Short-Circuit Dissipation of Static CMOS Circuitry and Its Impact on the Design of Buffer Circuits

HARRY J. M. VEENDRICK

Abstract — **This paper gives a detailed discussion of the short-circuit component in the total power dissipation in CMOS circuits, on the basis of an elementary CMOS inverter. Design considerations are given for CMOS buffer circuits, based upon the results of the dissipation discussion, to increase circuit performance.**

List of Parameters Used

A	process-determined constant defined in (11)
a	ratio between L_p and L_n [(A9)]
b	ratio between parasitic nodal capacitance and load capacitance
β	gain factor ($\mu A/V^2$) of an MOS transistor
β_n	β of NMOS transistor
β_p	β of PMOS transistor
β_N	β of nMOST and pMOST of the Nth (symmetrical) inverter of a string
$\beta_\square$	β of a transistor with equal channel length and channel width
C_{gN}	input gate capacitance of the Nth inverter of a string
C_L	load capacitance
C_N	total capacitance on node N
C_o	input capacitance of the first inverter of a string
C_{ox}	gate oxide capacitance
$C_{par\,N}$	parasitic capacitance on node N
f	frequency ($=1/T$)
I	short-circuit current
I_{mean}	mean value of the short-circuit current
I_{max}	maximum value of the short-circuit current
L_n	gate length of the nMOST
L_p	gate length of the pMOST
ΔL_n	gate length minus effective channel length of the nMOST
ΔL_p	gate length minus effective channel length of the pMOST
N	number of inverters
P	total power dissipation
P_1	dynamic power dissipation
P_2	short-circuit power dissipation
T	period-time of a signal ($=1/f$)
t	time
τ	rise or fall time of a signal[1]
τ_f	fall time of a signal[1]
τ_i	rise or fall time of an input signal[1]
τ_o	rise or fall time of an output signal[1]
τ_r	rise time of a signal
V_{dd}	supply voltage
V_{in}	input voltage
V_{out}	output voltage
V_T	threshold voltage
V_{T_n}	threshold voltage of nMOST
V_{T_p}	threshold voltage of pMOST
W_{nN}	channel width of the nMOST of the Nth inverter
W_{pN}	channel width of the pMOST of the Nth inverter

Introduction

During the last five years CMOS technology has become one of the most dominant technologies for VLSI circuits.

The most important reason for this is its low static power dissipation, due to the absence of dc currents during periods when no signal transients occur. However, during an edge of an input signal there will always be a short-circuit current flowing from supply to ground in static CMOS circuits. So far only limited analyses and discussions have appeared in the literature on this power component of static CMOS circuits [1].

In integrated circuits it is always necessary to drive large capacitances (bus lines, "off-chip" circuitry, etc.), often at high clock frequencies. Such driving circuits (buffers) will take a relatively large part of the total power consumption of the chip. It is clear that optimization of such circuits requires a different approach as compared to optimization of CMOS logic [2]. These buffer circuits need extra attention to obtain minimum power dissipation. Therefore, a detailed discussion on power dissipation of a basic CMOS inverter will be given first.

Manuscript received October 19, 1983; revised December 28, 1983.
The author is with Philips Research Laboratories, 5600 JA Eindhoven, The Netherlands.

[1]Although the rise and fall times are commonly defined to be the time between the 10 and 90 percent level of the signal extremes, in this paper these parameters are defined as the total duration of a linearized edge.

0018-9200/84/0800-0468$01.00

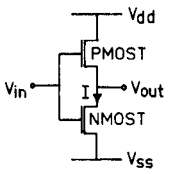


Fig. 1. Basic CMOS inverter.

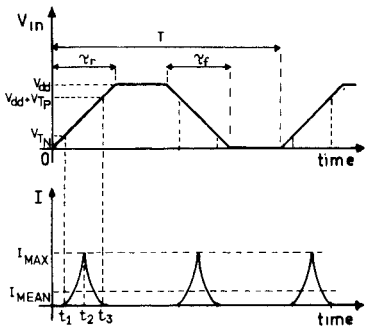


Fig. 2. Current behavior of an inverter without load.

