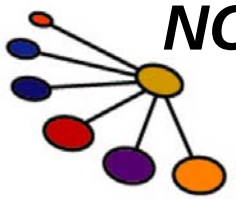
EE-612:
Lecture 26:
CMOS Limits

Mark Lundstrom
Electrical and Computer Engineering
Purdue University
West Lafayette, IN USA
Fall 2006



NCN

www.nanohub.org

Outline

1) Review: CMOS Metrics
2) MOSFET limits
3) Circuit limits
4) System Limits

CMOS metrics

1) Switching energy: $E_S = \frac{1}{2} C V_{DD}^2$

2) Switching delay: $\tau_S = \frac{C V_{DD}}{I_D(\text{ON})}$

3) Dynamic power: $P_D = \alpha f C V_{DD}^2$

4) Energy-delay product: $E_S \tau = \frac{1}{2} \frac{C^2 V_{DD}^3}{I_D(\text{on})}$

Outline

1) CMOS Metrics
2) **MOSFET limits**
3) Circuit limits
4) System limits

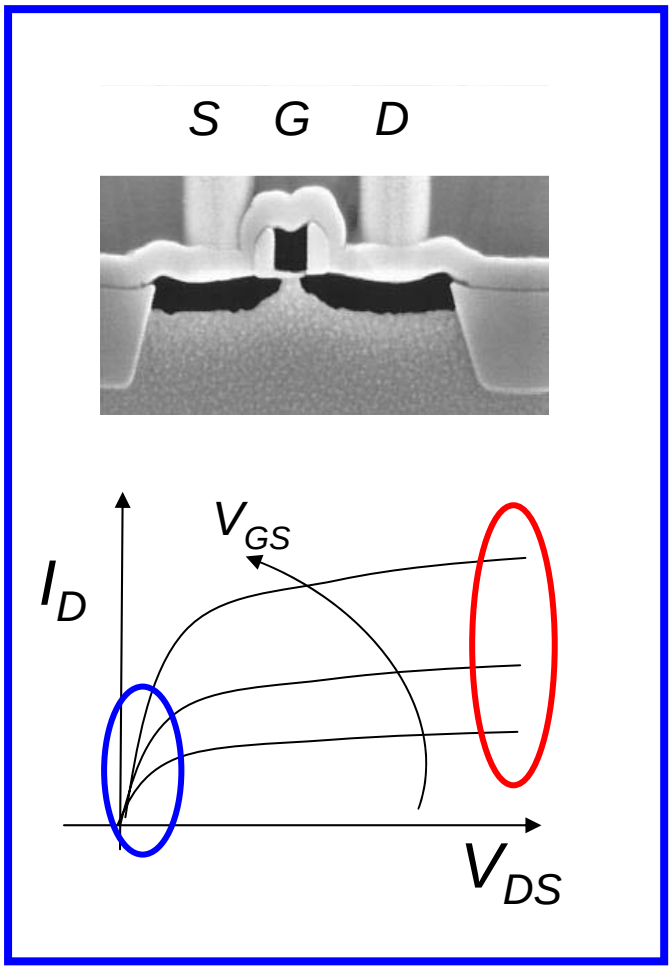
device limit questions

1) minimum energy dissipation per logic transition
2) minimum channel length
3) maximum device density
4) minimum device delay
5) power density
6) power-limited device density
7) CMOS vs. the ultimate switch

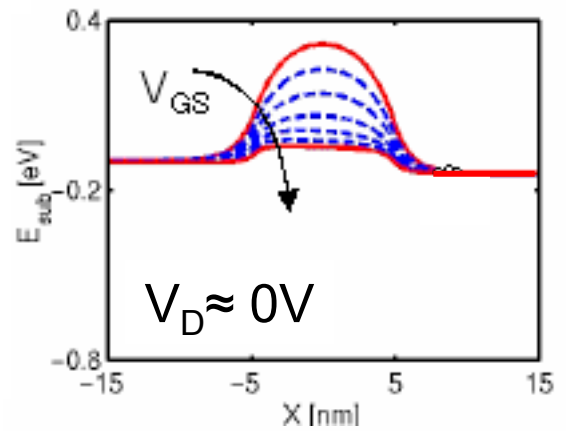
acknowledgment

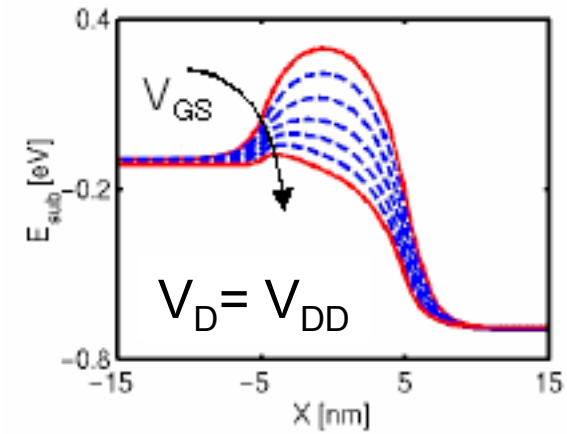
The approach that I take here is similar to the work of V.V. Zhirnov, R.K. Cavin, J.A. Hutchby, and G. Bourianoff, “Limits to Binary Logic Switch Scaling - A Gedankan Model,” *Proc. IEEE*, Special Issue on Nanoelectronics and Nanoscale Processing, Nov. 2003.

