Delay Components

Usually the delay consists of two major components

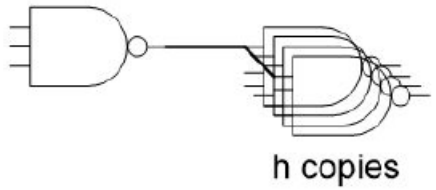
1. The ***parasitic delay*** is the time for a gate to drive its own internal diffusion capacitance
2. The ***effort delay*** depends on h and is the time for a gate to drive the load capacitance

In the last example, $\boldsymbol{t_{pdf}}$ was found to be (12 + 5h)RC or 12*RC + 5h*RC, where:

- 12*RC is the *parasitic delay* component
- 5h*RC is the *effort delay* component

Elmore Delay

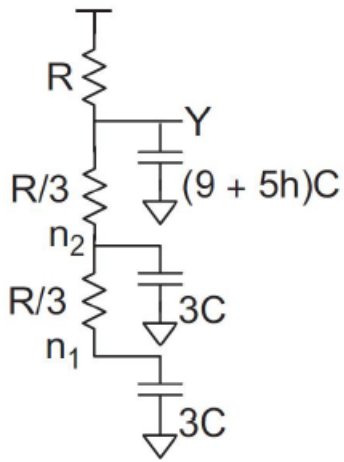
*Estimate t_{pdf}, t_{pdr}, t_{cdf}, and t_{cdr} for the 3-input NAND gate from Example 2 if the output is loaded with **h** identical NAND gates*

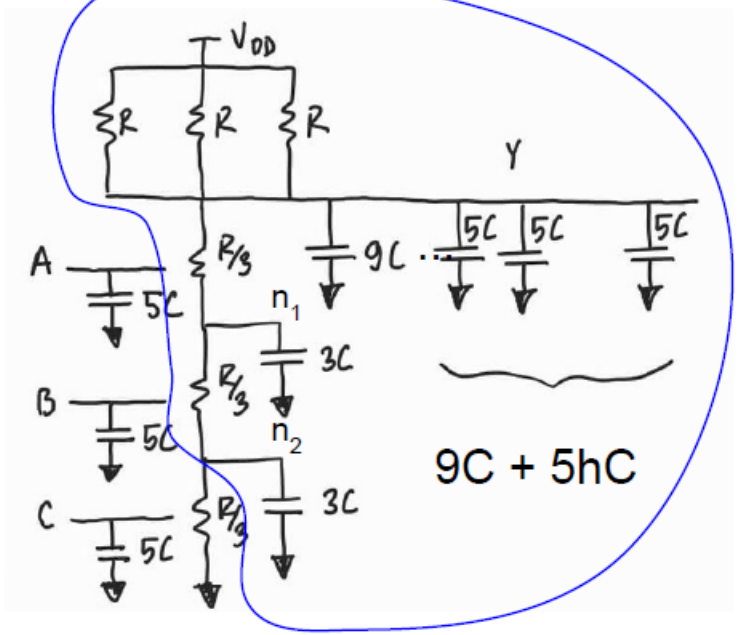


- t_{pdr} is the worst case rising delay
  - 2 inputs 1, 1 input 0 (A=1, B=1, C=0)
  - Nodes Y, n2 & n1 *all* have to charge

- The Elmore delay for the rising output is the sum of these RC products
- t_{pdr} = ((9 + 5h)C)(R) +
  (3C)(R) +
  (3C)(R)
  = (15 + 5h)RC

Here, the calculated Elmore delay is conservative and the actual delay is somewhat faster

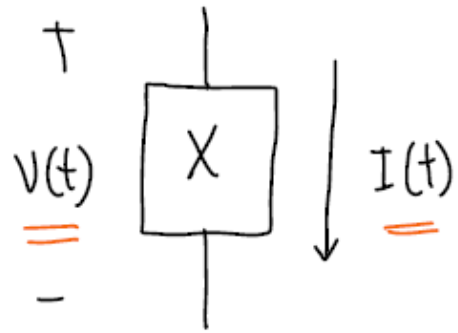




At any moment, an electronic element has ① a current through the element $I(t)$

② a voltage across the element $V(t)$

③ the relationship between I and V is called the I-V characteristics (usually governed by some law)



element "X"