Dissipation of a Basic CMOS Inverter

A static CMOS inverter does not dissipate power during the absence of transients on the input: when the input (Fig. 1) is at high level (V_{dd}), only the NMOS transistor conducts, and when the input is at low level, only the pMOST will conduct. However, during a transient on the input, there will be a time period in which both the nMOST and pMOST will conduct, causing a short-circuit (I) to flow from supply to ground, as shown in Fig. 2 for an inverter without load. This current flows as long as the input voltage (V_{in}) is higher than a threshold voltage (V_{T_n}) above V_{ss} and lower than a threshold ($|V_{T_p}|$) below V_{dd}.

If we load the inverter output of Fig. 1 with a capacitance C_L, then the dissipation of the circuit consists of two components:

$$\text{dynamic dissipation: } P_1 = C_L \cdot V^2 \cdot f \text{ and} \quad (1)$$

$$\text{short-circuit dissipation: } P_2 = I_{\text{mean}} \cdot V. \quad (2)$$

Clearly, the dynamic component P_1 does not depend on the inverter design (apart from contributions due to parasitic output capacitances, such as junction capacitances). The second component P_2, however, strongly depends on the inverter design.

Since there is a difference in the short-circuit dissipation of an inverter without load and that of an inverter with load, we start our discussions on the basis of an inverter with zero load capacitance. For simplicity we assume that the inverter is symmetrical (an asymmetrical inverter is not fundamentally different), which means that

$$\beta_n = \beta_p = \beta \text{ and } V_{T_n} = -V_{T_p} = V_T. \quad (3)$$

During the period (t_1–t_2; Fig. 2) in which the short-circuit current I increases from 0 to $I_{\max}$, the output voltage (V_{out}) will be larger than the input voltage (V_{in}) minus the threshold voltage (V_T) of the nMOST. As a consequence, the NMOS transistor will be in saturation during this period of time.

Using the simple MOS formula, this leads to

$$I = \frac{\beta}{2}(V_{\text{in}} - V_T)^2 \text{ for } 0 \leqslant I \leqslant I_{\max}. \quad (4)$$

This current will reach its maximum value when V_{in} equals half the supply voltage ($V_{\text{in}} = V_{dd}/2$), due to the assumption that the inverter was symmetrical. Another result of this assumption is that the current behavior during the time period t_1–t_3 will be symmetrical with respect to the time t_2.

The mean current during a time T (equal to one period of the input signal) can thus be written as

$$I_{\text{mean}} = 2 * \frac{2}{T}\int_{t_1}^{t_2} I(t)\,dt = \frac{4}{T}\int_{t_1}^{t_2} \frac{\beta}{2}\left(V_{\text{in}}(t) - V_T\right)^2 dt. \quad (5)$$

Assuming equal rise and fall times[2] ($\tau_r = \tau_f = \tau$) of the input signal (symmetrical) and a linear relation between the input voltage (V_{in}) and time (t) during its transients

$$V_{\text{in}}(t) = \frac{V_{dd}}{\tau} \cdot t, \quad (6)$$

it can be derived from Fig. 2 that

$$t_1 = \frac{V_T}{V_{dd}} \cdot \tau \quad \text{and} \quad t_2 = \frac{\tau}{2}. \quad (7)$$

Equations (5), (6), and (7) lead to

$$I_{\text{mean}} = \frac{2\beta}{T}\int_{\tau/2}^{V_T \cdot \tau / V_{dd}} \left(\frac{V_{dd}}{\tau} \cdot t - V_T\right) d\left(\frac{V_{dd}}{\tau} \cdot t - V_T\right), \quad (8)$$

which has the solution

$$I_{\text{mean}} = \frac{1}{12} \cdot \frac{\beta}{V_{dd}} \cdot (V_{dd} - 2V_T)^3 \cdot \frac{\tau}{T}. \quad (9)$$

From (2) and (9) the following expression can be derived for the short-circuit dissipation of a CMOS inverter without load:

$$P_2 = \frac{\beta}{12} \cdot (V_{dd} - 2V_T)^3 \cdot \frac{\tau}{T}. \quad (10)$$

As $1/T = f$, (10) shows that this dissipation component is also proportional to the frequency of switching. Because V_{dd} and V_T are process-determined, the only design parameters that affect P_2 are β and the input rise and fall times (τ) of the inverter.