MOSFETs control current with potential barriers



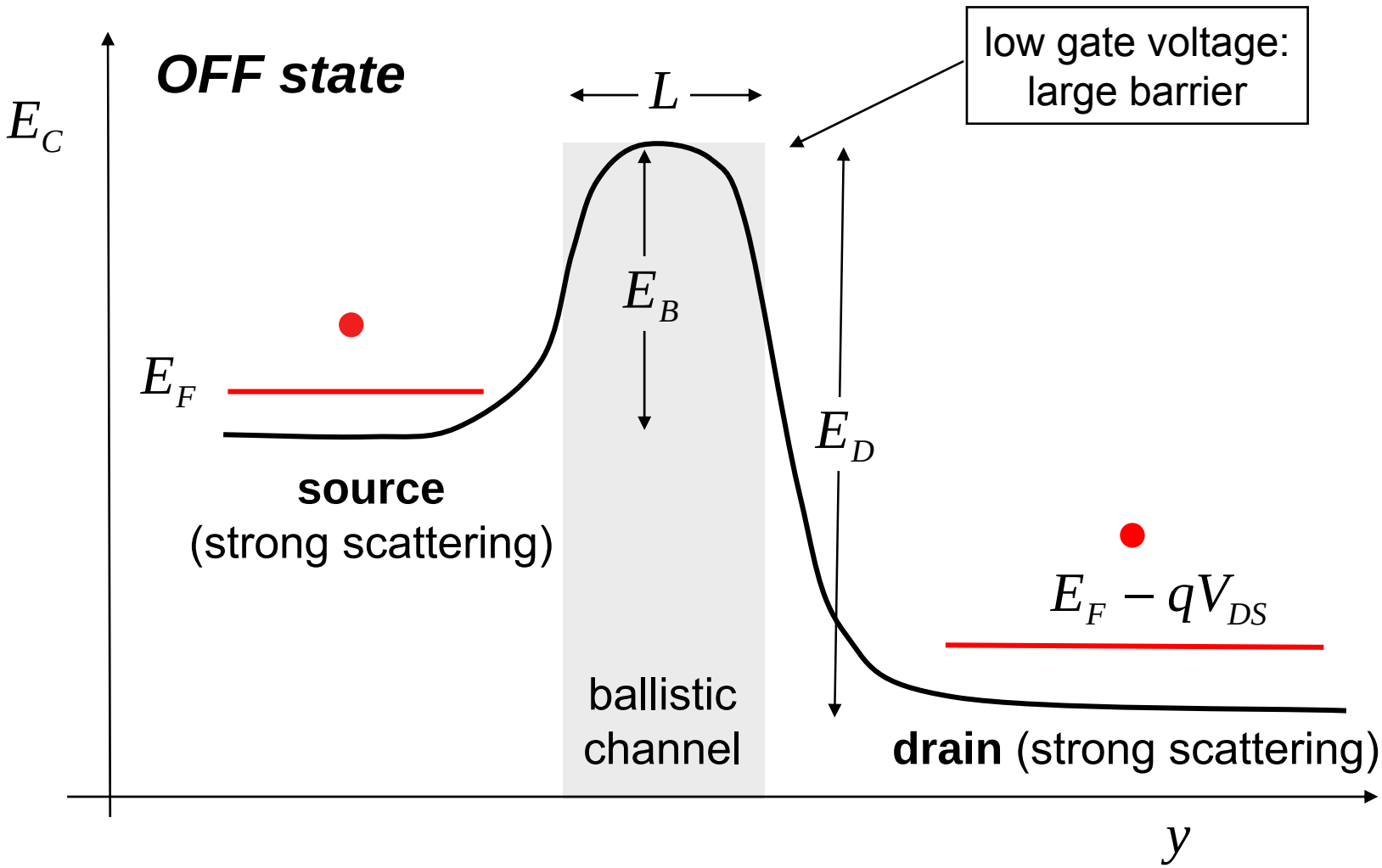
electron energy vs. position



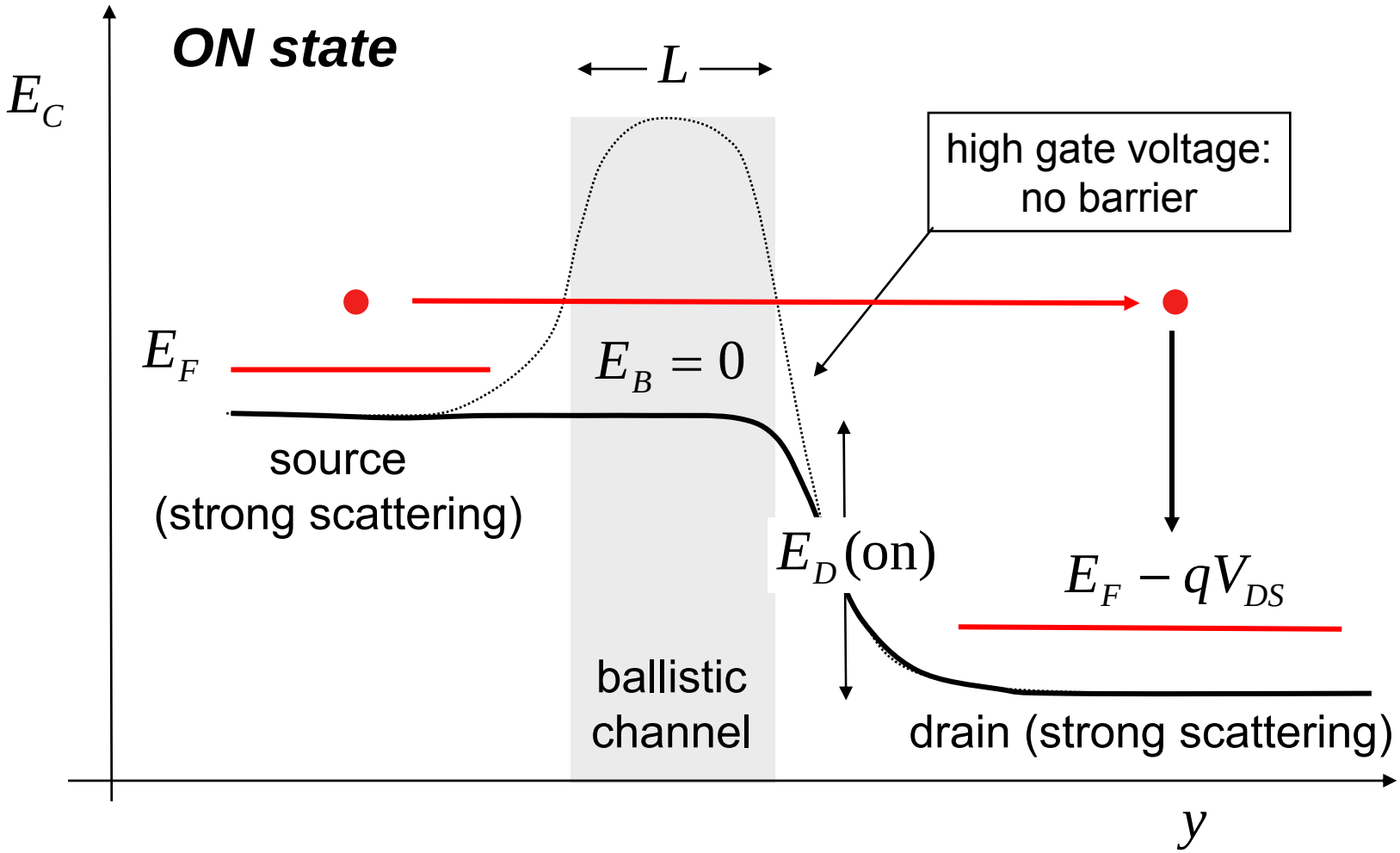


E.O. Johnson, *RCA Review*, 1971

ultimate MOSFET in the off-state



ultimate MOSFET in the on-state

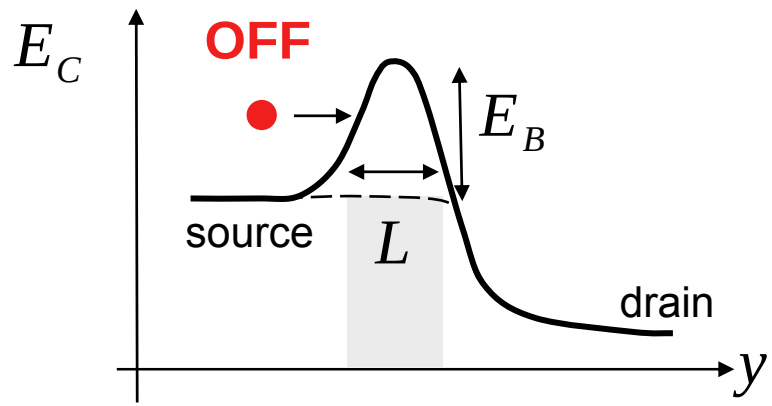


minimum switching energy

