Lecture 4
Tapered CMOS Inverter stages

Dynamic properties

Additional material w more details

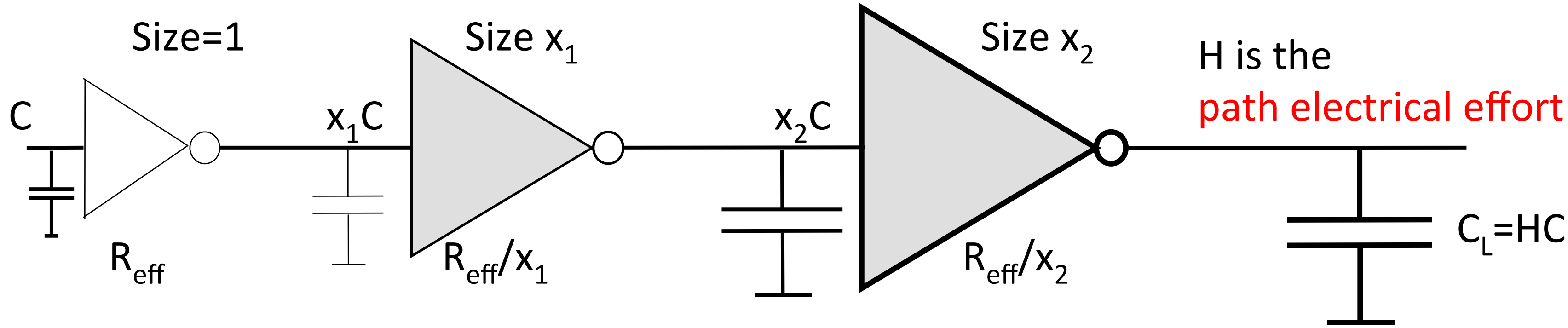
Introduction

- The tapered buffer is analyzed in detail in terms of path electrical effort *H*, and stage electrical effort *h*.
- A stage electrical effort, or fanout, of four is shown to be very close to the optimal solution for minimum delay.
- An H-tree clock distribution network is used to show how branching affects the fanout of a critical timing path. The path fanout, *F*, becomes the product of the path branching, *B*, and the path electrical effort, *H*!
- Conclusion $F=BH$ and stage fanout $f=\sqrt[N]{F}$!

The tapered buffer

Please note that H is the path electrical effort while *h* is the stage electrical effort.

Reference inverter . . .   and two inserted buffer inverters



With two intermediate buffer inverters we obtain a normalized delay relative to tau:

$D=(p_{inv}+h_1)+(p_{inv}+h_2)+(p_{inv}+h_3)$

where we have defined the stage electrical efforts, or fanouts, *h*.

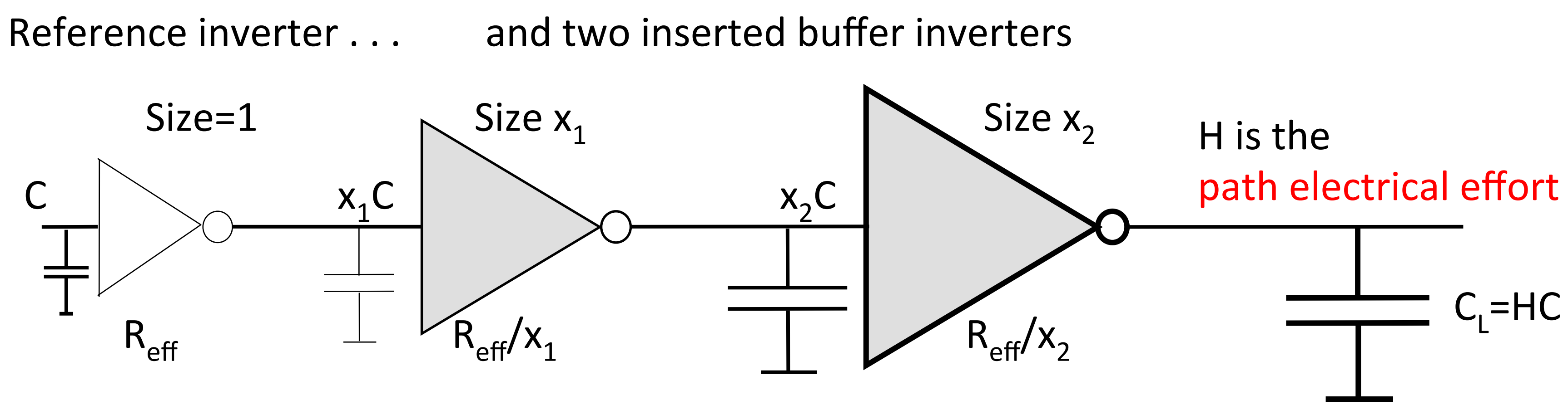
Here $h_1=x_1$, $h_2=x_2/x_1$, and $h_3=x_3/x_2$).

