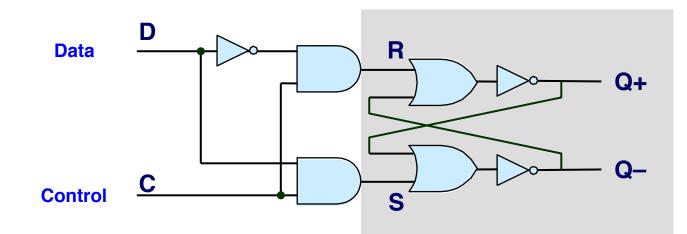


If R and S are different, Q+ is the same as S

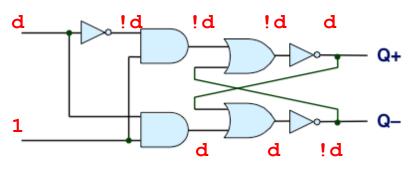


If R and S are different, Q+ is the same as S

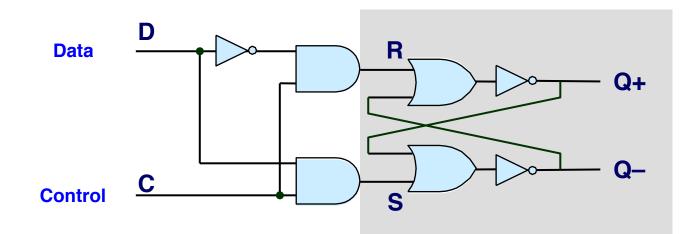




If R and S are different, Q+ is the same as S

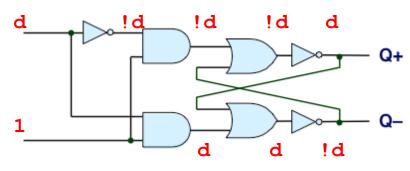


Q+ will continuously change as d changes

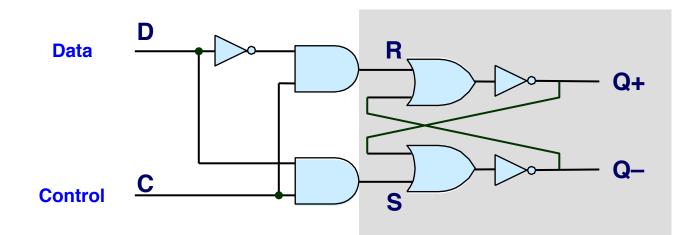


If R and S are different, Q+ is the same as S

Storing Data (Latching)

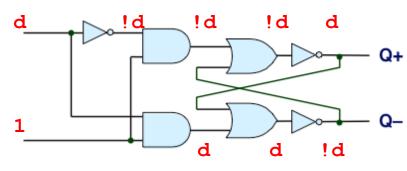


Q+ will continuously change as d changes

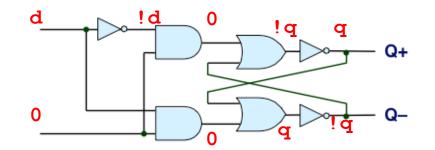


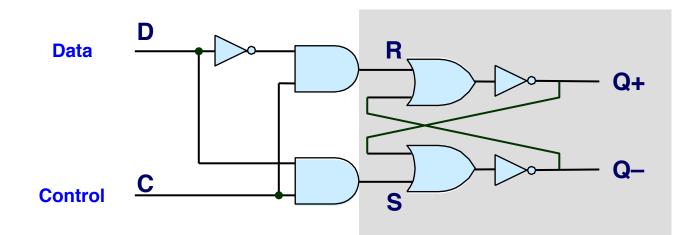
If R and S are different, Q+ is the same as S

Storing Data (Latching)



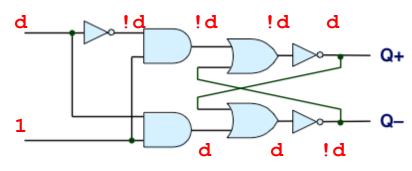
Q+ will continuously change as d changes



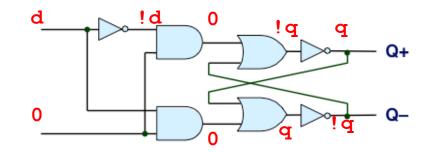


If R and S are different, Q+ is the same as S

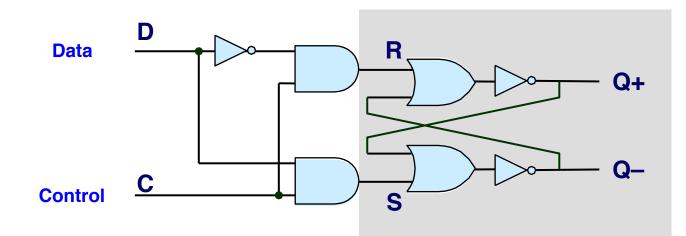
Storing Data (Latching)



Q+ will continuously change as d changes



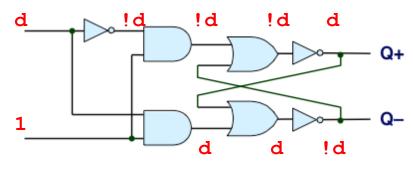
Q+ doesn't change with d



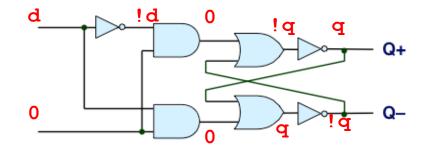
If R and S are different, Q+ is the same as S