Instantaneous power

The instantaneous power $P(t)$ consumed or supplied by a circuit element

$$p(t) = v(t)\, i(t)$$

Energy

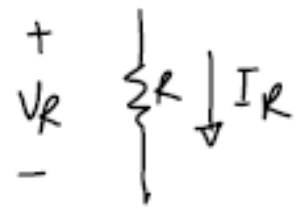
The energy consumed or supplied over some time interval "T", $E = \int_0^T p(t)\,dt$

Average power

The average power over this interval "T", $P_{avg} = \frac{E}{T} = \frac{1}{T}\int_0^T p(t)\,dt$

Examples

① Resistor

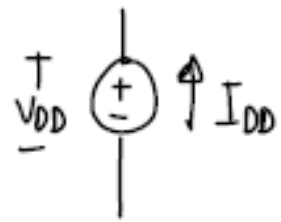


I-V : $V = IR$

Power : $P(t) = V_R(t) I_R(t) = \frac{V_R^2(t)}{R} = I_R^2(t) R$

Energy : $E = \int_0^T P(t) dt$

② Voltage source

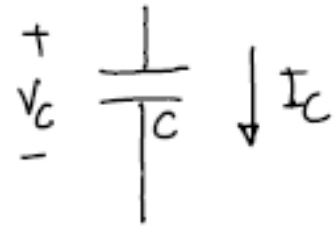


V = constant

$P(t) = V_{DD} \cdot I_{DD}(t)$

$E = \int_0^T P(t) dt = \int_0^T V_{DD} I_{DD}(t) dt$

③ Capacitor

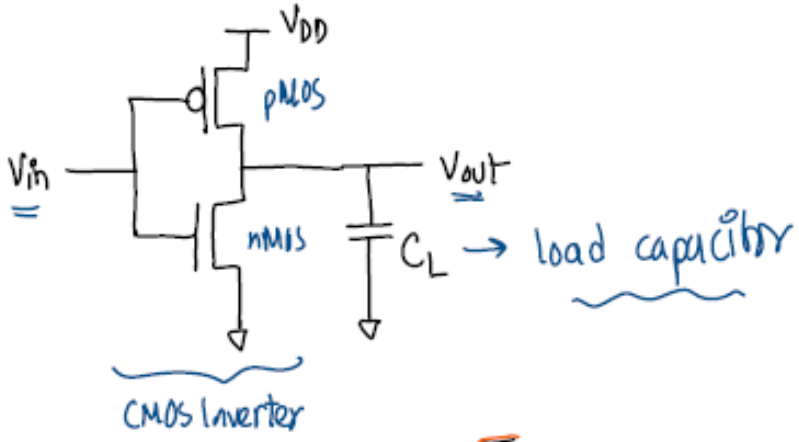


$I_C = C \frac{dV_C}{dt}$

$P(t) = V_C(t) I_C(t) = V_C(t) C \frac{dV_C(t)}{dt}$

$E = \int_0^T V_C(t) C \frac{dV_C}{dt} dt$

$= \int_0^{V_C} C V_C(t) dV_C = \frac{1}{2} C V_C^2$



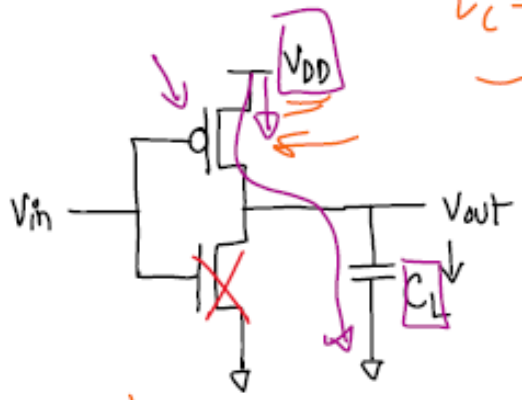
Case 1: Vin : 1 → 0 Vout: 0 → 1

Capacitor: Charging to V_{DD} through pMOS

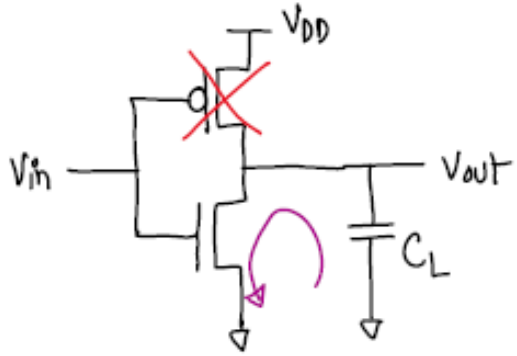
Case 2: Vin : 0 → 1 Vout: 1 → 0

Capacitor: Discharging to gnd through nMOS

$E_C = \frac{1}{2} C_L V_{DD}^2$



$I = C \frac{dV}{dt}$



$$E_{V_{DD}} = \int_0^\infty I(t) V_{DD}\, dt = V_{DD} \int_0^{V_{DD}} C_L \frac{dV_L}{dt}\, dt = V_{DD} \cdot C_L \cdot V_{DD}$$

$$= C_L V_{DD}^2$$

Switching power of CMOS :

V_{DD} charging & discharging C

Suppose, gate switches at some avererage frequency "f_{sw}" over some time "T".

∴ The load capacitor (C) will be charged and discharged : $f_{sw} \cdot T$ times

∴ The load capacitor (C) will be charged and discharged (switched): $f_{sw} \cdot T$ times

Energy required to charge & discharge a capacitor (C) with supply (V_{DD}) once: CV_{DD}^2

Total energy required for switching, $E_{switching} = f_{sw} T \cdot CV_{DD}^2$

Average power dissipated during this period, $P_{switching} = \frac{E_{switching}}{T} = \frac{f_{sw} T C V_{DD}^2}{T} = f_{sw} C V_{DD}^2$

Define $\alpha \triangleq \frac{f_{sw}}{f}$ where f is the clock frequency, α is activity factor

$\Rightarrow f_{sw} = \alpha \cdot f$