To distinguish **off** from **on**, electrons in the source must have less than a 50:50 chance of moving over the barrier from the source to drain.

$$e^{-E_B/k_BT} < \frac{1}{2}$$

$$E_B > E_{\min} = k_B T \ln(2)$$



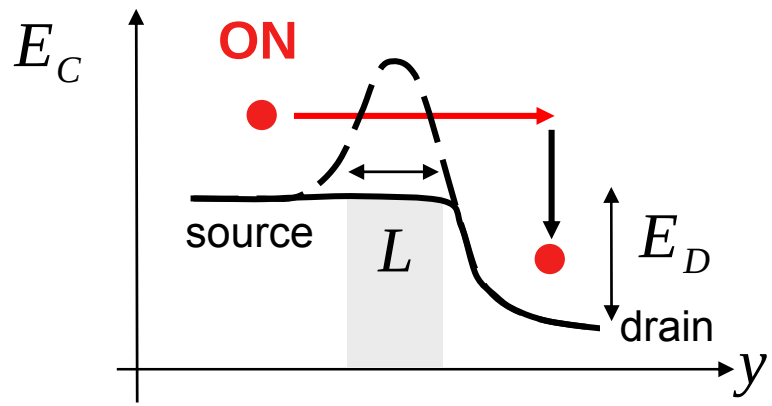
minimum switching energy (ii)

To distinguish **on** from **off**, electrons in the drain must have less than a 50:50 chance of moving over the barrier from the **drain to source.**