Storing Data (Latching)

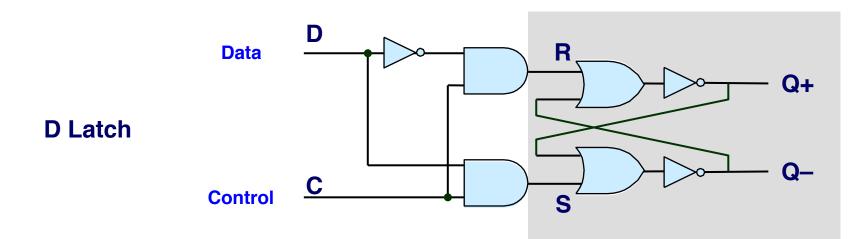
Holding Data



Q+ will continuously change as d changes



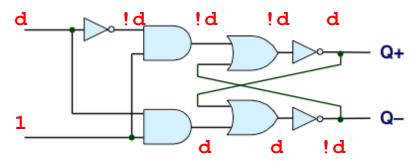
Q+ doesn't change with d



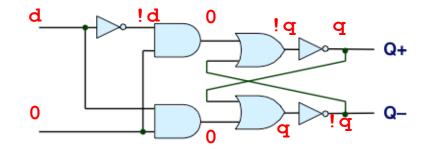
If R and S are different, Q+ is the same as S

Storing Data (Latching)

Holding Data

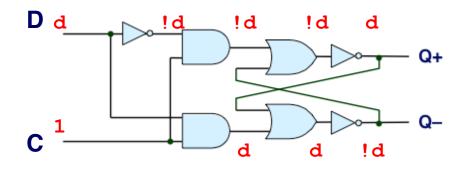


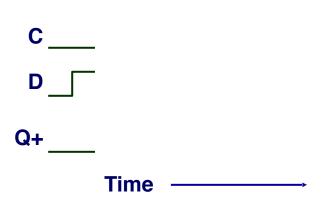
Q+ will continuously change as d changes