$P_{switching} = f_{sw} \cdot CV_{DD}^2 = \alpha f C V_{DD}^2$

$V_{DD} \rightarrow$ supply voltage

$C \rightarrow$ capacitance

$f \rightarrow$ clock frequency

$\alpha \rightarrow$ activity factor

$$\boxed{P_{switching} = \alpha f C V_{DD}^2}$$

Sources of power dissipation:

① Dynamic power dissipation

→ charging & discharging load capacitances as gates switch $P_{switching}$

→ "short-circuit" current when the PUN & PDN are momentarily "ON" $P_{shortcircuit}$

$$P_{dynamic} = P_{switching} + P_{shortcircuit}$$

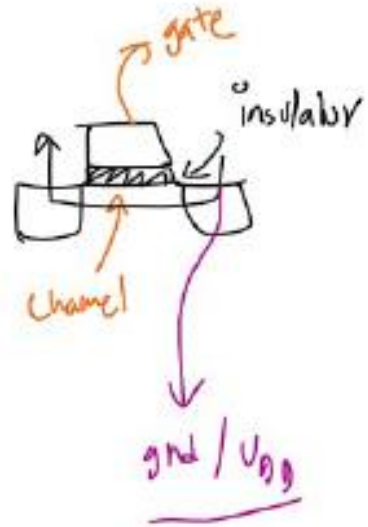
② Static power dissipation:

→ subthreshold leakage through "off" transistors I_{sub}

→ gate leakage through gate dielectric insulator I_{gate}

→ junction leakage from source/drain diffusions I_{junct}

→ contention current in ratioed circuits $I_{contention}$



$$P_{static} = \left(I_{sub} + I_{gate} + I_{junct} + I_{contention} \right) V_{DD}$$

$$P_{total} = P_{dynamic} + P_{static}$$

Problem

A digital system-on-chip in a **1 V** 65 nm process (with 50 nm drawn channel lengths and **λ = 25 nm**) has 1 billion transistors, of which 50 million are in logic gates and the remainder in memory arrays. The **average logic transistor width is 12 λ** and the **average memory transistor width is 4 λ**. The memory arrays are divided into banks and only the necessary bank is activated so the **memory activity factor is 0.02**. The static CMOS **logic gates have an average activity factor of 0.1**. Assume each transistor contributes **1 fF/μm of gate capacitance** and **0.8 fF/μm of diffusion capacitance**. Neglect wire capacitance for now (though it could account for a large fraction of total power). **Estimate the switching power when operating at 1 GHz.**

logic transistor

$N_{logic} = 50 \times 10^6$

$V_{DD} = 1\,V$

memory transistor

$N_{mem} = 950M = 950 \times 10^6$

$V_{DD} = 1\,V$

$V_{DD} = 1V$

$\lambda = 25nm$

$W_{logic} = 12\lambda$

$\alpha_{logic} = 0.1$

total cap of 1 transistor = gate + dif

of unit width

$= 1ff/\mu m + 0.8ff/\mu m$

$= 1.8 ff/\mu m$

total cap of 1 tran $= 1.8 ff/\mu m \times W_{logic}$

$= 1.8ff/\mu m \times 12\lambda$ ($\lambda = 25nm = 25\times10^{-3}\mu m$)