Only h_1 and h_2 are independent variables, the third h_3 becomes $h_3=H/h_1h_2$.

TASK: Show that minimum delay is obtained for $h_1=h_2=h_2=h=\sqrt[3]{H}$ >>> $D=3(p_{inv}+\sqrt[3]{H})$

The tapered buffer

Please note that H is the path electrical effort while *h* is the stage electrical effort.



With two intermediate buffer inverters we obtain a normalized delay relative to tau: $D=(p_{inv}+h_1)+(p_{inv}+h_2)+(p_{inv}+H/h_1h_2)$.

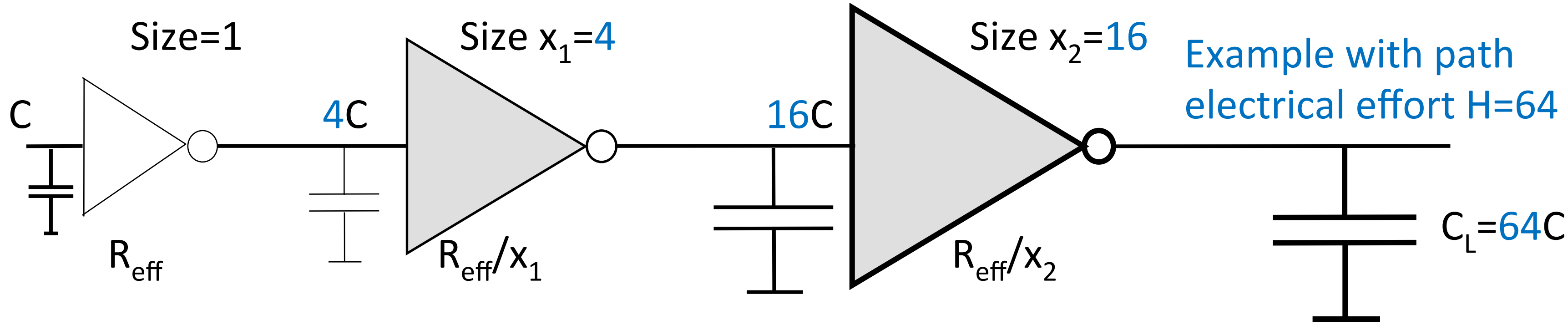
Taking partial derivatives wrt h_1 and h_2 we obtain

$$\frac{d}{dh_1}D(h_1,h_2)=1-\frac{H}{h_1^2h_2}=0,\ \text{and}\ \frac{d}{dh_2}D(h_1,h_2)=1-\frac{H}{h_1h_2^2}=0$$