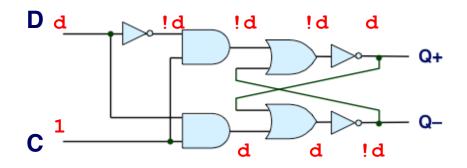
Q+ doesn't change with d

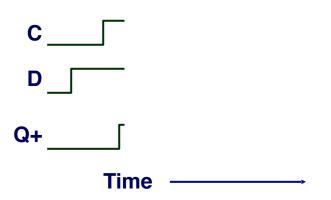
Latching



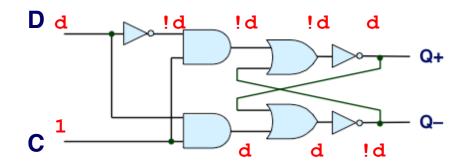


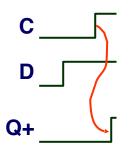
Latching





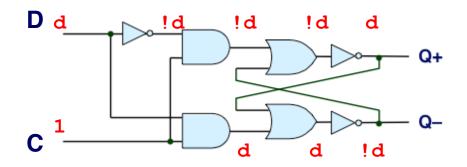
Latching

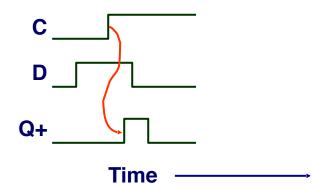




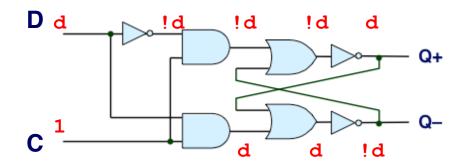


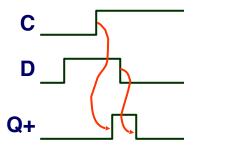
Latching





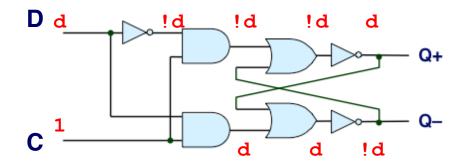
Latching

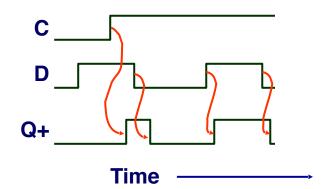




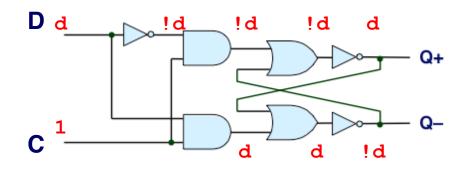


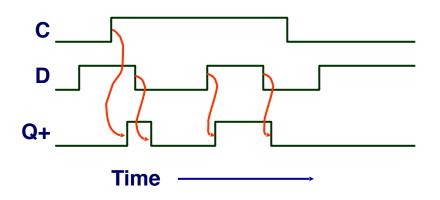
Latching





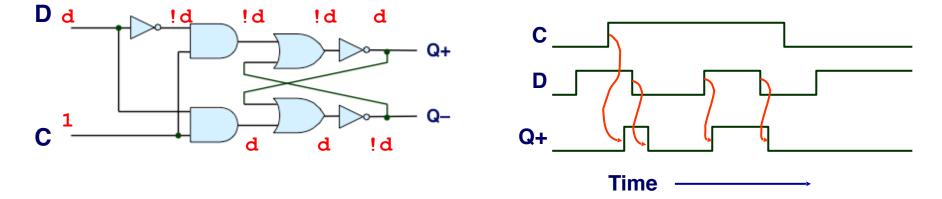
Latching





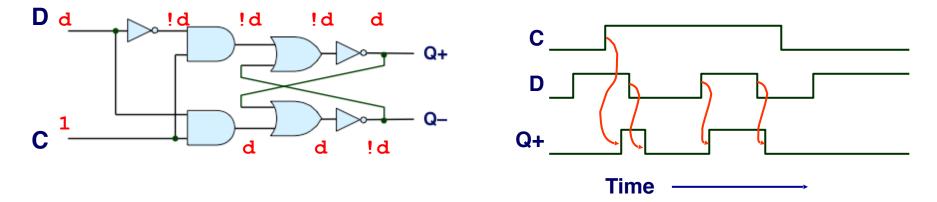
Latching

Changing D



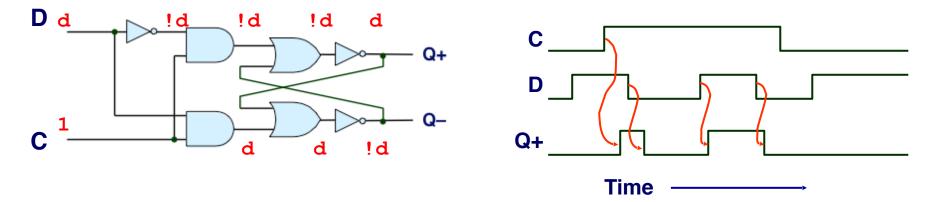
When you want to store d, you have to first set C to 1, and then set d

Latching



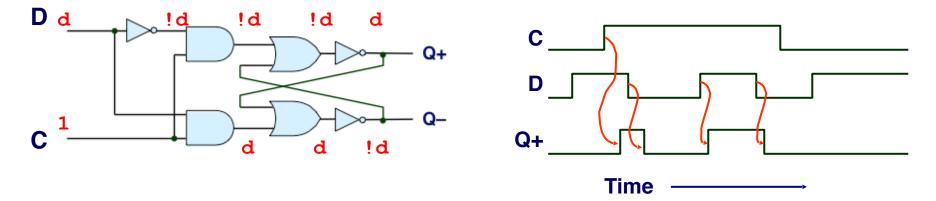
- When you want to store d, you have to first set C to 1, and then set d
- There is a propagation delay of the combinational circuit from D to Q+.
 So hold C for a while until the signal is fully propagated

Latching



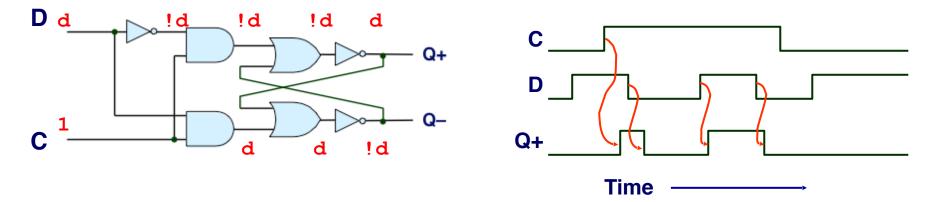
- When you want to store d, you have to first set C to 1, and then set d
- There is a propagation delay of the combinational circuit from D to Q+.
 So hold C for a while until the signal is fully propagated
- Then set C to 0. Value latched depends on value of D as C goes to 0

Latching



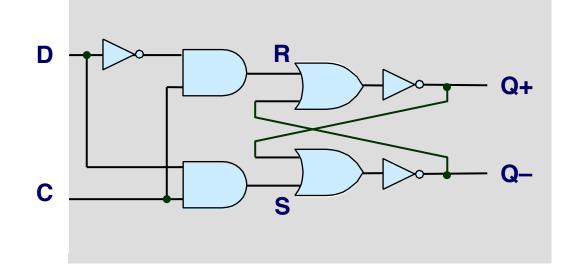
- When you want to store d, you have to first set C to 1, and then set d
- There is a propagation delay of the combinational circuit from D to Q+.
 So hold C for a while until the signal is fully propagated
- Then set C to 0. Value latched depends on value of D as C goes to 0
- D-latch is *transparent* when **C** is 1