$$e^{-E_D/k_BT} < \frac{1}{2}$$

$$E_D > E_{\min} = k_BT\ln(2)$$

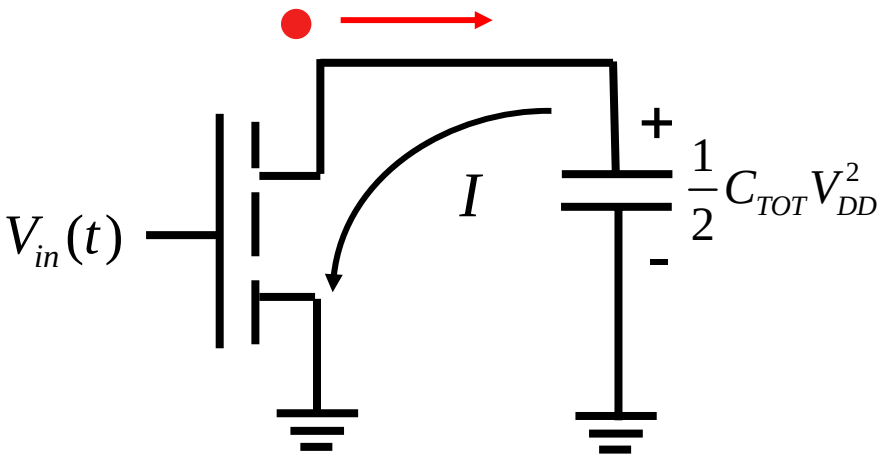
$$E_S > k_BT\ln(2) = 0.003\ \text{aJ}$$

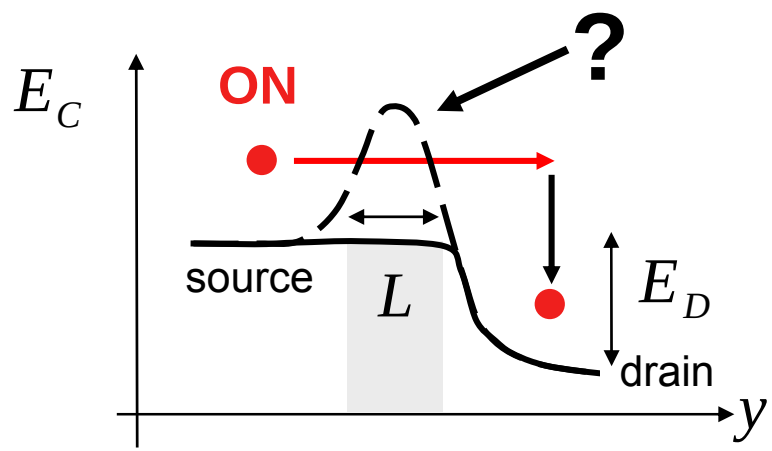


(minimum energy dissipation per logic transition)

minimum switching energy (iii)

Does it take additional energy to move the gate-controlled barrier up and down?





(k_BT ln(2) ***is*** the min energy dissipation per logic transition)

minimum device size

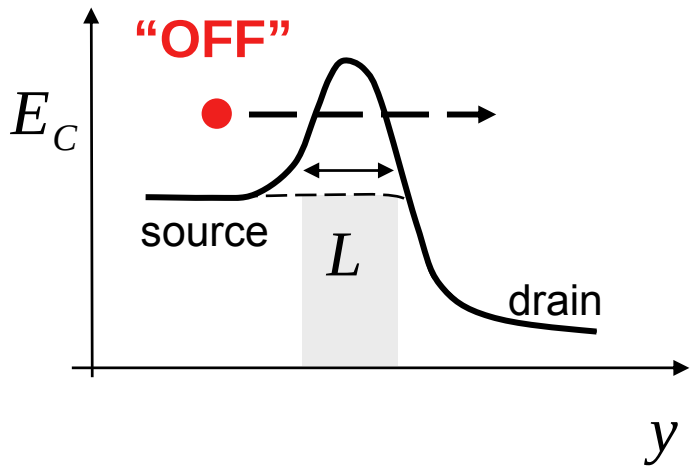
To distinguish **off** from **on**, the probability that an electron tunnels through the barrier must be less than a 50:50.

$$P(\text{WKB}) = \exp\left(-\frac{2\sqrt{2mE}}{\hbar}L\right) < \frac{1}{2}$$

$$L > \frac{\ln(2)}{2}\frac{\hbar}{\sqrt{2mE}}$$

$$E = k_B T = \frac{E_S(\text{min})}{\ln(2)}$$

$$L_{\text{min}} \approx \frac{\text{h}}{\sqrt{2mE_S(\text{min})}} = 1.5\,\text{nm}\,(300\text{K})$$



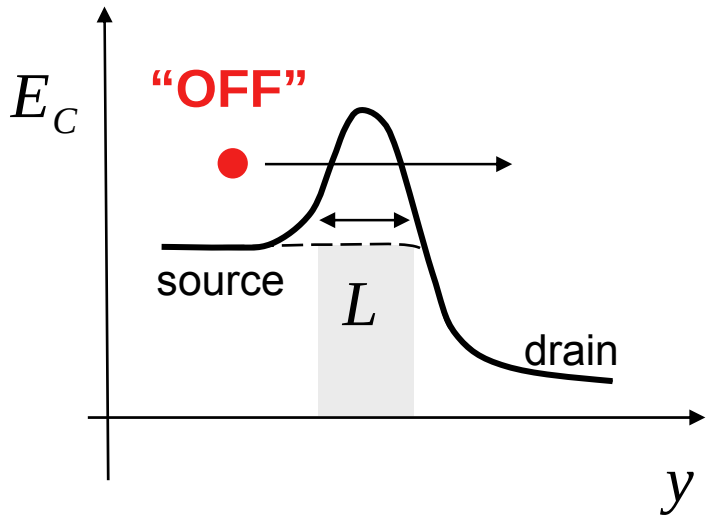
minimum device size (again)

$$\Delta p \Delta x > \mathrm{h}$$

$$\frac{\Delta p^2}{2m} = E_S(\min)$$

$$\Delta x = L$$

$$L_{\min} \approx \frac{\mathrm{h}}{\sqrt{2mE_S(\min)}} = 1.5\,\mathrm{nm}\,(300\mathrm{K})$$



minimum device size (iii)

Note also that the size of a device, *S*, must be larger than the size of its minimum element, *L*. (For a MOSFET, *S* ~ 10-15*L*.)

maximum device density

$$n_D(\max) \approx \frac{1}{\left(10 L_{\min}\right)^2} = 4.7 \times 10^{11}\ \text{cm}^{-2}$$

We will show later that device density is limited by the maximum power density that can be dissipated - not by device size.

