COMPARISON AND ANALYSIS OF DELAY ELEMENTS

Nihar R. Mahapatra[*], *Alwin Tareen*[†], *and Sriram V. Garimella*[‡]

University at Buffalo – The State University of New York
Buffalo, NY 14260, USA

ABSTRACT

This paper comprehensively reviews ten different delay element architectures for use in CMOS VLSI design. They can be categorized into three separate families: transmission gate based, cascaded inverter based, and voltage-controlled based. Six of these delay elements are already in use and we propose four new ones. We compare these delay elements, both analytically and using simulations, in terms of four important parameters: delay, signal integrity, power consumption, and area, and find that they have widely varying characteristics. The results presented in this paper, expressed as parameter ranges, will enable a designer to select the most appropriate delay element that meets delay, signal integrity, power consumption, and area specifications.

1. INTRODUCTION

A delay element is a circuit that produces an output waveform similar to its input waveform, only delayed by a certain amount of time. In this paper, we consider six delay elements that are currently in use and propose four new ones and compare all of them in terms of four relevant parameters: delay, signal integrity, power consumption, and area. These delay elements can be categorized into three families: transmission gate based (Fig.1(a), (b)), cascaded inverter based (Fig.1(c), (d), (e), (f)), and voltage-controlled based [3] (Fig.1(g), (h), (i), (j)). Based on these architectures, we propose four new delay elements, two of which are based on the voltage-controlled principle, and two that utilize a Schmitt trigger in the output stage. These new architectures are compared with the other previously proposed delay elements. Each delay element has its own advantages and disadvantages. Comparison results presented in this paper should prove useful to designers in selecting the best delay element for their application. In the following sections, we describe the functionality of the various delay elements, analyze their four parameters, and present simulation results.

2. TRANSMISSION GATE BASED

2.1. Transmission Gate

A transmission gate is a bidirectional switch consisting of a parallel connection of an NMOS and a PMOS transistor that are controlled by complementary control signals as shown in Fig. 1(a). The NMOS and PMOS transistors pass a logic 0 and 1, respectively. The delay of a transmission gate is effectively determined by the time to charge or discharge a load capacitance C_L at its output through the equivalent resistance R_{eq} of the two transistors connected in parallel. That is, $V_{out}(t) = (1 - e^{-t/R_{eq}C_L})V_{DD}$, so that propagation delay ($V_{out}(t_p) = V_{DD}/2$) is given by [2]:

$$t_p = \ln(2)R_{eq}C_L = \ln(2)\frac{2V_{DD}}{k_n(V_{DD}-V_{Tn})^2+k_p(V_{DD}-|V_{Tp}|)^2}C_L. \quad (1)$$

Here V_{Tn} and V_{Tp} denote NMOS and PMOS transistor threshold voltages, respectively, and k_n and k_p denote gain factors ($\propto$ ratio of width and length of the channel between source and drain ($\frac{W}{L}$)) of the two transistors. For a given fan-out, delay may be increased compared to that of a minimum-size transmission gate by increasing L of the transistors, which linearly increases R_{eq}. There is very little power consumption in a transmission gate since it is not driven by any of the supply rails. Most of the power consumed is that for charging and discharging the output load capacitance from the input. Since the transmission gate is not driven by any of the supply rails, the output load capacitance is charged or discharged by the input. This causes its signal integrity to be not very good. Since $V_{out}(t) = (1 - e^{-t/R_{eq}C_L})V_{DD}$, signal integrity, which is the time for the output to transition between 10% and 90% of V_{DD}, is given by: $t_{r/f} = \ln(9)R_{eq}C_L$. The transmission gate requires relatively less area with just two transistors, although it does require two complementary control signals.