This yields $h=h_1=h_2$ and $H=h^3 \rightarrow h=\sqrt[3]{H}$

The tapered buffer

Reference inverter . . .   and two inserted buffer inverters

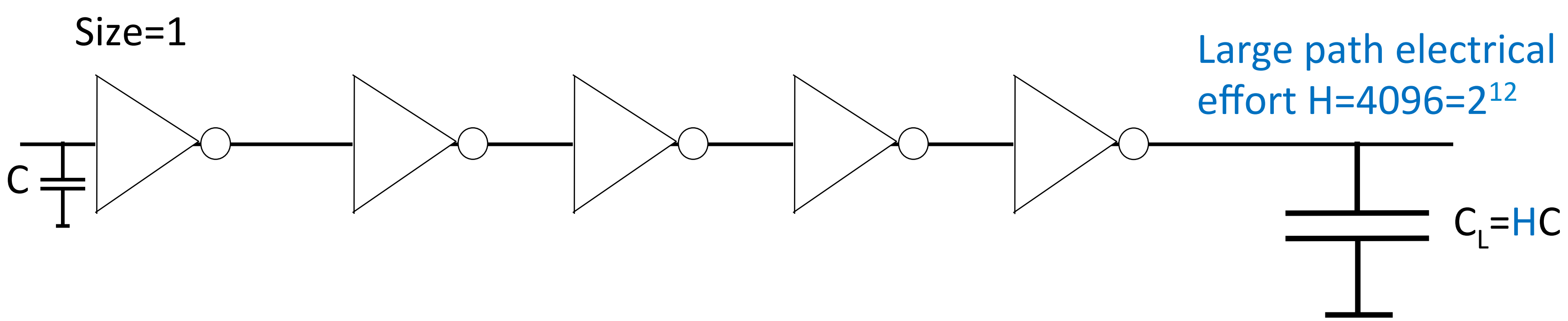


Sharing the load equally between the inverters yields equal stage fanouts $h=\sqrt[3]{64}=4$

The total delay is then equal to 3 FO4 delays, i.e. 15tau=75 ps (assuming p_{inv}=1).

The Tapered Buffer

- What if the path electrical effort, for some reason, is very large, e.g. H=4096.
- How many inverters, *N*, are needed to minimize the delay?



Minimum delay $D=N(p_{inv}+h)$? Determine best N and h!

The Tapered Buffer

Solving this problem we start by accepting that minimum delay occurs when stage electrical efforts, *h*, are equal.

Hence path propagation delay is given by $D = N(p_{inv} + h)$

Furthermore, $h=\sqrt[N]{H}$, i.e. $H= h^N$.

Taking natural logarithms we obtain number of inverters $N = \dfrac{\ln H}{\ln h}$

We rewrite path delay equation as $D = \dfrac{p_{inv} + h}{\ln h} \ln H$

Looking for minimum path delay by taking derivatives of *D* wrt *h*

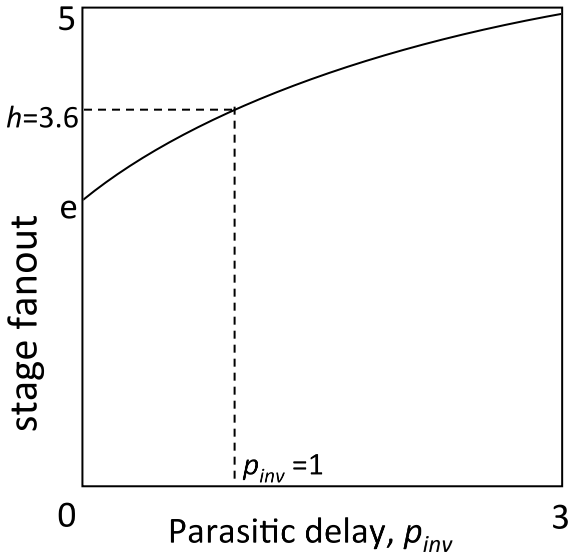
we obtain $\ln H \dfrac{\ln h - \left(p_{inv} + h\right)/h}{\left(\ln h\right)^2} = 0$, i.e. $\ln h = \dfrac{p_{inv} + h}{h}$