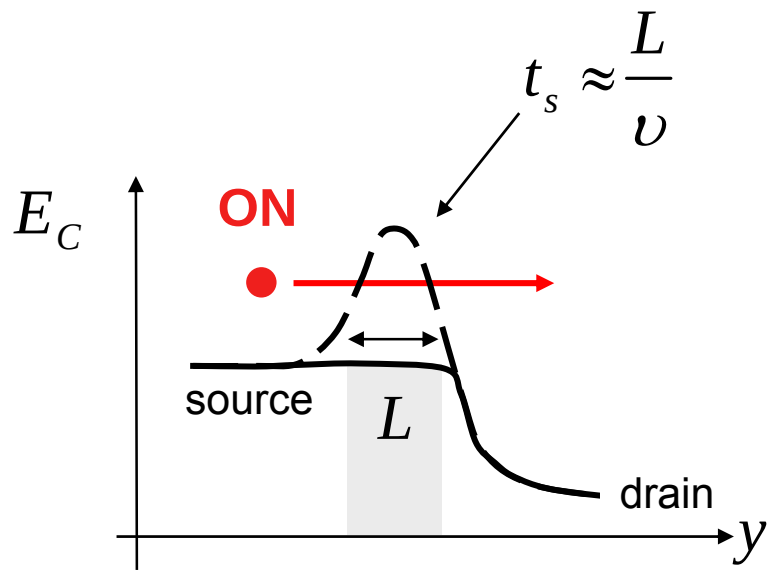
maximum device speed

The minimum device transit time sets the maximum speed.

$$t_S(\min) \approx \frac{L_{\min}}{\upsilon} = \frac{L_{\min}}{\sqrt{2E/m}}$$

$$E = k_B T = \frac{E_S(\min)}{\ln(2)}$$

$$L_{\min} \approx \frac{h}{\sqrt{2mE_S(\min)}}$$



$$t_S(\min) \approx \frac{h}{E_S(\min)} = 0.04 \text{ ps} \quad (300\text{K})$$

$$\Delta E \Delta t = \hbar$$

power dissipation

$$P_D = \frac{\alpha n_D E_S}{t_S}$$

$$\alpha = 1, \quad n_D(\max), \quad E_S(\min), \quad t_S(\min)$$

$$P_D = 3.7 \times 10^4 \ \text{W/cm}^2$$

surface of the sun: 6×10^3 W/cm^2

forced water cooling: < 800 W/cm^2

ITRS: < 100 W/cm^2

power-limited device density

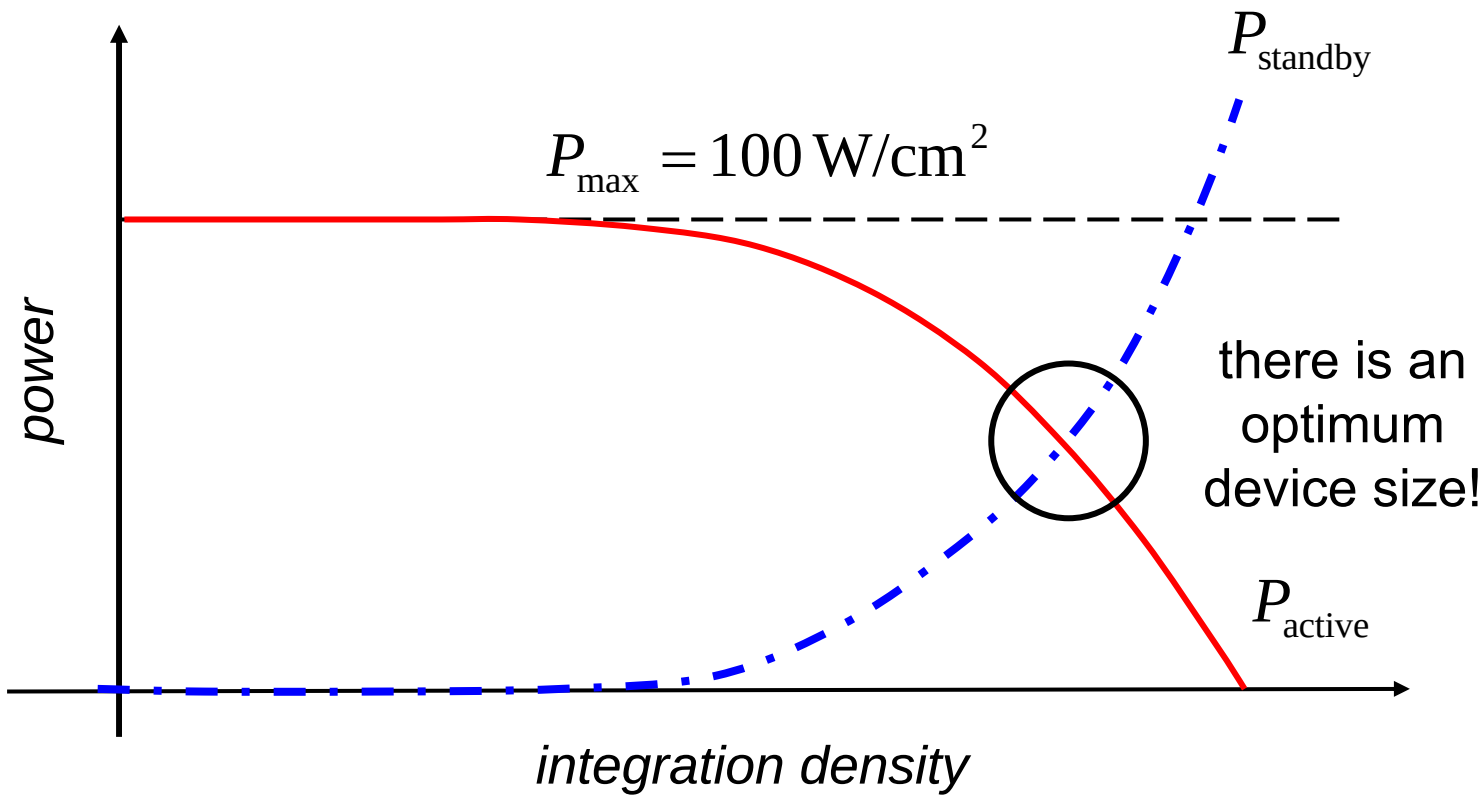
maximum power dissipation per unit area limits density
- not our ability to make devices small.