For an inverter with capacitive load, the β's of the transistors are determined by requirements on output rise and fall times. In this case the short-circuit dissipation depends only on the duration of the input signal edges. As will be shown further on, these edges should not be too long, especially in the case of driver circuits that have a large β value. In the derivation of (10), we started with an

[2] The definitions of rise and fall times used here are different from those in common use (see the list of parameters used).

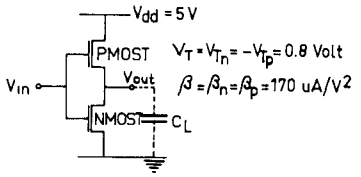


Fig. 3. The inverter used in the example.

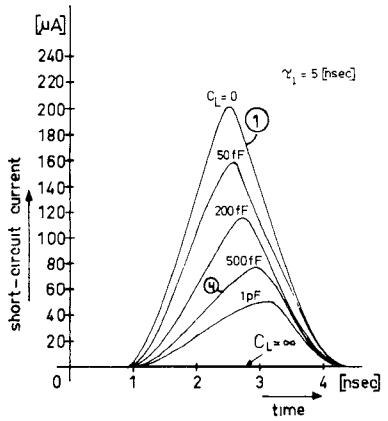


Fig. 4. Short-circuit current as a function of different inverter load capacitances.

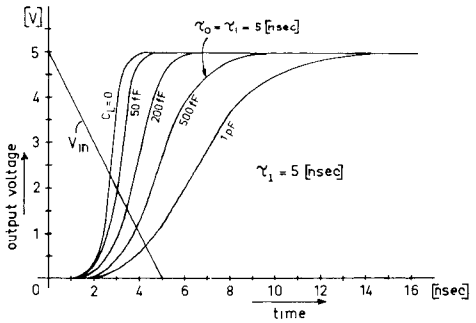


Fig. 5. Inverter output voltage behavior for different inverter load capacitances.

inverter without load. The following example examines what happens when we load the inverter with different capacitances.

Example

The discussions that follow are based upon the inverter whose parameter values are shown in Fig. 3. Its operation was simulated with a circuit analysis program. Some results are presented in Fig. 4. The figure shows the short-circuit current behavior, during a time interval t_1-t_3 (see Fig. 2), as a function of the load capacitance C_L, for input rise and fall times of 5 ns. Curve ① shows the behavior of the inverter without load. At any time this current is the maximum short-circuit current that can occur. This means that all other current characteristics for different load capacitances must be within this curve. Curve ④ shows the short-circuit current behavior of the inverter when it is loaded with a characteristic capacitance C_L of 500 fF. In this case the rise and fall times on the output node are equal to the rise and fall times on the input. Fig. 5 shows

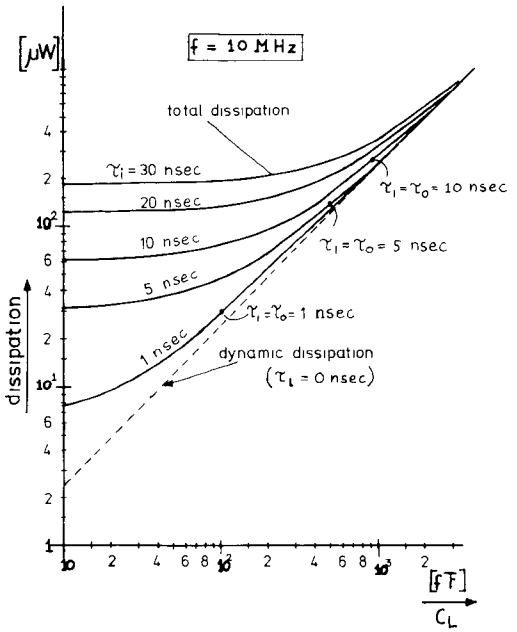


Fig. 6. Inverter dissipation as a function of the inverter load capacitance.

the output transient curves for different values of the load capacitances when the input is switched from a high level to ground with a fall time of 5 ns.

As expressed in (10) for a CMOS inverter with zero load capacitance, it is obvious that the dissipation versus load capacitance characteristic will depend on the rise and fall times of the input signal. Fig. 6 shows this characteristic for different values of the input rise and fall times (τ_i). The dashed line shows the dynamic dissipation ($f = 10$ MHz), while the solid lines show the actual inverter dissipation (dynamic plus short-circuit dissipation). The points where the load capacitance corresponds to equal input and output rise and fall times for the different characteristics are indicated on the figure.