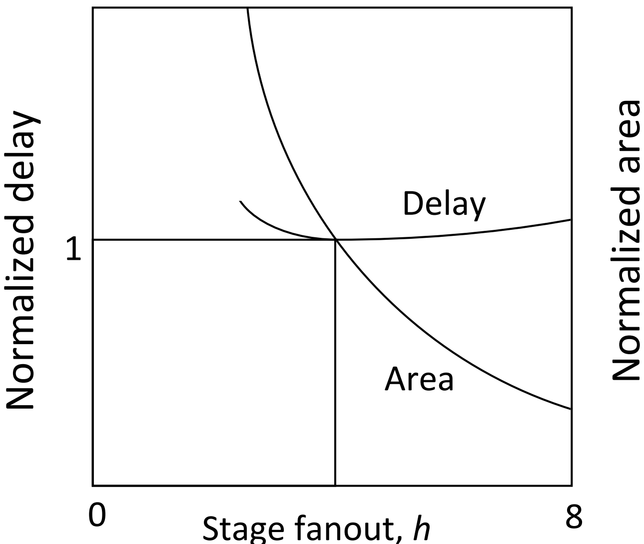
Analytical solution is possible only for $p_{inv}=0$:

h=e=2.72 which gives N = ln H

The Tapered Buffer

- For $p_{inv} \neq 0$ the equation has to be solved numerically





[Hedenstierna & Jeppson 1987]

For typical values of p_{inv} the optimum tapering factor is between 3.6 and 5. Typically a FO4!

Please, note that the propagation delay minimum is rather flat,
while total inverter area on the silicon decreases rapidly when larger stage fanout is used.
Silicon real estate (=cost) can be saved for relatively little loss of speed!

The tapered buffer – area comparison

So for a path electrical effort $H=4096=4^6$, minimum delay is obtained for N=6 inverters with a stage fanout of four!!

Minimum normalized path delay $D=6(p_{inv}+4)=\mathbf{30}$, and $\boldsymbol{t_{pd}}\mathbf{=150\ ps}$

What if we choose N=4? That is $h=8$.

What would be the propagation delay?

Normalized path delay $D=4(p_{inv}+8)=\mathbf{36}$, and $\boldsymbol{t_{pd}}\mathbf{=180\ ps}$.

At the same time the inverter area would be reduced from

A=1+4+16+64+256+1024=**1365 units** to

A=1+8+64+512=**585 units**.

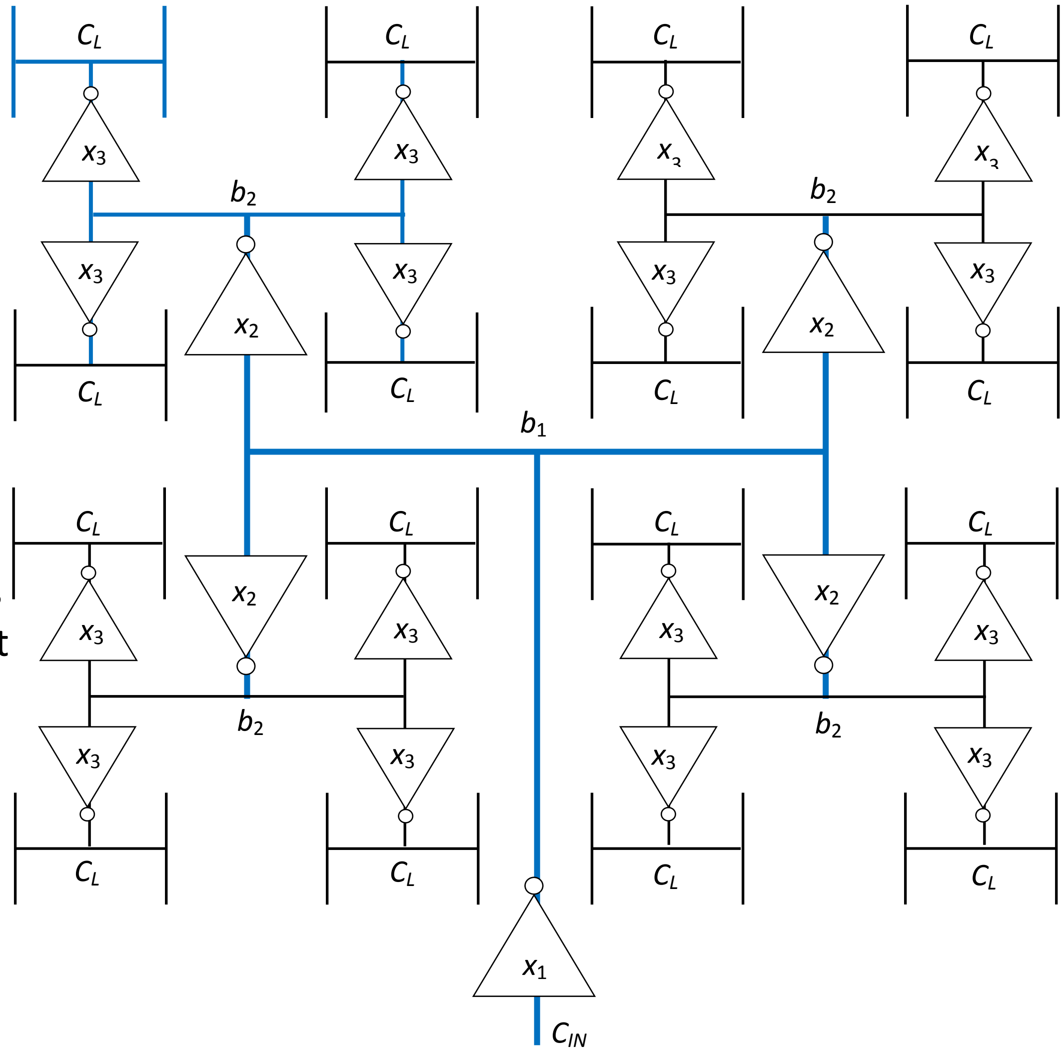
That is a reduction by more than 55%.