$= 1.8 ff/\mu m \times 12 \times 25\times10^{-3}\mu m$

$= 0.54\ ff$

$V_{DD} = 1V$

$\lambda = 25nm$

$W_{mem} = 4\lambda$

$\alpha_{mem} = 0.02$

total cap. of 1 unit width

tran. = gate + dif

$= 1.8 ff/\mu m$

total cap. 1 tran $= 1.8 \times W_{mem}$

$= 1.8ff/\mu m \times 4\lambda$

$= 1.8ff \times 4 \times 25\times10^{-3}$

$= 0.18\ ff$

total cap $= N_{logic} \times 0.54ff$

$P_{switching(logic)} = \alpha_{logic} C_{logic} f V_{DD}^2$

$\times 0.54 \times 10^{-15}$

$= 0.1 \times 50 \times 10^6 \times f V_{DD}^2$

total cap $= N_{mem} \times 0.18ff$

$P_{switching(mem)} = \alpha_{mem} \cdot C_{mem} \cdot f V_{DD}^2$

$= 0.02 \times 950 \times 10^6 \times 0.18$

$\times 10^{-15} \times f \times V_{DD}^2$

$P_{switching} = P_{switching(logic)} + P_{switching(memory)}$

Suppose the technology you are using to design a VLSI system has $\lambda = 80$ nm, a clock frequency of 5 MHz, and a supply is 5 V. The chip you are designing has 5 million transistors, of which 1 million remain active at any given time. The activity factor is defined by the fraction of total components which remain active. The gate and diffusion capacitances are 12 fF/μm and 5 fF/μm, respectively for all the 5 million transistors. The gate width is 20λ. You also obtain the following power consumption data:

- Short circuit power = 0.5 W
- Leakage power = 0.01 W
- Subthreshold power = 0.02 W

The **acceptable TOTAL power consumption** of a chip is **3 W.**

(a) **Find** the activity factor and load capacitance of the system described above.

(b) **Calculate** the switching power consumption of the chip.

(c) **Calculate** the dynamic and static power of the chip. Thereafter calculate the **TOTAL** power consumption.

(d) Is the **TOTAL** power consumption within the acceptable range? If not, **find** the maximum clock frequency to keep the **TOTAL** power within the acceptable range?

(e) For a **10-input NAND gate**, what should be the ratio of width scaling factors – $\boldsymbol{k_p/k_n}$ - for the NMOS and PMOS to keep the rise and fall resistances equal?

4 no. Ans

(a) Out of 5 million transistor only 1 million remain active.

So, activity factor, $\alpha = \frac{1 \text{ million}}{5 \text{ million}}$

$= \frac{1}{5}$

$= 0.2$

gate width, $w = 20\lambda$

Load capacitance, $C_{load\ unit} = (C_g + C_{diff}) \times w$

$= (12+5)\ fF/\mu m \times 20\lambda$

$= 17\ fF/\mu m \times 20 \times 10^{-15} \times 80 \times 10^{-3}\ \mu m$

~~$= 2.72 \times 10^{-8}\ F$~~

$= 2.72 \times 10^{-14}\ F$

02

(ii) $C_{load\ total} = 5 \times 10^6 \times C_{load\ unit}$

$= 0.02772\ C$ ~~$1.36 \times 10^{-7}\ C$~~

$V_{DD} = 5\ V$

$f = 5 \times 10^6\ Hz$

(iii) $P_{switching} = \alpha\ C_{total}\ V_{DD}^2\ f$

$= 0.2 \times 1.36 \times 10^{-7} \times (5)^2 \times 5 \times 10^6$

$= 680000\ watt$ ~~$3.4\ W$~~

~~C_{load}~~ 03

(c)

$P_{dynamic} = P_{switching} + P_{shortcircuit}$

$= (3.4 + 0.5)\ W$

$= 3.9\ W$

~~Stati~~ $P_{static} = P_{leak} + P_{sub}$

$= \cancel{0.02 + 0}\ 0.01 + 0.02$

$= 0.03\ W$

$P_{total} = P_{dynamic} + P_{static}$

$= 3.9 + 0.03$ (03)

$= 3.93\ W$

(d) No, It is not in acceptable range.