From these characteristics we can conclude that if the operation of the inverter is such that the output signal and input signal have equal rise and fall times, the short-circuit dissipation will be only a fraction (< 20 percent) of the total dissipation. However, if the inverter is more lightly loaded, causing output rise and fall times that are relatively short as compared to the input rise and fall times, then the short-circuit dissipation will increase to the same order of magnitude as the dynamic dissipation. Therefore, to minimize dissipation, an inverter used as part of a buffer should be designed in such a way that the input rise and fall times are less than or about equal to the output rise and fall times in order to guarantee a relatively small short-circuit dissipation.

Fig. 7 shows the linear relationship [according to (10)] between the short-circuit dissipation and the input rise and fall times (τ_i) derived by means of circuit simulations. In this case the inverter of Fig. 3 was loaded with a capacitance of 500 fF. From Fig. 5 it was known that a load capacitance of 500 fF causes 5 ns rise and fall times of the output signal, when the inverter input rise and fall times

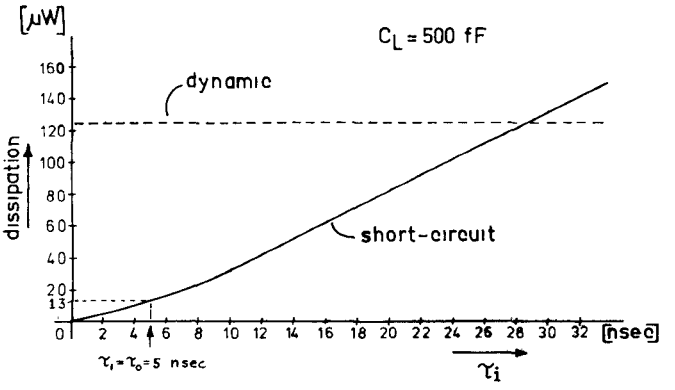


Fig. 7. Inverter dissipation as a function of the input rise and fall times.

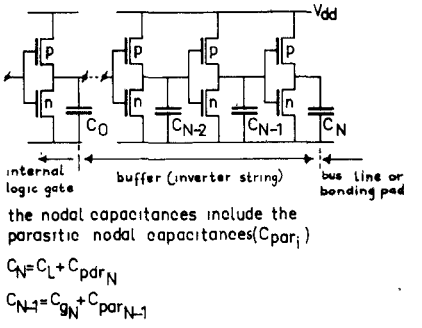


Fig. 8. Inverter string acting as a buffer circuit.

are equal to 5 ns. This point, $\tau_i = \tau_0 = 5$ ns, is indicated on Fig. 7 and the corresponding short-circuit dissipation is only about 10 percent of the dynamic dissipation.

This result of designing a string of inverters in such a way that the input and output rise and fall times of each inverter are equal to obtain minimum dissipation can very well be applied to the design of static CMOS buffers.

Although (10) was derived for an inverter with zero load capacitance and therefore is for the maximum short-circuit dissipation, it can also be applied to inverters designed with $\tau_i = \tau_0$. It has empirically been found that for such designs the short-circuit dissipation is half of the maximum, calculated with (10).

Design Considerations

In integrated circuits it is always necessary to drive large capacitances, like bus lines or "off-chip" circuitry. Moreover, this must often occur at high speed, which will take a relatively large part of the total power dissipation of the chip. Particularly in the case of bus lines, which control a large number of inputs of the different subcircuits on a chip, it is necessary to have short signal rise and fall times [(10)] to minimize dissipation.

Suppose we want to drive such a bus interconnection line or "off-chip" circuitry with a signal coming from an internal node A. Let us assume that the logic gate, having node A as output, is capable of charging a capacitance C_0 in a time t to 95 percent of the supply voltage. From the foregoing results we know that if we use an inverter string as a buffer circuit between node A and the bus line, the rise and fall times on each node of the string should be equal to the required rise and fall times on the bus line (or bonding pad) to be driven.

The problem now is how to design an inverter string (Fig. 8) loaded with a capacitance C_L, with τ ns rise and fall times on each node, driven from an internal logic gate capable of charging and discharging the input capacitance C_0 in the same time. In [3] it was derived that a factor of e between the β's of the successive inverters (tapering factor) was needed to guarantee a minimum propagation delay time for such an inverter string. However, it is well understood that in terms of dissipation and silicon area this will not lead to an optimum design. Design optimization for minimum dissipation and silicon area requires a different approach, as will be shown in the following.