$$\hat{n}_D(\max) = \frac{P_{\max} t_S(\min)}{\alpha E_S(\min)}$$

for $P_{\max} = 100\,\mathrm{W/cm^2}$ and $\alpha = 1$

$$\hat{n}_D(\max) \sim 1.5 \times 10^9 \quad \text{devices/cm}^2$$

power-constrained design



(Dave Frank, IBM)

power-delay product

$$P_D \tau = \frac{E_S}{t_S} t_S = E_S$$

$$P_D \tau\big|_{\min} = E_S\big|_{\min} = k_B T \ln 2$$

But….this metric does not capture the fact that we usually want to runs circuits fast and at low power.

energy-delay product

$$E\tau\big|_{\min} = E_S t_S\big|_{\min} = E_S(\min)\frac{\mathsf{h}}{E_S(\min)} = \mathsf{h}$$

$$\left(\Delta E \Delta t > \mathsf{h}\right)$$

summary

1) Switching energy: $E_S > k_B T \ln(2)$

2) Switching delay: $t_S > \mathrm{h}/E_S(\min)$

3) Device size: $L_{\min} > \mathrm{h}/\sqrt{2mE_S(\min)}$

4) Device density: $\hat{n}_D < P_{\max} t_S(\min)/E_S(\min)$

5) Energy-delay: $E\tau > \mathrm{h}$

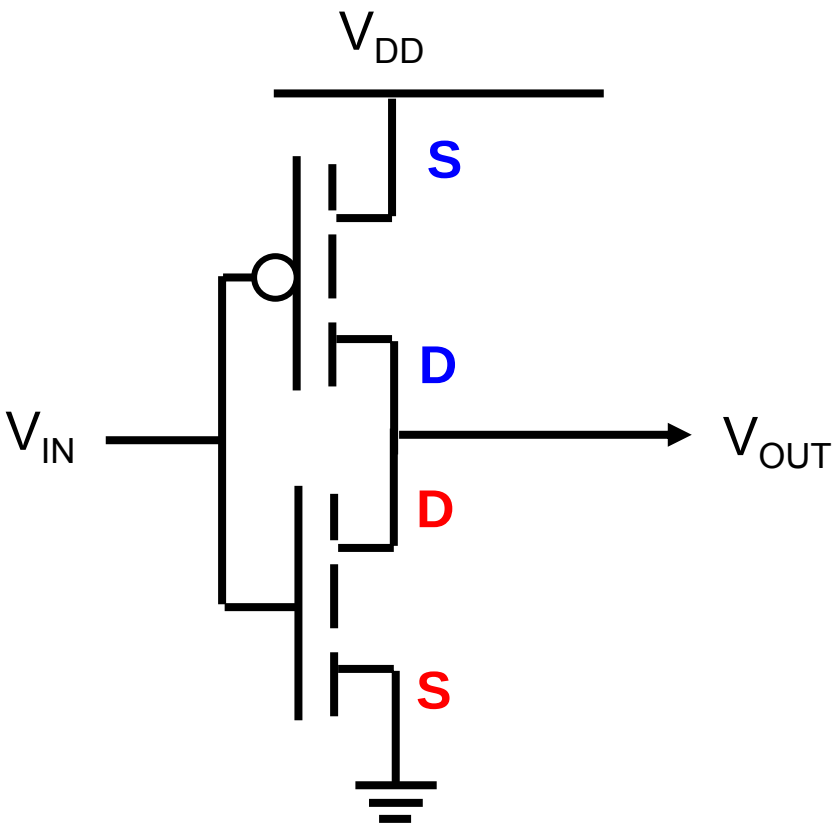
comparison to 65nm CMOS

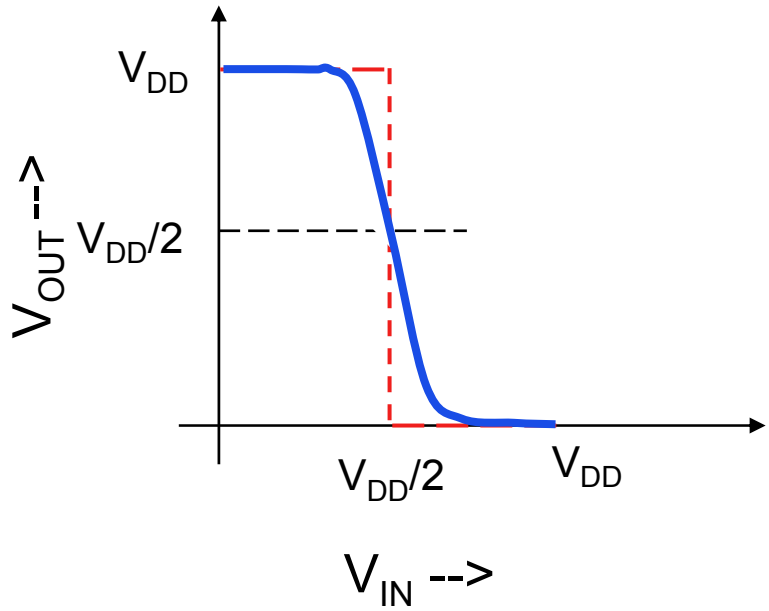
parameter	65nm ITRS	Limit	65nm /Limit
E_s (aJ)	23	0.003	8000
τ (fs)	640	40	16
L (nm)	25 nm	1.5 nm	17
$\hat{n}_S$ (cm^{-2})	0.8 x 10^{9}	1.5 x 10^{9}	0.5
$E\tau$ (J-s)	1.5 x 10^{-29}	1.1 x 10^{-34}	136,000

outline

1) CMOS Metrics
2) MOSFET limits
3) **Circuit limits**