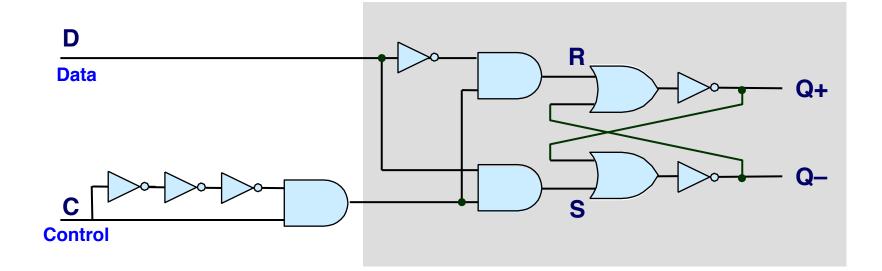
Latching

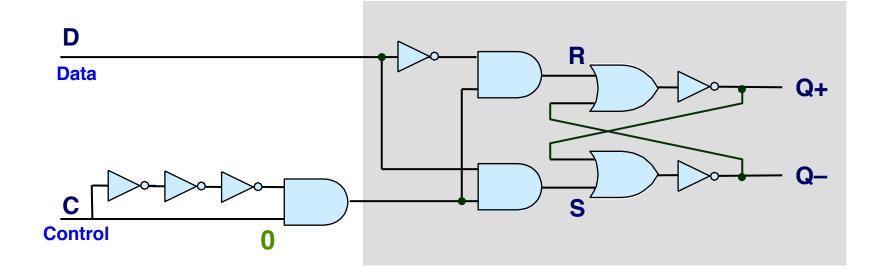


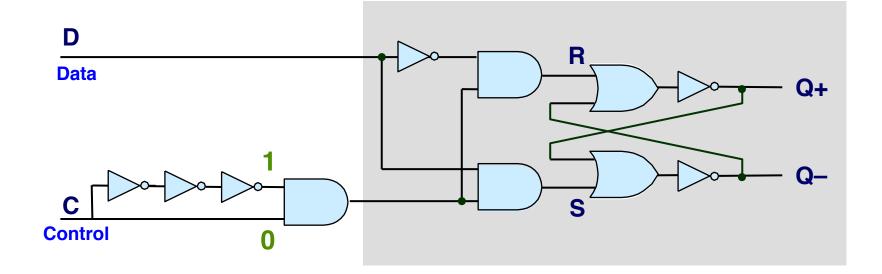
- When you want to store d, you have to first set C to 1, and then set d
- There is a propagation delay of the combinational circuit from D to Q+.
 So hold C for a while until the signal is fully propagated
- Then set C to 0. Value latched depends on value of D as C goes to 0
- D-latch is *transparent* when **C** is 1
- D-latch is "level-triggered" b/c **Q** changes as the voltage level of **C** rises.

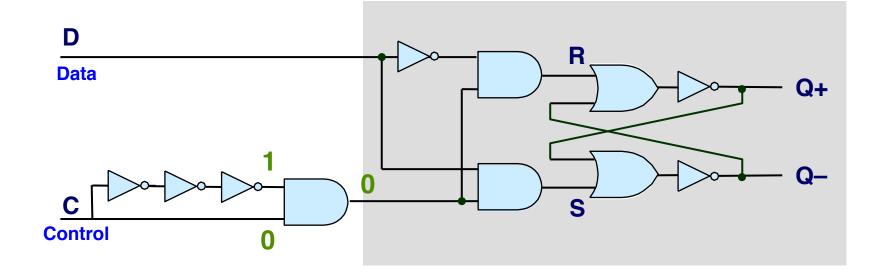
Edge-Triggered Latch (Flip-Flop)