For keep in power rang we need to find the $P_{switching}$ value

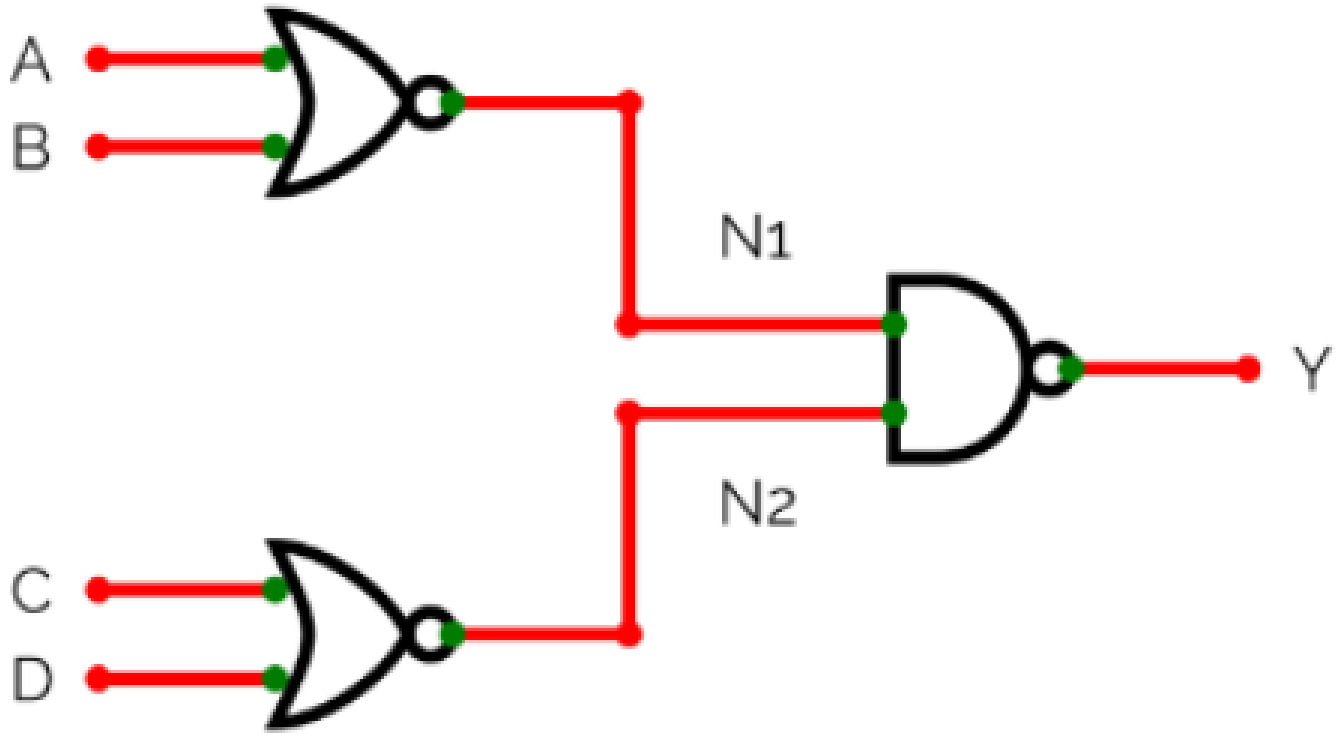
$P_{switching} = P_{total} - P_{static} - P_{short\ circuit}$

$= (3 - 0.03 - 0.5)\ W$

$= 2.47\ W$

Q3. [CO3, CO4] - Marks 10

Consider the following logic circuit, which is being driven by a system frequency of 1 GHz and a supply voltage of 3.2 V. The inputs to the circuit are A, B, C, and D, while the output node is Y. N1 and N2 are two intermediate nodes. The activity factors of nodes N1, N2 , and Y with respect to the system frequency are given in table 1. The input and output capacitances of the individual logic gates are listed in table 2.