4) System limits

minimum V_{DD}





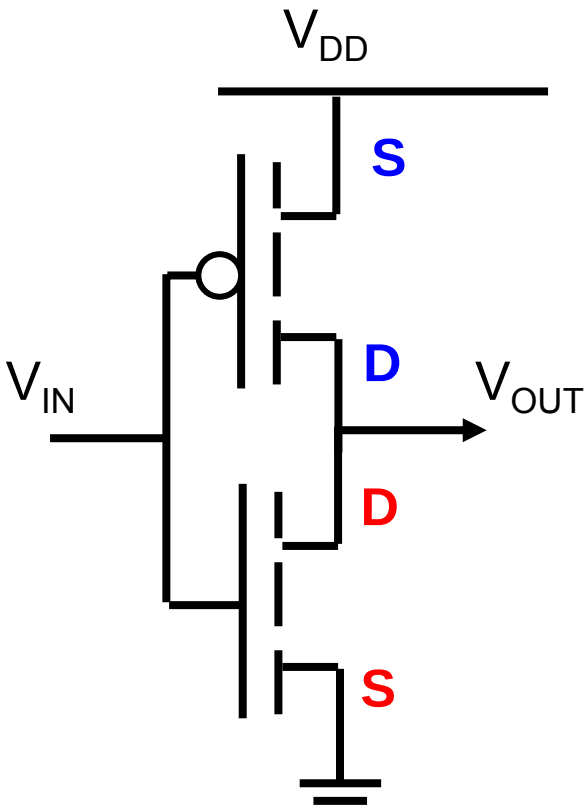
Question: What is the smallest V_{DD} for a CMOS inverter?

Answer: The smallest V_{DD} *that gives a gain > 1*

reference

J.D. Meindl and J.A. Davis, "*The* Fundamental Limit on Binary Switchin Energy for Terascale Integration, *IEEE J. Solid-State Circuits*, **35**, pp. 1515-1516, 2000.

minimum V_{DD}



assume subthreshold operation:

$$I_D = Ke^{q(V_{GS}-V_T)/mk_BT}\left(1-e^{-qV_{DS}/k_BT}\right)$$

$$I_{DN} = K_N e^{q(V_{in}-V_{TN})/mk_BT}\left(1-e^{-qV_{out}/k_BT}\right)$$

$$I_{DP} = K_P e^{q(V_{DD}-V_{in}+V_{TP})/mk_BT}\left(1-e^{q(V_{out}-V_{DD})/k_BT}\right)$$

$$I_{DN} = I_{DP} \rightarrow V_{out}\left(V_{in}\right)$$

minimum V_{DD} (ii)

$$K_N e^{q(V_{in}-V_{TN})/mk_BT}\left(1-e^{-qV_{out}/k_BT}\right)=K_P e^{q(V_{DD}-V_{in}+V_{TP})/mk_BT}\left(1-e^{q(V_{out}-V_{DD})/k_BT}\right)$$

$$\left.\begin{aligned} K_N &= K_P = K \\ V_{TN} &= V_T = -V_{TP} \\ m &= 1 \end{aligned}\right\} \quad \text{solve for} \quad A_V = \frac{dV_{out}}{dV_{in}} = f\left(V_{DD}\right)$$

$$\left|A_V\right| > 1 \Rightarrow V_{DD} > 2\ln(2)\frac{k_BT}{q}$$

$$V_{DD}\big|_{\min} = 2\ln(2)\frac{k_BT}{q}$$

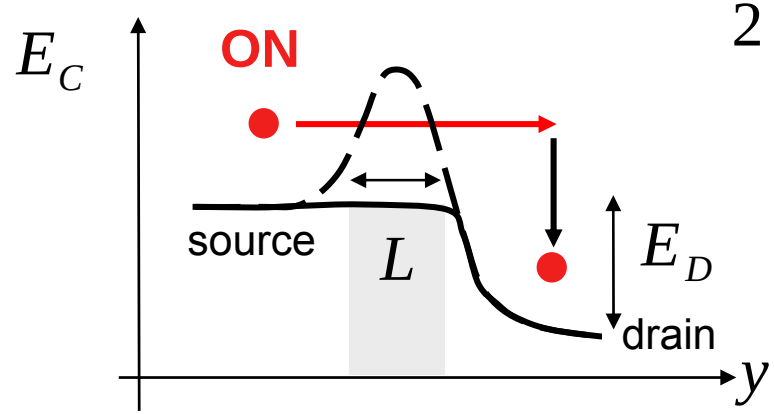
minimum V_{DD} (iii)