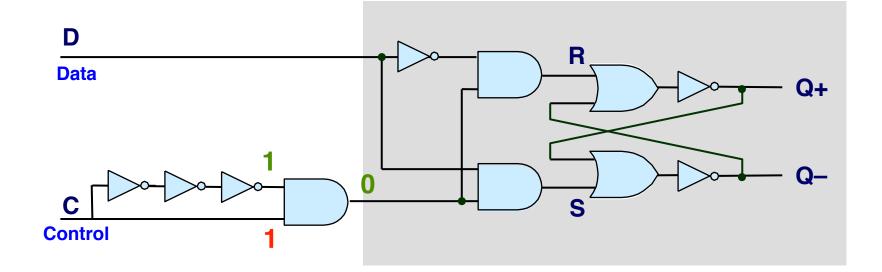


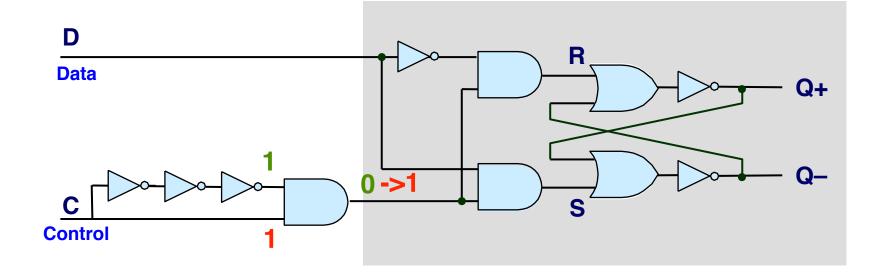


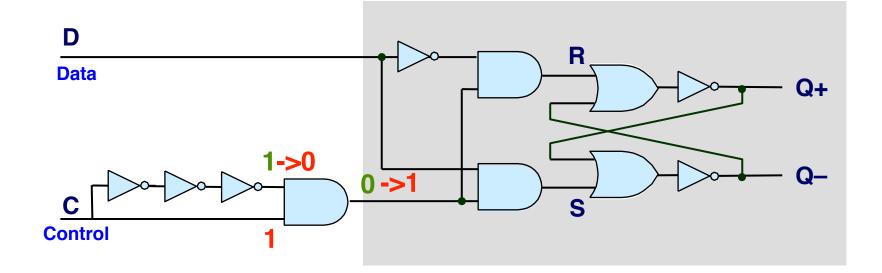


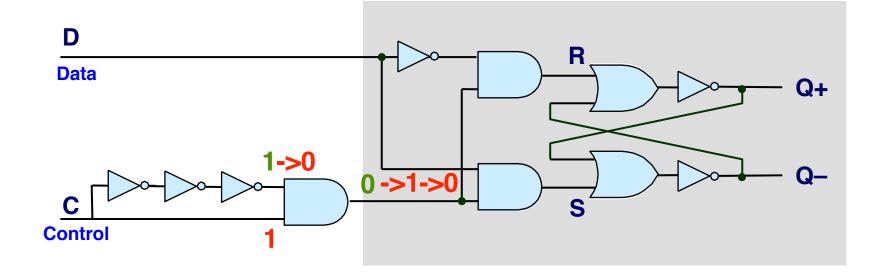


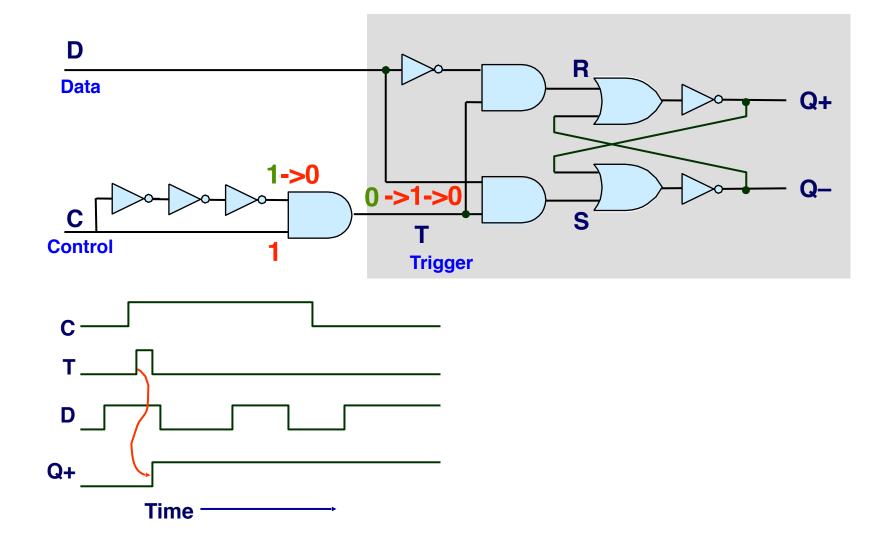


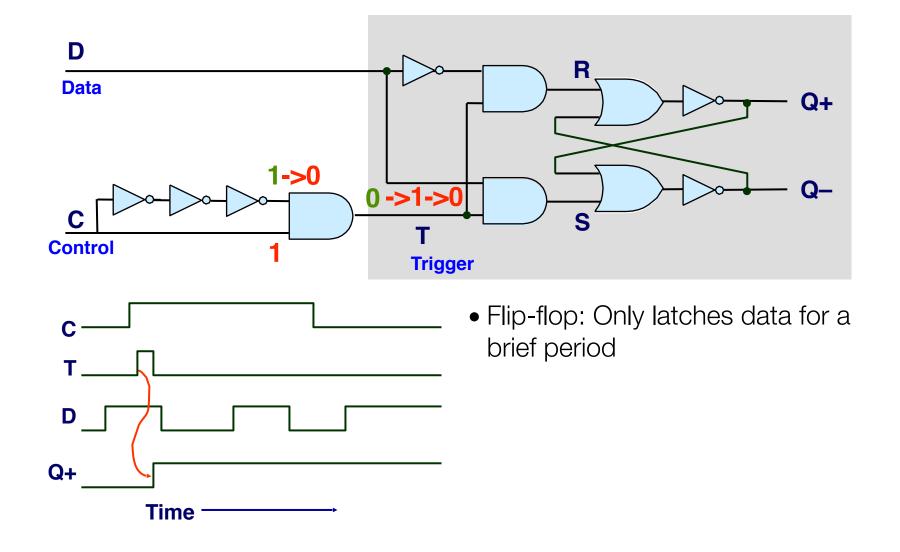


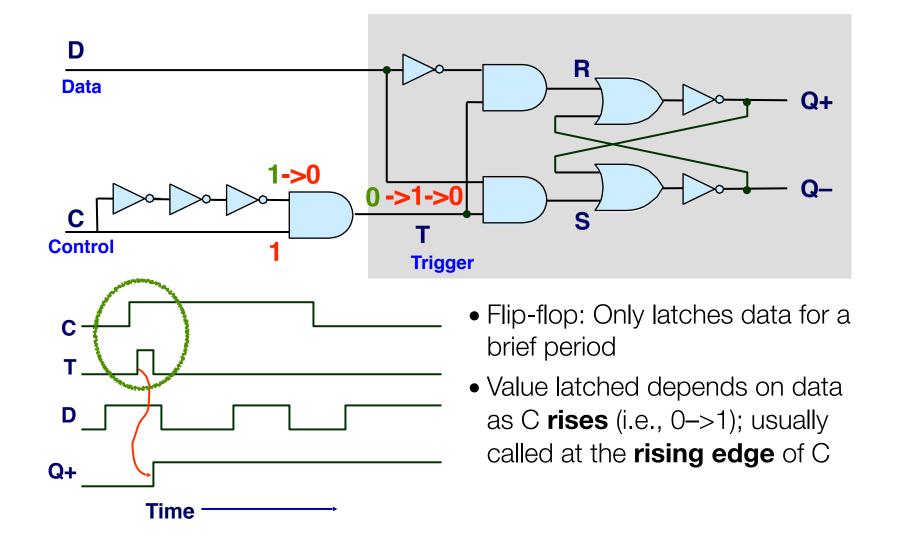


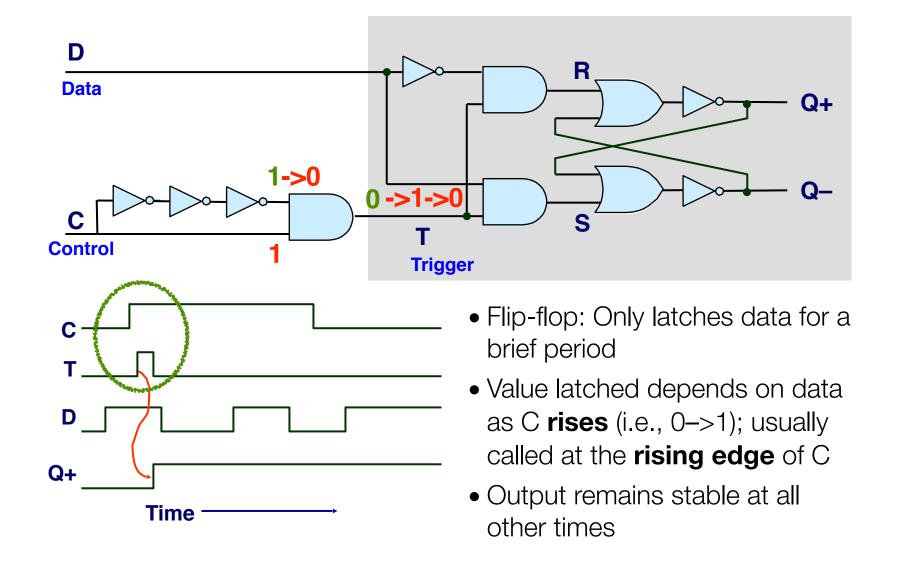




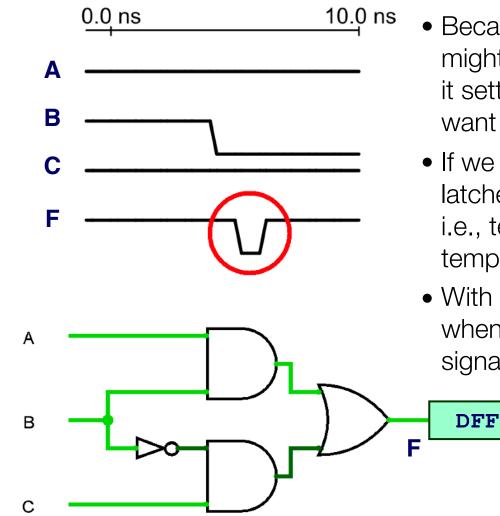








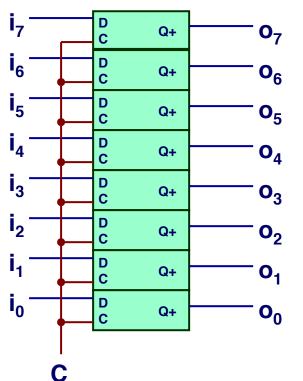
Why Use a Flip-Flop?



- Because the data we want to store might be temporarily changing before it settles down (due to glitch). We want to capture only the final value.
