ELEX 2117 : Digital Techniques 2
2024 Winter Term

Implementation of Digital Logic Circuits

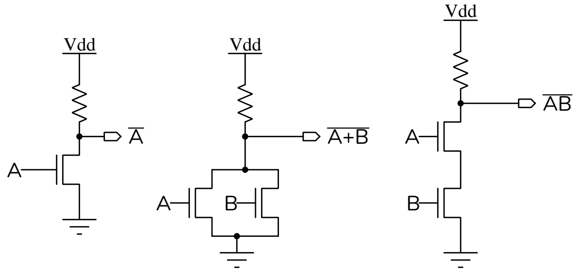
This lecture gives an overview of how of digital logic circuits are implemented.
After this lecture you should be able to: state which transistors are on and off in a CMOS totem-pole output; determine the direction of current flow between driver and receiver; determine from a data sheet: if an input or output voltage would be high, low or invalid and calculate noise margin; compute the effect of frequency and voltage changes on the power consumption of CMOS logic circuits; determine the RC time constant and current consumption of an open-collector output; describe the causes and consequences of ESD; design simple circuits to convert between logic levels; distinguish between DIP, TQFP, BGA and CSP packages.

Transistor Switches

Modern digital logic circuits are implemented using (enhancement-mode) MOSFETs.

As you've learned in other courses, an n-channel MOSFET allows current flow from drain to source when the gate is more positive than the source by more than a threshold voltage. Logic circuits are designed so the transistor is either fully on ($V_{\text{GS}} >> V_{\text{th}}$) or fully off ($V_{\text{GS}} \approx 0$).

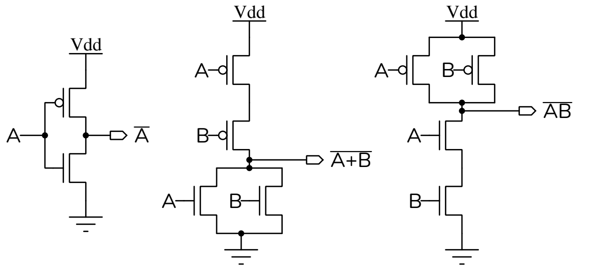
The following circuits shows how NOT, NAND and NOR gates could be implemented using n-channel MOSFETS[1]:



When one (for NOT), either A or B (for OR) or both A and B for (AND) inputs are high ($> V_{\text{th}}$) the transistors pull the output low. The resistor (typically implemented with a transistor structure) pulls the output high otherwise.

These "nMOS" circuits are relatively simple to manufacture since they use only n-channel MOSFETS. However, they draw current through the current-limiting resistance when the output is low. This resistance increases the rise time of the output which in turn limits the maximum clock rate, typically to a few MHz.

Modern digital circuits use a combination of n- and p-channel MOSFETS in complementary pairs, called CMOS, to eliminate the current-limiting resistance. This reduces power consumption and increases switching speeds. Examples of CMOS NOT, NAND and NOR gate circuits are:



In each case the appropriate combination of NMOS transistors are turned on to pull the output low. Otherwise the PMOS transistor(s) pull the output high.

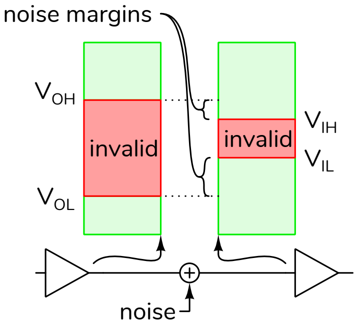
Exercise 1: In which direction does the output current flow when the output is high? When it is low? Which transistors are on in each case?

Due to the vertical arrangement of the p- and n-channel transistors in the schematic, this is often called a "totem-pole" output.

Logic levels and Noise Margins

Logic ICs are manufactured using different types of transistors – bipolar transistors (also known as Transistor-Transistor Logic or TTL) or MOSFET transistors. They are also designed for different power supply voltages (5V, 3.3V, 2.5V). Each of these "logic families" has different input and output levels corresponding to high and low logic levels. The diagram below shows the specifications for output and input logic levels:

[1] These MOSFET symbols are often used in IC schematics.