2.2. Transmission Gate Cascaded With Schmitt Trigger

One of the disadvantages of using a transmission gate as a delay element is that the signal integrity of its output waveform is poor. We can overcome this deficiency by placing a Schmitt trigger at the output of the transmission gate. A Schmitt trigger (Fig. 1(b)) is a circuit that generates a fast, clean output signal from a noisy or slowly varying input signal. This is not only useful for noise suppression, but also the steep output minimizes power consumption due to direct-path currents. If the output of the Schmitt trigger is initially low, then the output will go high only when the rising input signal reaches V_{M+}. Similarly, if the output is high, then it will go low only when the falling input signal reaches V_{M-}. The delay of this element can be changed in two ways. The first is by altering the $\frac{W}{L}$ ratio of the transistors of the transmission gate. Decreasing the $\frac{W}{L}$ ratio will increase the gate's delay. The second is by changing the switching thresholds of the Schmitt trigger. If we raise the switching threshold that is active during a rising input transition, V_{M+}, the rise delay will increase. Lowering the switching threshold V_{M-} that is active during a falling input transition will have the same effect on the fall delay. Due to the use of positive feedback, the signal integrity of this delay element is very good. A particularly desirable characteristic is that the signal integrity remains virtually unchanged as the delay value increases. The power consumed by a transmission

[*] This research was supported by startup funds from the University at Buffalo and US National Science Foundation Grant # 0102830. A preliminary version of this paper appears in *Proc. IEEE Computer Society Annual Workshop on VLSI (WVLSI 2000)*, pp. 81-86, Orlando, FL, Apr. 27-28, 2000. N.R. Mahapatra is with the Dept. of Comp. Sc. & Eng.; e-mail: mahapatr@cse.buffalo.edu.

[†] Dept. of Electrical Eng.; e-mail: tareen@eng.buffalo.edu.

[‡] Andiamo Systems, Inc., San Jose, CA 95134, USA; e-mail: sgarimel@andiamo.com.

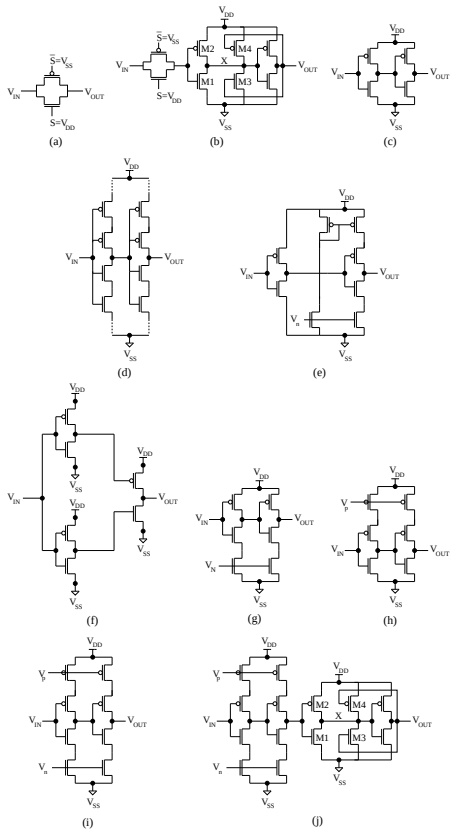

Figure 1: (a) Transmission gate, (b) T-gate with Schmitt trigger, (c) Cascaded inverter, (d) m-transistor cascaded inverters, (e) Current starved cascaded inverters, (f) Staged cascaded inverters, (g) n-voltage controlled, (h) p-voltage controlled, (i) np-voltage controlled, (j) np-voltage controlled with Schmitt trigger.

gate was discussed in the previous section. Since the structure of a Schmitt trigger is very similar to that of a cascaded inverter, the power consumption analysis is similar. This delay element requires eight transistors.

3. CASCADED INVERTER BASED

3.1. Cascaded Inverters

A pair of cascaded inverters can also function as a simple delay element that delays the input signal by an amount equal to the combined propagation delays of the two inverters (see Fig. 1(c)). The propagation delay of an inverter depends upon the time taken to (dis)charge the load capacitance. An exact computation of this delay is nontrivial because of the nonlinear dependence of the (dis)charging current on the output voltage. An approximate expression is derived by using an average value of this current equal to the saturation current of the PMOS (NMOS) transistor given by [2]:

$$I_{av} = \frac{k_p}{2}(V_{GS} - |V_{Tp}|)^2 = \frac{k_p}{2}(V_{DD} - |V_{Tp}|)^2 \approx \frac{k_p}{2}V_{DD}^2. \quad (2)$$