- If we had a transparent D latch, the latched value would change with F, i.e., temporal glitches will be temporarily stored as well.
- With a flip flop, we can store data only when its value settles: raise the control signal of the flop when F settles.

Registers

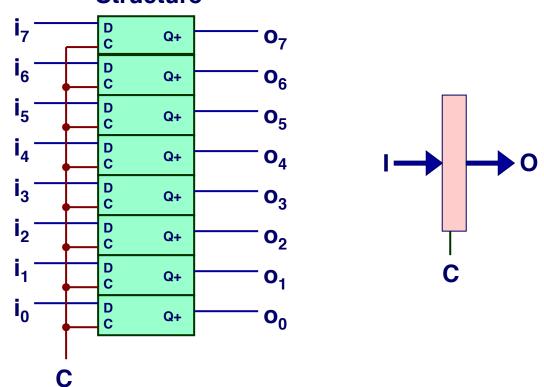
Structure



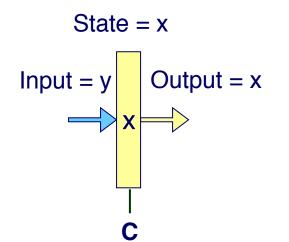
- Stores several bits of data
- Collection of edge-triggered latches (D Flip-flops)
- Loads input on rising edge of the C signal