Exercise 2: Which of these specifications does the manufacturer guarantee? Which are requirements?

Noise is an unpredictable voltage or current superimposed on (added to) a signal. Noise in digital circuits results from voltage drops along shared power supply and ground conductors and inductive or capacitive coupling between conductors. As shown above, noise causes the input voltage to be different than the output voltage.

The noise voltage required for an output voltage to fall into the invalid input voltage range is called the *noise margin*:

- noise margin(low) = $V_{IL(max)}$ - $V_{OL(max)}$
- noise margin(high) = $V_{OH(min)}$ - $V_{IH(min)}$

A negative noise margin means that the interface may work properly because some output logic levels may not be recognized correctly.

Exercise 3: A logic family has V_{OH}(min) = 5 V, V_{OL}(max) = 0.5 V, V_{IH}(min) = 4 V and V_{IL}(max) = 1.5 V. What are the noise margins?

Power Consumption

MOSFET gates are insulated from both the source and the drain and can be considered as small capacitors that do not draw current except when being charged or discharged. Thus the average power consumption of a CMOS logic circuit will be proportional to the frequency at which the gates are switched and proportional to the *square* of the voltage (since both the voltage (V) and the current ($i = C\frac{dV}{dt}$) required to charge a capacitance (C) increase linearly with the voltage (V)).

When comparing power consumption in CMOS digital logic circuits for two operating conditions, the ratio of the power consumption is thus:

$$\frac{P_2}{P_1} = \frac{f_2}{f_1} \cdot \left(\frac{V_2}{V_1}\right)^2$$

Exercise 4: All else being equal, by how much would we expect to decrease power consumption when reducing logic levels from 5 V to 3.3 V? What would be the effect on power consumption in reducing the clock frequency from 50 MHz to 1 MHz?

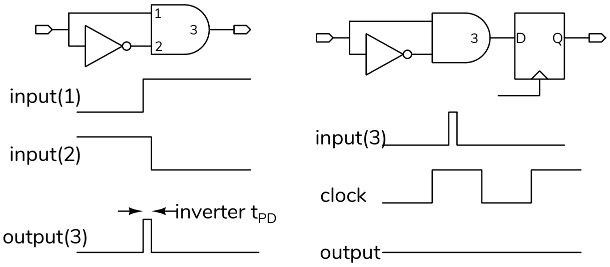
Exercise 5: The energy stored in a battery (its "capacity") is measured in Watt-hours. If a circuit draws 100 mA for 100 µs per second and draws 100 mA the rest of the time, how long will a 1000 mAh battery last?

Multi-Voltage ICs

Some IC's have different supply voltages for their internal logic circuits and for I/O. By using a lower voltage (e.g. 1.2 V) for internal ("core") logic and a higher voltage for I/O (e.g. 3.3 V) power consumption can be reduced while maintaining noise immunity.

Glitches

Glitches are short pulses that result from different propagation delays in a circuit. For example:



Glitches on IC outputs increase power consumption and noise. Increased power consumption results from charging and discharging interconnect capacitances for each glitch. The short glitch duration allows it to be readily coupled to other signals and to be radiated. This increases noise.

For these reasons it's important to avoid glitches on IC outputs. Putting a flip-flop at the output eliminates the glitch because the flip-flop is only updated at clock edges. It's good practice to "register" outputs as shown above.

Bidirectional Buses

A bidirectional pin is implemented by connecting both an input and a an output to the same pin:

To connect multiple outputs in parallel we must ensure that only one output is enabled at a time. Otherwise two devices could try to drive the same net to different levels. This is called contention and results in excessive current and, potentially, damage.

There are two approaches to avoiding contention between outputs.

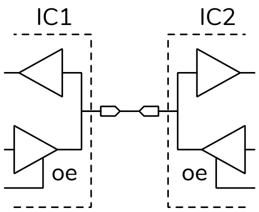
The first is a “tri-state” output together with an output-enable control. When the output is not enabled both totem-pole output transistors are turned off so that the output neither sources nor sinks current.