It has been derived in the Appendix that a "minimum dissipation design" of the inverter string is completely determined by the following three equations:

$$\beta_N = \frac{C_N}{\tau_N} * A,$$

where

$$A = \frac{1}{V_{dd} - V_{T_n}} \cdot \left\{ \frac{2V_{T_n}}{V_{dd} - V_{T_n}} + \ln\left[\frac{2(V_{dd} - V_{T_n}) - V}{V} \right] \right\} \tag{11}$$

is constant for a given technology.

$$\frac{\beta_{N-1}}{\beta_N} = \frac{A}{\tau \cdot \beta_\square} \cdot (1+b) \cdot C_{ox} \cdot L_{n_N} \cdot \left\{ L_{n_N}(1+3a^2) - \Delta L_{n_N} - 3a \cdot \Delta L_{p_N} \right\} \tag{12}$$

and

$$\left(\frac{\beta_N}{\beta_{N-1}} \right)^N = \frac{C_N}{C_0} \tag{13}$$

where β_N represents the β of the last inverter stage, β_{N-1}/β_N is the tapering factor at equal input and output rise and fall times and N is the number of inverters of the string. For a practical application the following assumptions are made for the parameters of (11), (12), and (13):

$$\left.\begin{aligned} &V_{T_n} = -V_{T_p} = 1 \text{ V} \\ &V = 0.05 * V_{dd} \text{ [in (11)]} \\ &V_{dd} = 5 \text{ V} \end{aligned}\right\} A \approx 1$$

$$\beta_{n_\square} = 42\ \mu\text{A/V}^2 \qquad \beta_{p_\square} = 14\ \mu\text{A/V}^2$$

$$C_{ox} = 700\ \mu\text{F/m}^2$$

$$L_n = 2.5\ \mu\text{m} \qquad \Delta L_n = 0.5\ \mu\text{m} \qquad \Delta L_p = 1\ \mu\text{m}.$$

Required rise and fall times: $\tau = 5$ ns. Practical values for the constants a and b are $a = 3/2.5$ and $b = 0.1$.

With (12) this leads to

$$\frac{\beta_N}{\beta_{N-1}} = 11.5.$$

This tapering factor, as it is often called, is strongly process dependent; nearly all parameters in (12) are determined by the process. Different CMOS processes may therefore lead to different tapering factors.

If, in a practical situation, $C_o = 100$ fF and C_L (Fig. 8) equals 10 pF and we want 5 ns rise and fall times (τ) on the output of the driver (inverter string), then the design procedure is as follows (with the above assumptions):

according to Fig. 8:

$$C_N = C_L + C_{\text{par}_N} = (1+b)\cdot C_L = 11 \text{ pF}$$

according to (11):

$$\beta_N = \frac{C_N}{\tau_N}\cdot A = \frac{C_N}{\tau}\cdot A = 4.4*10^{-3} \text{ A/V}^2.$$

Thus, the last inverter stage (N) is determined by

$$\beta_{n_N} = \beta_{p_N} = \beta_N = 4.4*10^{-3} \text{ A/V}^2,$$

or (see Appendix)

$$\left(\frac{W_n}{L_n - \Delta L_n}\right)_N = \frac{\beta_N}{\beta_{n\square}} = 105, \qquad \text{so} \left(\frac{W_n}{L_n}\right)_N = \frac{210\ \mu\text{m}}{2.5\ \mu\text{m}}.$$

and

$$\left(\frac{W_p}{L_p - \Delta L_p}\right)_N = \frac{\beta_N}{\beta_{p\square}} = 315, \qquad \text{so} \left(\frac{W_p}{L_p}\right)_N = \frac{630\ \mu\text{m}}{3\ \mu\text{m}}.$$

The $(N-i)$ inverters are now determined by the tapering factor

$$\frac{\beta_N}{\beta_{N-1}} = 11.5.$$

With the given $C_0 = 100$ fF and $C_N = 11$ pF we find from (13) that the number of inverters needed for this example will be equal to $N = 1.83$. (This, of course, has to be rounded off to $N = 2$.)