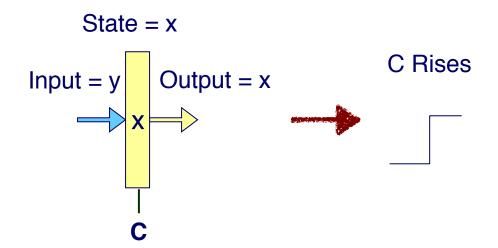
Registers

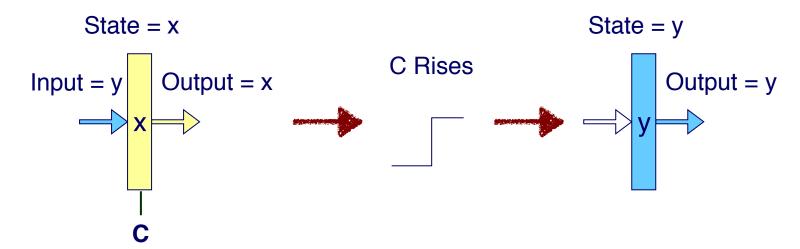
Structure

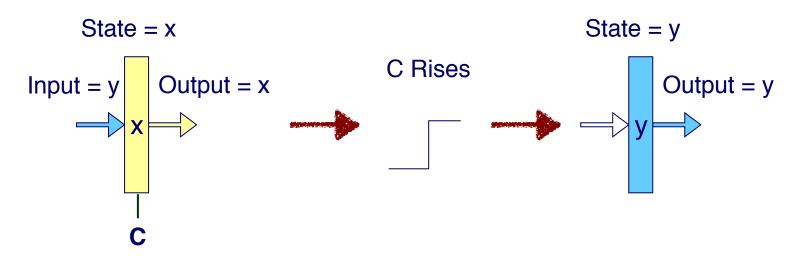


- Stores several bits of data
- Collection of edge-triggered latches (D Flip-flops)
- Loads input on rising edge of the C signal

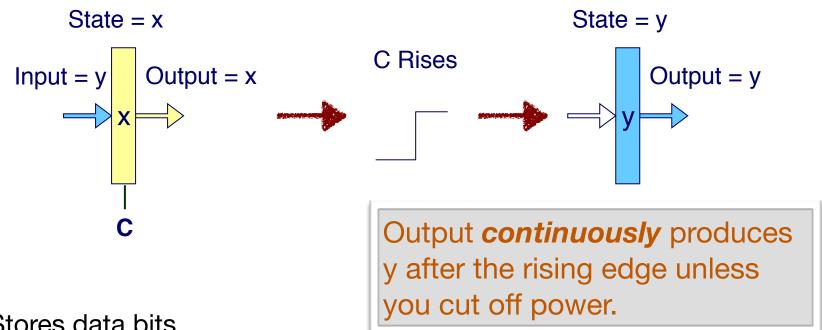




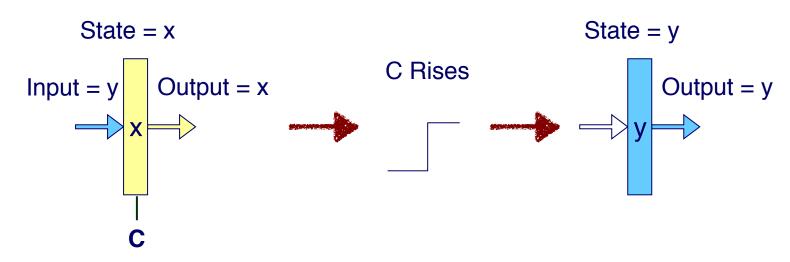




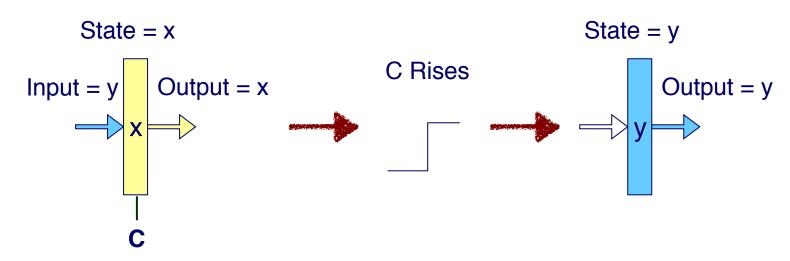
- Stores data bits
- For most of time acts as barrier between input and output
- As C rises, loads input
- So you'd better compute the input before the C signal rises if you want to store the input data to the register



- Stores data bits
- For most of time acts as barrier between input and output
- As C rises, loads input
- So you'd better compute the input before the C signal rises if you want to store the input data to the register

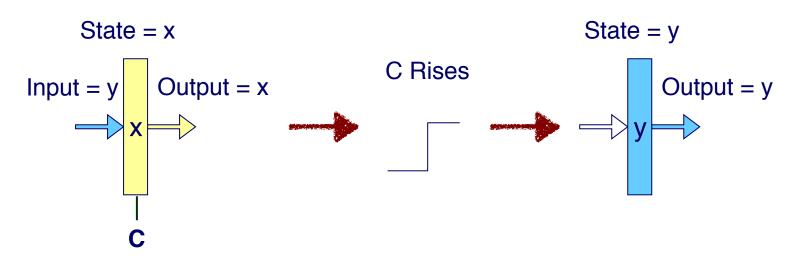


- A special C: periodically oscillating between 0 and 1
- That's called the **clock** signal. Generated by a crystal oscillator inside your computer.