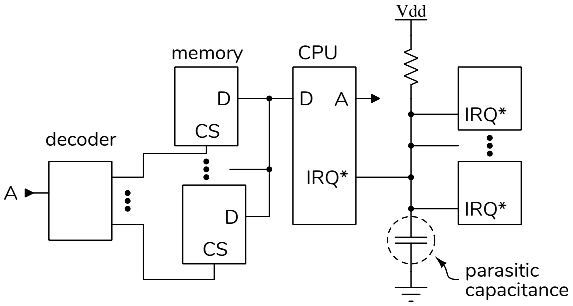
A circuit using tri-state outputs must be designed so that only one of the device’s enable signals is asserted at a time. A common application for tri-state outputs is a memory data bus that is connected in parallel to multiple memory devices. The bank-select (chip-select) signal is used to ensure that only one memory device’s output is enabled at any time.

The second approach is an “open-drain” output (called open-collector for bipolar output transistors). In this case the output can only sink, not source current and an external pull-up resistor is used to pull the bus high when no device is pulling it low.

A common application for this “wired-or” configuration is a microprocessor active-low interrupt input that can be connected in parallel to multiple devices’ interrupt request outputs. One disadvantages of an open-collector/open-drain output approach is that the circuit draws current whenever the signal is being pulled low. Another is that the rise time, determined by the RC time constant of the net, can be significant.

The following diagram shows examples of both types of bidirectional buses:





Exercise 6: What are the active-state current and the RC time constant for a wired-or interrupt-request line using a 10kΩ resistor pulling up a circuit with 50 pF capacitance to 3.3 V?

ESD Damage

If there are two objects with a charge difference (Q) and capacitance (C) between them[2], the result is voltage difference equal to $V = Q/C$. When these two objects are connected – equivalent to connecting the two leads of a capacitor – a current will flow from one to the other to equalize the charge. This is called electrostatic discharge (ESD).

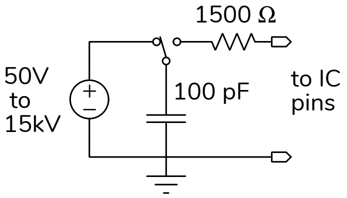
If the voltage is applied to an IC pin, it may be enough to damage the thin dielectric (insulating layer) in a semiconductor device. This will damage the device.

The effect of the damage may be a parameter such as increased leakage current or reduced input voltage tolerance that is not detected during manufacturing and testing but could cause failures when the device is stressed when in service.

Although most ICs have protection diodes to discharge excessive voltages and reduce the risk of ESD damage, this protection can be overwhelmed. Manufacturers and repair facilities thus follow procedures to ensure that semiconductors are not subject to ESD. Failure to follow these procedures could lose a company its quality certification and make it liable for in-service failures.

A “Human Body Model” is used for testing devices’ protection against ESD. This is a small capacitor charged to a test voltage and then discharged through a resistor to the pins device being tested. A sequence of pulses of different polarities is applied and then the functionality of the device is re-tested.

[2] Fun fact: depending on where you’re standing and your footwear, your body forms a capacitor of about 50–250 pF with “ground”.



Strict ESD procedures might be ignored when the consequences of device failure are small (e.g. teaching and prototyping labs).

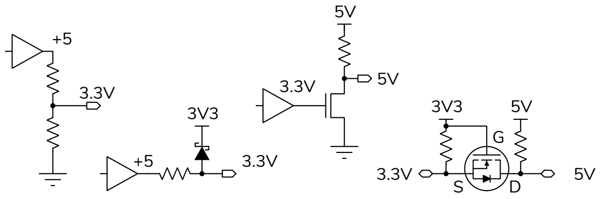
Latch-Up

Latch-up is a feature of CMOS devices that can cause PN junctions that are normally reverse-biased to turn on and short the power supply bus to ground. Latch-up results in high currents, can damage the device and can only be corrected by shutting off the power. A common source of latch-up is an input exceeding the supply voltage. This can happen when circuits with multiple power supplies are not powered up in the correct sequence. Although most ICs have protection against latch-up, circuits should be designed so that manufacturer's specifications for input levels relative to the ground and supply voltages are always met. This may require adding protection diodes and circuits to ensure that multiple supply voltages are turned on in an order that ensures each device's requirements are always met.