The above holds since $V_{DD} \gg |V_{Tp}|, V_{Tn}$. Based on this I_{av} value, the propagation delay is as follows [2]:

$$t_p = \frac{1}{2}(t_{pLH} + t_{pHL}) = \frac{C_L}{2V_{DD}}\left(\frac{1}{k_p} + \frac{1}{k_n}\right), \quad (3)$$

where t_{pLH} and t_{pHL} denote propagation delays for low to high and high to low output transitions, respectively. The above expression is valid when the input signal makes an abrupt transition from V_{DD} to V_{SS} or vice versa. The effect of a nonzero input rise time $t_r > t_{pHL}$ on propagation delay t_{pHL} is captured by the following equation [2]:

$$t_{pHL(actual)} = \sqrt{t_{pHL(step)}^2 + (t_r/2)^2} \quad (4)$$

A similar expression is valid for t_{pLH}. An approximate expression for signal integrity is computed using the average (dis)charging current expression in Eq. 2 and finding out the time to charge from $0.1V_{DD}$ to $0.9V_{DD}$, or vice versa, as follows:

$$t_{r/f} = \frac{0.9V_{DD}C_L}{|I_{av}|} \approx \frac{0.9C_L}{V_{DD}}\left(\frac{1}{k_p} + \frac{1}{k_n}\right). \quad (5)$$

The signal integrity depends upon the same factors that the propagation delay depends upon. Since the PMOS and NMOS transistors are never on simultaneously in steady-state operation, the static power consumption in an inverter occurs only due to leakage currents and is generally small: $P_{stat} = I_{leak}V_{DD}$. Most of the power is consumed during switching. This dynamic power consists of two components. The major component is due to charging and discharging of the load capacitance. During a low-to-high transition, a certain amount of power is consumed, half of which is dissipated in the PMOS transistor and the other half stored on the load capacitance. During a high-to-low transition, the stored energy in the load capacitance is discharged and dissipated in the NMOS transistor. If an inverter switches on and off f times per second, this capacitive power is given by [2]: $P_{cap} = C_L V_{DD}^2 f$. Hence this power depends quadratically upon V_{DD} and linearly upon C_L and f. The second component of dynamic power arises due to nonzero rise and fall times of the input signal, which results in both NMOS and PMOS transistors being on (a short circuit) briefly. Let the peak current flow during this period be I_{peak}, then the short-circuit power consumption is [2]: $P_{sc} = \frac{t_r+t_f}{2}V_{DD}I_{peak}f$. This means that short-circuit power consumption depends linearly upon the input signal integrity and the supply voltage. It has been shown that in a chain of inverters with equal input rise and fall times, the short-circuit dissipation is less than 20% of the dynamic power dissipation [2]. The total power is therefore given by [2]:

$$\begin{aligned} P_{tot} &= P_{stat} + P_{cap} + P_{sc} \\ &= I_{leak}V_{DD} + C_L V_{DD}^2 f + \frac{t_r+t_f}{2}V_{DD}I_{peak}f. \end{aligned} \quad (6)$$

Of this, the power consumed during capacitive charging (P_{cap}) is the dominant one, followed by P_{sc}, and then P_{stat}. A generalized cascaded inverter requires $4m$ transistors.

3.2. m-Transistor Cascaded Inverters

This delay element is a modified version of the simple cascaded inverter configuration. It has m series-connected NMOS transistors and m series-connected PMOS transistors in its pull-down and pull-up networks, respectively. The gates to all of these transistors are connected to the input. Increasing the fan-in (m) not only increases the effective (dis)charging resistance, but also increases the gate and diffusion capacitances, which contribute to more capacitance at the input and output, respectively. Therefore, the propagation delay (proportional to R and C) depends quadratically on fan-in or m [2]. Further, it increases the delay of the fan-in gate by presenting it a larger load capacitance. Consequently, more delay per unit area may

be obtained by using a generalized inverter with m series-connected transistors than by using a chain of m simple inverters. The increased fan-in (m) has a similar effect on the signal integrity as it does on the delay. That is, the increased effective (dis)charging resistance and increased gate and diffusion capacitances lead to a larger signal integrity value. The increases are roughly the same for both the rise (low-to-high) and fall (high-to-low) times. This delay element consumes slightly more power than the cascaded inverter element. Since there are more transistors in the pull-up and pull-down networks, more energy will be dissipated in these additional devices. An m-transistor cascaded inverter requires 4m transistors.