For this example, therefore, the inverter string should be designed as shown in Fig. 9 to guarantee a very small short-circuit dissipation and a minimum area consumption. The parasitic nodal capacitances are also depicted in Fig. 9.

By means of circuit simulations the mean power dissipation has been calculated at a clock frequency of $f = 1/T =$ 10 MHz. Table I shows the results of a comparison of two tapering factors: a factor of 11.5, which is derived in this paper (process-determined) from optimization of power dissipation and area, and a factor e, which is derived [3] from optimization of the propagation delay.

In this example the most important improvement due to choosing a tapering factor equal to 11.5 instead of a factor e is a much smaller area ($<1/4$) and a reduced parasitic power consumption ($\approx 1/4$). This parasitic power consumption is the total power consumption minus the power dissipation which was actually meant ($C_L V^2 f$).

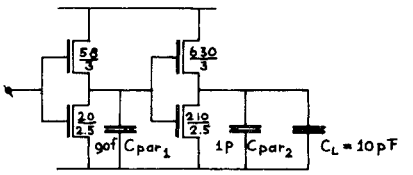


Fig. 9. The designed inverter string for a practical example.

TABLE I
COMPARISON OF THE PERFORMANCES OF TWO INVERTER STRINGS WITH DIFFERENT TAPERING FACTORS

	factor 11.5	factor e
number of inverters	2	5
size of the PMOST in the last inverter	630/3	2660/3
dyn. power dissipation par. " "	2.5 mW 1.4 mW	2.5 mW 5.4 mW
propagation delay	6.5 nsec	5.5 nsec

In our case, the propagation delay has only been increased by 1 ns as compared to a tapering factor equal to e.

In summarizing, we can state that optimization of the power dissipation of CMOS driving circuits, like buffers, will lead to a better overall circuit performance (power, delay, and area) then can be achieved by optimization of the propagation delay.

Summary and Conclusions

In this paper a simple formula is derived for a quick calculation of the maximum short-circuit dissipation of static CMOS circuits. A detailed discussion of this short-circuit dissipation is given based upon the behavior of the inverter when loaded with different capacitances. It was found that if each inverter of a string is designed in such a way that the input and output rise and fall times are equal, the short-circuit dissipation will be much less than the dynamic dissipation (<20 percent). This result has been applied to a practical design of a CMOS driving circuit (buffer), which is commonly built up of a string of inverters. An expression has also been derived for a tapering factor between two successive inverters of such a string to minimize parasitic power dissipation. Finally, it is concluded that optimization in terms of power dissipation leads to a better overall performance (in terms of speed, power, and area) than is possible by minimization of the propagation delay.

Appendix

It can easily be derived [1] what transistor is needed to discharge a capacitive load C_N from the supply voltage V_{dd}

to a voltage V in the time τ_N:

$$\beta_N = \frac{1}{\tau_N} \cdot \frac{C_N}{V_{dd} - V_{T_n}} \cdot \left\{ \frac{2V_{T_n}}{V_{dd} - V_{T_n}} + \ln\left[\frac{2(V_{dd} - V_{T_n}) - V}{V} \right] \right\} \tag{A1}$$

where

$$\frac{1}{V_{dd} - V_{T_n}} \cdot \left\{ \frac{2V_{T_n}}{V_{dd} - V_{T_n}} + \ln\left[\frac{2(V_{dd} - V_{T_n}) - V}{V} \right] \right\} = A \tag{A2}$$

is a constant for a given technology. Thus,

$$\beta_N = \frac{C_N}{\tau_N} \cdot A \tag{A3}$$

and for the $(N-1)$th inverter:

$$\beta_{N-1} = \frac{C_{N-1}}{\tau_{N-1}} \cdot A. \tag{A4}$$

Again, assuming the inverter to be symmetrical:

$$\beta_n = \beta_p = \beta; \qquad V_{T_n} = -V_{T_p} \quad \text{and} \quad \tau_r = \tau_f = \tau \tag{A5}$$

and, because of the difference in mobility of holes and electrons:

$$\frac{W_p - \Delta W_p}{L_p - \Delta L_p} = 3 * \frac{W_n - \Delta W_n}{L_n - \Delta L_n}. \tag{A6}$$