$$E_D > E_{D\min} = k_B T \ln(2)$$

$$V_{DD}\big|_{\min} = E_{D\min} / q = \frac{k_b T}{q} \ln(2)$$

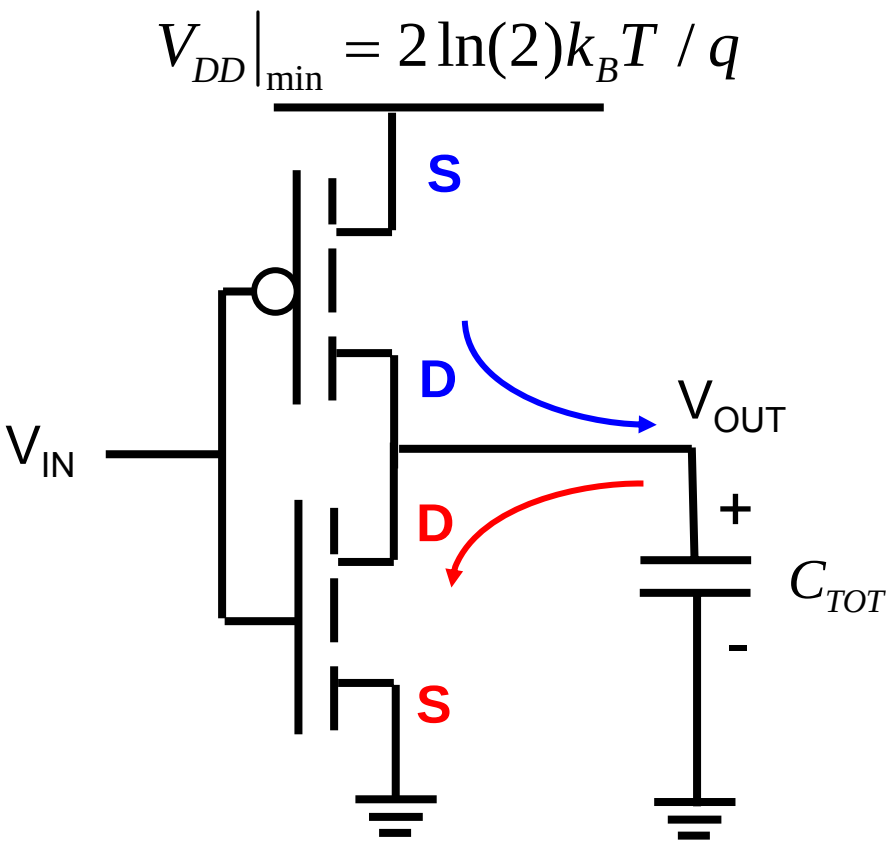
2 MOSFETS in series $\Rightarrow$

$$e^{-E_D / k_B T} < \frac{1}{2}$$



$$V_{DD}\big|_{\min} = 2\ln(2)\frac{k_B T}{q}$$

minimum switching energy



$$E_S = \frac{1}{2} C_{TOT} V_{DD}^2$$

$$E_S = \frac{1}{2} Q V_{DD}$$

$$Q\big|_{\min} = q$$

$$E_S = \frac{1}{2} Q\big|_{\min} V_{DD}\big|_{\min}$$

$$E_S = k_B T \ln(2)$$

outline

1) CMOS Metrics
2) MOSFET limits
3) Circuit limits
4) **System limits**

circuit performance

	device	CPU circuit	increase
delay:	0.64 ps	250 ps	~ 400 ×
switching energy:	23 aJ	6000 aJ	~300 ×
energy-delay:	$\sim 10^{-29}$ J-s	$\sim 10^{-24}$ J-s	~100,000 ×

circuits and fundamental limits

parameter	CPU circuit	Limit	circuit /limit
E_s (aJ)	6000	0.003	2,000,000
τ (fs)	250,000	40	6250
$E\tau$ (J-s)	1.4 x 10^{-24}	1.1 x 10^{-34}	~10^{10}

Since 1960, switching energy has decreased by about 5 orders of magnitude. (J.D. Meindl, Q. Chen, and J.A. Davis, "Limits on Silicon Nanoelectronics for Terascale Integration," *Science*, **239**, pp. 2044-2049, 2001)

chip performance index

$$B = \frac{N}{\tau_S} \quad \text{ops / cm}^2\text{-s}$$

$$CPI = \frac{N}{\tau_S} \times \frac{1}{\left(E_S / t_S\right)} = \frac{N}{E_S} \quad \text{ops / s-cm}^2\text{-W}$$

$CPI: \ 10^{23}$ (today's high-performance logic)

$CPI\big|_{\text{ultimat}}: \ 10^{32}$ (since 1960, the CPI has increased by factor of ~ 10^{14})

J.D. Meindl, "Low Power Microelectronics: Retrospect and Prospect," *Proc. IEEE*, **83**, 619-635, 1995

summary

1) Key device metrics include size, switching energy, and speed.

2) Key system metrics include density, switching energy, speed, and power.

3) Device metrics are 'approaching' fundamental limits.

4) System metrics are a long way from fundamental limits.

outline

1) CMOS Metrics
2) MOSFET limits
3) Circuit limits
4) System limits