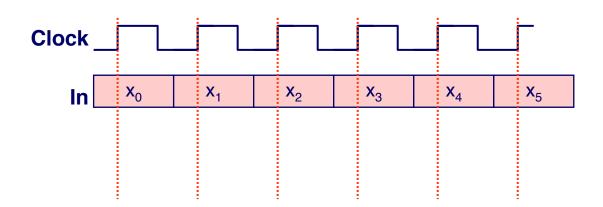


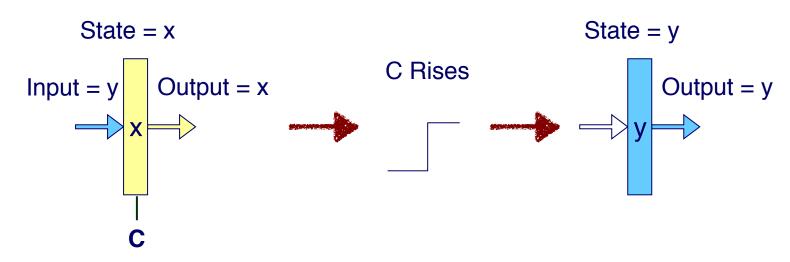
- A special C: periodically oscillating between 0 and 1
- That's called the **clock** signal. Generated by a crystal oscillator inside your computer.



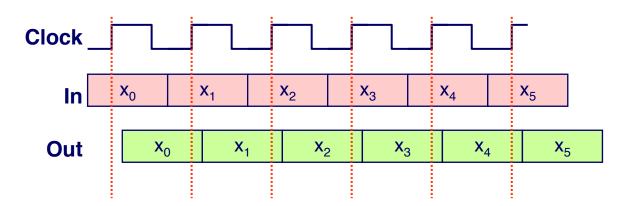


- A special C: periodically oscillating between 0 and 1
- That's called the **clock** signal. Generated by a crystal oscillator inside your computer.

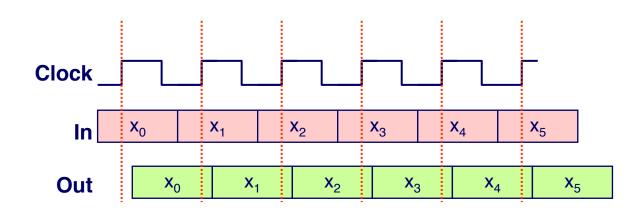




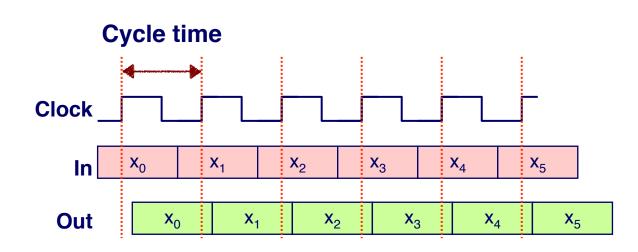
- A special C: periodically oscillating between 0 and 1
- That's called the **clock** signal. Generated by a crystal oscillator inside your computer.



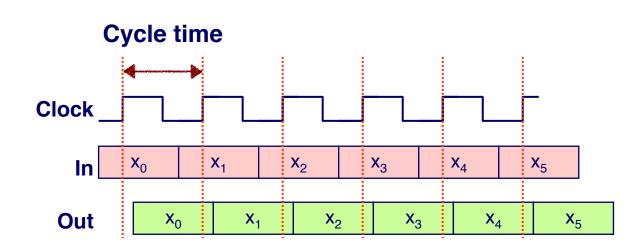
• Cycle time of a clock signal: the time duration between two rising edges.



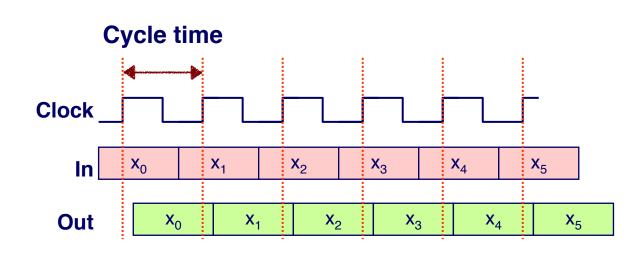
• Cycle time of a clock signal: the time duration between two rising edges.



- Cycle time of a clock signal: the time duration between two rising edges.
- Frequency of a clock signal: how many rising (falling) edges in 1 second.



- Cycle time of a clock signal: the time duration between two rising edges.
- Frequency of a clock signal: how many rising (falling) edges in 1 second.
- 1 GHz CPU means the clock frequency is 1 GHz



- Cycle time of a clock signal: the time duration between two rising edges.
- Frequency of a clock signal: how many rising (falling) edges in 1 second.
- 1 GHz CPU means the clock frequency is 1 GHz
 - The cycle time is $1/10^9 = 1$ ns