Logic Level Conversion

A common problem is interfacing logic families that use different, and incompatible, logic levels.

Although there are dedicated ICs designed for interfacing between different voltage levels, some simple solutions are shown below for situations where fast switching speeds are not required:



If the high output voltage is too high, a resistive voltage divider or a current-limiting resistor and a (Schottky) diode clamp[3] to the lower supply voltage can be used. If the high output voltage is too low[4], an open-collector output or n-channel MOSFET and pull-up resistor can be used.

If a bidirectional level conversion is required, an NMOS transistor can be used that allows either side to pull the other side low (when the 3.3V side is low the MOSFET is on and pulls the 5V side low; when the 5V side is low the body diode between the source terminal and the drain pulls the 3.3V side low).

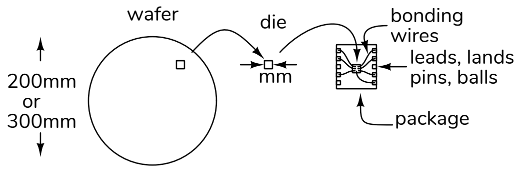
The disadvantages of these circuits, shared with open-collector outputs, is the steady-state low-level current drain and the RC time constant for the rise time of the signal.

IC Packages

The die that are cut from silicon wafers after manufacturing must be packaged to that they can be connected to circuits. Many packaging options have been developed.

In the simplest approach, "chip on board," the die is glued to a printed circuit board (PCB) and thin wires are bonded (welded) from conductive pads on the die to the PCB. A protective blob of epoxy is then applied over the die.

A more flexible approach is to attach the die to a plastic or ceramic package and use wire bonds to attach the IC's pads to pins, leads, lands or balls:

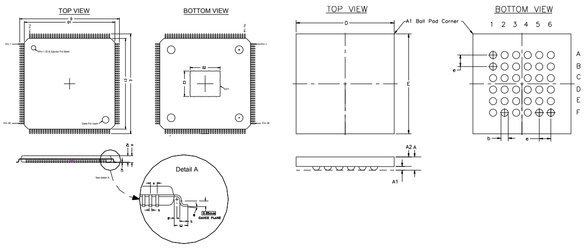


There are a wide variety of packages available. You've used dual-in-line (DIP) packages which are suitable for "through-hole" mounting. The IC on your CPLD or FPGA board uses a TQFP (thin quad flat package) which is suitable for "surface-mount" circuit boards. Higher densities can be achieved with BGA (ball-grid array) packages which have an rectangular array of connecting solder balls on the bottom.

[3] Sometimes a suitable clamp diode is included in the input of the lower-voltage device.

[4] Note that V_{OH}(min) for CMOS logic outputs with 3.3 V supplies is typically high enough to drive 5 V inputs directly.

The diagrams below show examples of the outlines for TQFP and BGA packages[5]. The e dimensions (pitch) are 0.5 mm and 0.4 mm and the D and E dimensions (width or length) are 22 mm and 3.5 mm for the 144-pin TQFP and the 64-pin BGA packages respectively.



Exercise 7: How many square mm of PCB area does each package require? Which packages have their pins accessible when the package is placed on the PCB?

BGA and TQFP packages have approximately the same pitch (distance between pins) but the number of BGA pins is proportional to the product of the two dimensions while the number of TQFP pins is proportional to the sum of the two dimensions. BGA's require more complex assembly but they are useful when the smallest possible package is required. They are the only choice for IC's with a large number (hundreds) of pins.

Modern ICs use "chip-scale" packages (CSP) that are approximately the same size as the die (millimetres). The die is flipped over and pads along the top or edge of the die are soldered to the PCB.

[5] From Intel datasheets.