H-tree clock distribution

Problem: To distribute a clock signal across the chip area so that it arrives simultaneously to all chip corners, i.e no clock skew, and with sharp clock edges!

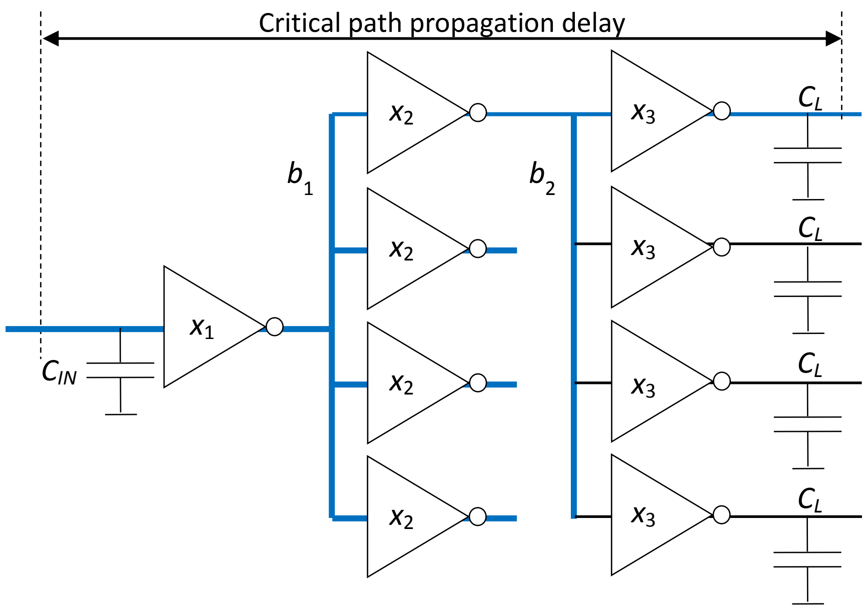
One approach is to use balanced H-trees!

The load capacitances C_L are due to the input capacitances of all clocked register cells.