3.3. Current Starved Cascaded Inverters

A larger fan-in and a clever configuration of control transistors is utilized in the delay element proposed in [1]. The basic architecture is similar to the cascaded inverters. However, two additional PMOS and NMOS devices are added to extend the delay value. The gate voltage V_n is applied to the additional NMOS devices, and they control the maximum current available to the inverter. The gates of the additional PMOS devices are connected to the source of the first additional NMOS device. The delay that can be achieved with this delay element is considerably greater than that achieved with the cascaded inverter or m-transistor cascaded inverter configurations. This is primarily because of its control transistors that limit the amount of current that can (dis)charge the load capacitor CL. The extended (dis)charging time is what gives this element an impressive value of delay. One of the drawbacks of this circuit is that its signal integrity value is very poor, due to the same property that makes its delay very good: the current-limiting capability of its control transistors. An extended (dis)charging time on the load capacitor means that the output signal slopes will be much less steep. Due to their similar architectures, the current starved cascaded inverter has a similar rate of power consumption as the cascaded inverter. The extra fan-in of the current-limiting transistors will increase this parameter slightly. A current starved cascaded inverter requires eight transistors.

3.4. Staged Cascaded Inverters

This delay element consists of an arrangement of three inverters, in two stages. The first stage consists of two inverters, where each inverter output controls a transistor on the second stage's inverter. The key intuition in this design is that the two large transistors in the output stage are never on at the same time, thus eliminating short circuit power dissipation. The delay is obtained by dimensioning the resistances of the two inverters in the first stage. The transition that controls the output edge is always produced by the transistor in series with the resistance and it can be slowed down using large values of R1 and R2. The penalty is in less sharp output edges (although the gain of the output inverter mitigates this effect), and, when both output transistors are off, the input line is susceptible to cross-talk. Both of these effects are greatly reduced by adding another output stage (i.e., two inverters). The power consumption analysis for this delay element is very similar to one for the cascaded inverters. However, this delay element architecture virtually eliminates the component due to short circuit power consumption. Therefore, it only consumes static and capacitive power consumption. A generalized staged cascaded inverter requires 6m transistors.

4. VOLTAGE-CONTROLLED BASED

4.1. n-Voltage Controlled

An n-voltage-controlled delay element, proposed in [3], is shown in Fig. 1(g). It consists of a cascaded inverter pair with an additional series-connected NMOS transistor in the pull-down of each inverter controlled by a global control voltage V_n, varying which changes the delay of this delay element. Note that this delay element may be generalized in a manner similar to the m-transistor cascaded inverter case (Sec. 3.2) by having m-series connected NMOS and PMOS transistors per inverter in addition to the controlled NMOS transistor. There are three ways to change the delay of this element. First is by changing the sizes of the transistors and the second by changing the fan-in m—similar to the cascaded inverter case (Sec. 3.2). The third way is to change the control voltage V_n. Note that during any transition of the input V_{in}, one of the two inverters of the delay element will be discharging through a controlled transistor, and the other charging normally through a PMOS transistor. Therefore, the overall delay is the sum of the normal inverter delay (see Eq. 3) and a controlled inverter delay. The latter delay is inversely proportional to the discharging drain current I_D through the control transistor. Approximating the average discharging current by the saturation current of the controlled NMOS transistor ($I_{av} = \frac{k_n}{2}(V_{GS} - V_{Tn})^2 \approx \frac{k_n}{2}V_{GS}^2 = \frac{k_n}{2}V_n^2$), the propagation delay of this delay element becomes:

$$t_p = \frac{1}{2}(t_{pLH} + t_{pHL}) = \frac{C_L}{2}\left(\frac{1}{k_p V_{DD}} + \frac{V_{DD}}{k_n V_n^2}\right) \tag{7}$$

This shows that t_p is inversely proportional to V_n^2. In [3], the value of the delay assumed is that corresponding to $V_n = 3.5$V, and then V_n is varied after fabrication to fine tune the delay. A delay variation of up to 30% was obtained by changing V_n in 1.2 μm CMOS technology in [3]. The signal integrity is computed in a manner similar to that for the cascaded inverter case. The only difference is that the rise time is similar to the cascaded inverter, but the fall time depends upon V_n. Therefore:

$$t_{r/f} = 0.9C_L\left(\frac{1}{k_p V_{DD}} + \frac{V_{DD}}{k_n V_n^2}\right). \tag{8}$$