Node	Activity factor
N_1	$\frac{1}{4}$
N_2	$\frac{1}{4}$
Y	$\frac{15}{16}$

Table 2

Gate	Input capacitance	Output capacitance
NAND-2	4 pF	6 pF
NOR-2	5 pF	6 pF

- 1 GHz = 10^9 Hz
- 1 pF = 10^{-12} F

(a) Find out the total capacitances at nodes N1, N2, and Y. **[3]**

(b) Compute the total switching power of the logic circuit, assuming no power is being consumed at nodes A, B, C, and D. **[8]**

(c) Prove that a pMOS transistor cannot properly pass a low voltage signal (corresponding to a logical 0). **[4]**

However, he has messed up the layout design of a **3-input NOR gate** (Figure 2(a)) because of which, the NOR gate has **an equivalent rise resistance that is four times its equivalent fall resistance**. Also, the NOR gate's **equivalent rise resistance is twice the rise resistance of a unit inverter** for that design process. But at least he made sure to keep the source-drain terminals shared where possible to reduce capacitance. The gate capacitance and parasitic (diffusion) capacitance of are **0.5 fF/µm.**

This **NOR gate** is loaded with a **Unit inverter** and the output node **Y** has a voltage waveform as shown in Fig. 2(c).

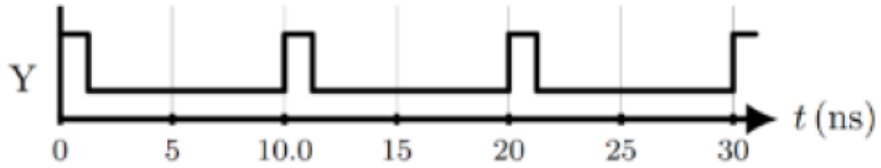


Figure 2(c): Voltage waveform at node **Y**

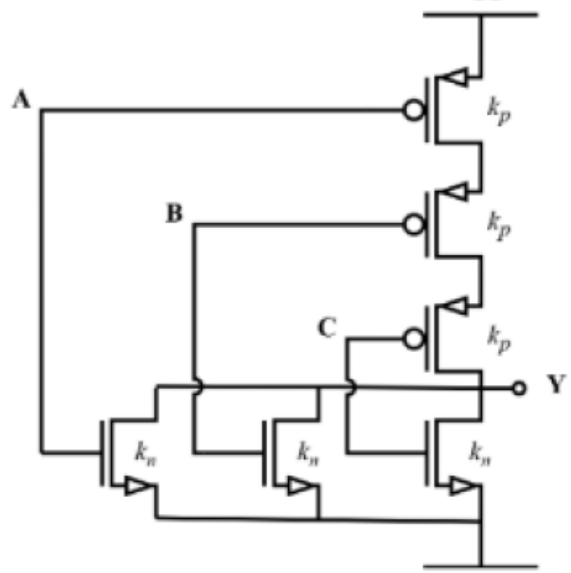


Figure 2(a): CMOS implementation of a 3 - input NOR gate

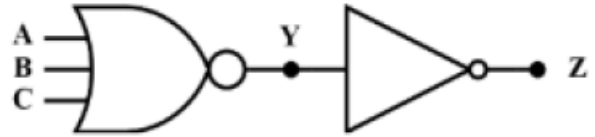


Figure 2(b): A 3-input NOR gate loaded with a **Unit Inverter**

(a) **Find** the value of width scaling factors - k_p and k_n of the NOR gate designed by Brook. Thereafter, calculate the actual width of the **pMOS** and **nMOS**. **[2]**

(b) **Draw** the RC equivalent model of the above 3-input NOR gate. Lump the capacitances together where possible. **[5]**

(c) **Find** the capacitance of the **Y** node (in **fF** units), when it is loaded with a unit inverter, as shown in Fig. 2(b). **[3]**

(d) Find the best-case falling edge **contamination delay (t_{cdf})** of **Y**-node in terms of **R** and **C** of unit MOS. **[2]**
[**Hint:** You do not need Elmore modelling. All three PMOS are off. Three NMOS resistances are in parallel.]

(e) **Find** the switching power of the **Y**-node, if the system has a supply voltage of **5 V**. **[3]**