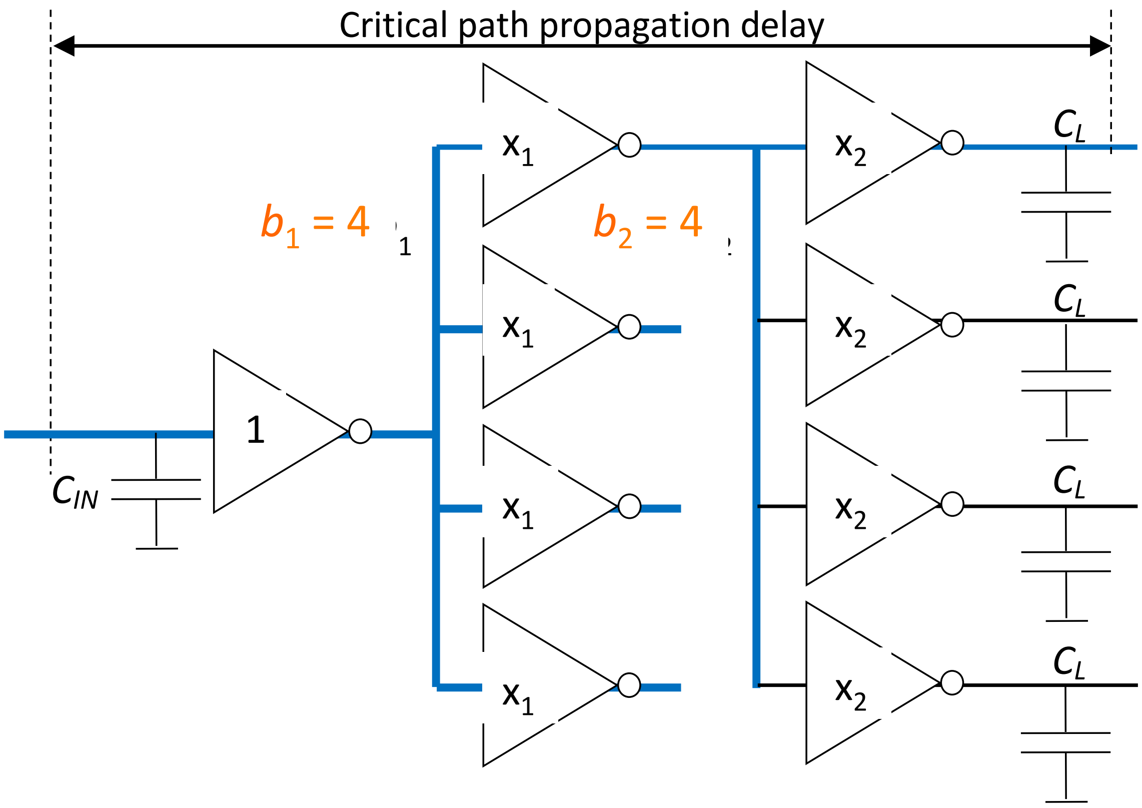
H-tree clock distribution

- What is the electrical effort of the timing paths considering also the branching?
- What sizes to choose for the inverters of the H-tree?



H-tree clock distribution

Note: Here we change the inverter sizes to match the problem in slides 3-5



H-tree clock distribution

As before, assume stage fanouts but now call them f_1, f_2, f_3

Path delay is then given by $D = \left(p_{inv} + f_1\right) + \left(p_{inv} + f_2\right) + \left(p_{inv} + f_3\right)$

Define them as: $f_1 = b_1 x_1,\ f_2 = b_2 \frac{x_2}{x_1},\ f_3 = \frac{C_L}{x_2 C_{IN}} = \frac{H}{x_2}$

Now we have: $f_1 f_2 f_3 = b_1 x_1 b_2 \frac{x_2}{x_1} \frac{H}{x_2} = b_1 b_2 H$

Let us introduce the path branching effort B= $b_1 b_2$!

The path fanout is now $F=B\ H$, and stage fanout $f=\sqrt[3]{F}$.

In an example with B=16 and H=256 we obtain F=4096=16^3

Branching effort definitions

Stage branching effort ***b***:

$$b = (C_{onpath} + C_{offpat})/C_{onpath}$$

Path branching effort ***B:***

$$\boldsymbol{B} = b_1 \times b_2 \times ..... \times b_N$$

More about these concepts next week

Example solution for C_L=256 C_{IN}!

H-tree clock distribution