The signal integrity depends upon the same factors that the propagation delay depends upon. It will be worse than that for the cascaded inverter case because of the additional resistance and diffusion capacitance of the controlled transistors. The power consumption is computed similar to the cascaded inverter case. It will be slightly more because of the additional diffusion capacitance of the controlled transistors that contributes to the total load capacitance. A generalized voltage-controlled delay element requires $4m + 2$ transistors, two more transistors compared to cascaded inverters.

4.2. p-Voltage Controlled

The voltage-controlled technique can be applied to a delay element in several different manners. This delay element uses a cascaded inverter pair with an additional series-connected PMOS transistor in the pull-up network of each inverter. The gates of these additional transistors are controlled by a control voltage V_p, and this value can be varied to control the amount of delay. Changing the delay of this element can be accomplished by altering the sizes of the transistors, or increasing the fan-in m of the gate (m-transistor cascaded inverter case, Sec. 3.2). Another way is to change the gate voltage of the control transistor, V_p. In this case, the delay can be analyzed in a manner similar to the n-voltage controlled element. During an input transition of V_{in}, one inverter will charge its load capacitance through a controlled PMOS transistor, and the other will discharge regularly through an NMOS transistor. The overall delay is the sum of a controlled inverter delay and a normal inverter delay. The former delay is inversely proportional to the charging drain current I_D

through the control transistor. Approximating the average charging current by the saturation current of the controlled PMOS transistor:

$$I_{av} = \frac{k_p}{2}\left(V_{GS} - V_{T_p}\right)^2 \approx \frac{k_p}{2}V_{GS}^2 = \frac{k_p}{2}V_p^2 \quad (9)$$

The propagation delay becomes:

$$t_p = \frac{1}{2}\left(t_{pLH} + t_{pHL}\right) = \frac{C_L}{2}\left(\frac{V_{DD}}{k_p V_p^2} + \frac{1}{k_n V_{DD}}\right) \quad (10)$$

Hence, t_p is proportional to ${V_p}^2$. The signal integrity analysis is very similar to the cascaded inverter delay element. In the p-voltage controlled element, the rise time depends on the value of V_p, and the fall time is similar to that of a regular inverter.

$$t_{r/f} = 0.9C_L\left(\frac{V_{DD}}{k_p V_p^2} + \frac{1}{k_n V_{DD}}\right) \quad (11)$$

The power consumption is computed similar to the cascaded inverter case. It will be slightly more because of the additional diffusion capacitance of the controlled transistors that contributes to the total load capacitance. Similar to the n-voltage controlled delay element, this one requires 4m + 2 transistors.

4.3. np-Voltage Controlled

This delay element is a combination of the n-voltage controlled and p-voltage controlled configurations. It employs control transistors in both the pull-up and pull-down networks. The delay for this element can be altered by using the methods outlined in the n-voltage controlled and p-voltage controlled sections. One advantage is that the delay can be altered by changing V_p or V_n. The delay analysis is similar to both n-voltage controlled and p-voltage controlled elements. The only difference is that all (dis)charging takes place through a controlled transistor. The propagation delay of this element is:

$$t_p = \frac{1}{2}\left(t_{pLH} + t_{pHL}\right) = \frac{C_L V_{DD}}{2}\left(\frac{1}{k_p V_p^2} + \frac{1}{k_n V_n^2}\right) \quad (12)$$

Note that in this case, t_p is proportional to both ${V_p}^2$ and ${V_n}^2$. The presence of control transistors in both pull-up and pull-down networks influence the rise and fall components of the signal integrity. That is, the rise time depends on V_p and the fall time depends on V_n.

$$t_{r/f} = 0.9C_L V_{DD}\left(\frac{1}{k_p V_p^2} + \frac{1}{k_n V_n^2}\right) \quad (13)$$

The power consumption is computed similar to the cascaded inverter case. It will be slightly more because of the additional diffusion capacitance of the controlled transistors that contributes to the total load capacitance. This delay element requires 4m + 4 transistors.