As $W_n \gg \Delta W_n$ and $W_p \gg \Delta W_p$, (A6) reduces to

$$\frac{W_p}{L_p - \Delta L_p} = 3 * \frac{W_n}{L_n - \Delta L_n}. \tag{A7}$$

With

$$W_n = (L_n - \Delta L_n) \frac{\beta_N}{\beta_{n_\square}} \tag{A8}$$

and given a linear relation between L_n and L_p

$$L_p = a \cdot L_n \tag{A9}$$

we find

$$W_p \cdot L_p = W_n \cdot L_n \cdot 3a \cdot \frac{a \cdot L_n - \Delta L_p}{L_n - \Delta L_n}. \tag{A10}$$

From Fig. 8 it is known that

$$C_{N-1} = C_{g_N} + C_{\mathrm{par}_{N-1}}. \tag{A11}$$

In a practical design the parasitic capacitance $C_{\mathrm{par}N-1}$ of node $N-1$ will be proportional to its load capacitance C_{gN}, so that

$$C_{\mathrm{par}_{N-1}} = b \cdot C_{g_N} \quad \text{and} \quad C_{\mathrm{par}_N} = b \cdot C_L. \tag{A12}$$

From (21) and (22) we derive

$$C_{N-1} = (1+b) \cdot (W_{p_N} \cdot L_{p_N} + W_{n_N} \cdot L_{n_N}) \cdot C_{\mathrm{ox}}. \tag{A13}$$

This, combined with (A10) yields

$$C_{N-1} = (W_{n_N} \cdot L_{n_N}) * \left\{ 1 + a \cdot \frac{(3a \cdot L_{n_N} - 3\Delta L_{p_N})}{L_{n_N} - \Delta L_{n_N}} \right\} \cdot (1+b) \cdot C_{\mathrm{ox}}. \tag{A14}$$

Equations (A4), (A5), and (A14) result in

$$\beta_{N-1} = \frac{A}{\tau} \cdot (1+b) \cdot C_{\mathrm{ox}} \cdot \frac{W_{n_N} \cdot L_{n_N}}{L_{n_N} - \Delta L_{n_N}} \cdot \left(L_{n_N} - \Delta L_{n_N} + 3a^2 \cdot L_{n_N} - 3a \cdot \Delta L_{p_N} \right). \tag{A15}$$

Finally, from (18) and (25) we derive

$$\frac{\beta_{N-1}}{\beta_N} = \frac{A}{\tau \cdot \beta_{n_\square}} * (1+b) \cdot C_{\mathrm{ox}} \cdot L_{n_N} \cdot \left\{ L_{n_N}(1+3a^2) - \Delta L_{n_N} - 3a \cdot \Delta L_{p_N} \right\}. \tag{A16}$$

With equations (A1) and (A16) the inverter string is completely determined.

The number of inverters (N), which depends on the ratio in (26), is now determined by

$$\left(\frac{\beta_N}{\beta_{N-1}} \right)^N = \frac{C_N}{C_0} \tag{A17}$$

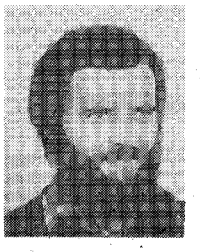
where C_0 and C_N represent the input capacitance and the output load capacitance of the inverter string, respectively (Fig. 8).

Acknowledgment

The author wishes to acknowledge the helpful suggestions and contributions made by A. T. van Zanten and C. Hartgring.

References

[1] M. I. Elmasry, "Digital MOS integrated circuits: A tutorial," *Digital MOS Integrated Circuits*. New York: IEEE Press pp. 4–27.
[2] A. Kanuma, "CMOS circuit optimization," *Solid-State Electron.*, vol. 26, no. 1, pp: 47–58, 1983.
[3] C. Mead and L. Conway, *Introduction to VLSI Systems*. New York: pp. 12–15.



Harry J. M. Veendrick was born in Hummelo en Keppel, The Netherlands, on March 8, 1950. He received the Ing. degree from the Zwolle Polytechnical College, Zwolle, The Netherlands, in 1971, and graduated from the Department of Electronic Engineering at the Eindhoven University of Technology, Eindhoven, The Netherlands, in 1977.

In 1977, he joined Philips Research Laboratories, Eindhoven, The Netherlands, where he is working in a digital integrated-circuit design group on MOS circuit design.