4.4. np-Voltage Controlled Cascaded with Schmitt Trigger

The poor signal integrity characteristic of the np-voltage controlled delay element can be improved with the addition of a Schmitt trigger to its output. As described in Sec. 2.2, a Schmitt trigger can produce a fast, clean signal from a noisy, slowly varying input. A special advantage of this delay element is that its delay can be changed in four ways. The first way involves altering the $\frac{W}{L}$ ratios of the transistors in the np-voltage controlled gate. Another way is to change the gate voltages applied to the control transistors in the pull-up and pull-down networks. Also, the fan-in m of the gate can be increased, producing a greater delay. The final way is to change the switching thresholds of the Schmitt trigger, and this method is discussed in more detail in Sec. 2.2. This delay element produces the best signal integrity characteristic out of all the delay elements proposed in this chapter. As the delay is altered by changing the $\frac{W}{L}$ ratios of the np-voltage controlled gate, the signal integrity remains virtually unchanged. The main drawback of this circuit is that it consumes a great deal of power, much more than the other delay elements. The power analysis is similar to that of a cascaded inverter. The addition of a Schmitt trigger means that this delay element requires 4m + 10 transistors.

5. COMPARISONS

Experiments were performed for these delay elements using the SpectreS tool from Cadence in 0.18μm technology. The parameters that were taken into account include the delay, the signal integrity, the power dissipation, and the area. We altered the lengths of the delay elements' appropriate transistors, while keeping their widths at a constant minimum value. Then, we extracted the necessary parameters. We began with L = 0.2μm, and increased the transistor length in steps of 0.2μm to L = 1.2μm. A standard cell inverter was used as the fan-in and fan-out for each delay element. Due to space constraints, we summarize only the main results here. We found that overall, the cascaded inverters and n-voltage controlled elements gave a reasonable amount of delay without large power costs. The np-voltage controlled with Schmitt trigger gave the best delay and signal integrity results, but it consumed the most power. The transmission gate based elements proved to be unreliable.

6. CONCLUSIONS

Several principles can be identified in the design of delay elements. of the most reliable ways to increase a circuit's delay is to increase the length L of one or more of its transistors. Another is to create a network of transistors to be placed in series with the n or p network and one of the supply lines. This strategy was successfully utilized in the m-transistor cascaded inverters. The use of a series transistor whose gate voltage can be varied to regulate the current has been shown to create delays. Also, adding a Schmitt trigger to the output of an existing delay element can improve its delay and signal integrity. Overall, the cascaded inverters and n-voltage controlled elements gave a reasonable amount of delay without large power costs. The np-voltage controlled with Schmitt trigger gave the best delay and signal integrity results, but it consumed the most power. The transmission gate based elements proved to be unreliable.

7. REFERENCES

[1] G. Kim, M.-K. Kim, B.-S. Chang, and W. Kim, "A low-voltage, low-power CMOS delay element," *IEEE Journal of Solid State Circuits*, Vol. 31, No. 7, pp. 966-971, Jul. 1996.

[2] J.M. Rabaey, "Digital integrated circuits: A design perspective," *Prentice-Hall Book Company*, 1st edition, ISBN #0-13-178609-1, Upper Saddle River, NJ, 1996.

[3] Yuk-Wah Pang, Wing-yun Sit, Chiu-sing Choy, Cheong-fat Chan and Wai-kuen Cham, "An asynchronous cell library for self-timed system designs," *IEICE Transactions on Information and Systems*, Vol. E80-D, No. 3, pp. 296-305, Mar. 1997.

[4] M.F. Aburdene, J. Zheng and R.J. Kozick, "New recursive VLSI architectures for forward and inverse discrete cosine transform,"*Proceedings of SPIE - The International Society for Optical Engineering*, Vol. 2669, pp. 59-65, 1996.

[5] N.R. Mahapatra, S.V. Garimella, and A. Tareen, "Effi cient techniques based on gate triggering for designing static CMOS ICs with very low glitch power dissipation," *Proc. 2000 IEEE International Symposium on Circuits and Systems*, Vol. II, pp. 537-540, Geneva, Switzerland, May 28-31, 2000.