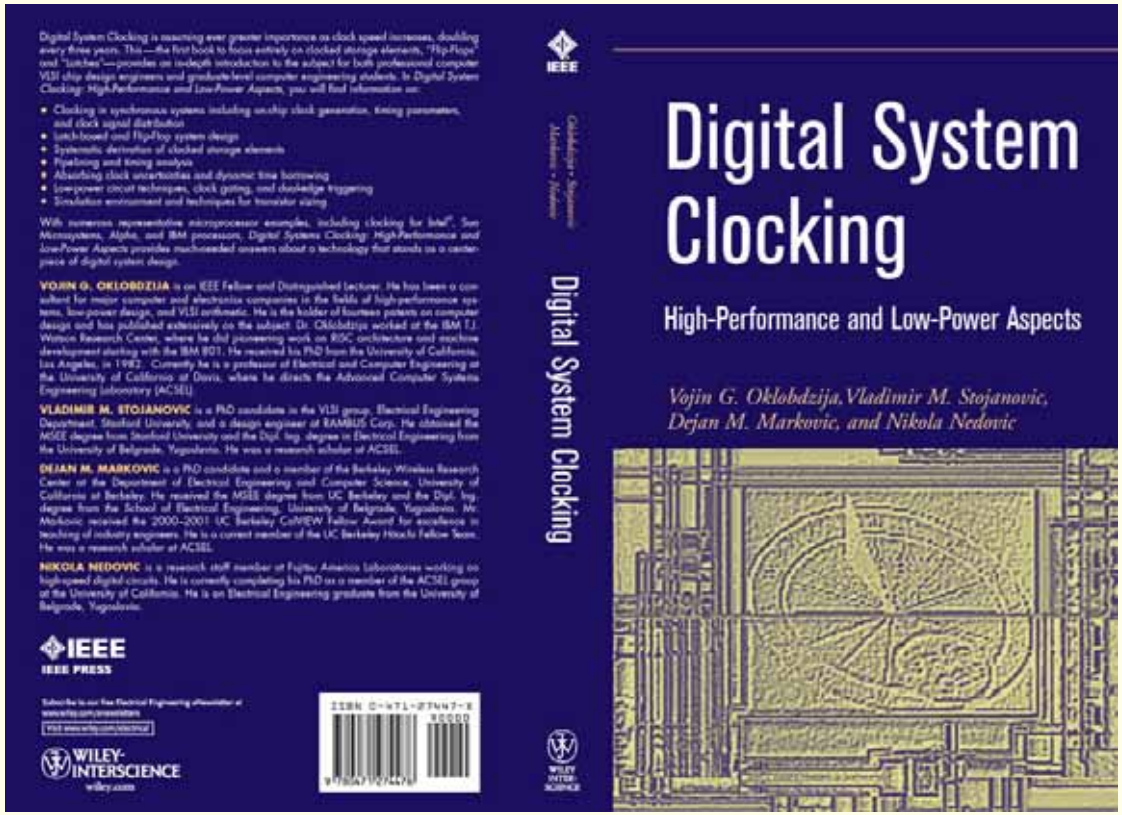
Digital System Clocking:

High-Performance and Low-Power Aspects

Vojin G. Oklobdzija, Vladimir M. Stojanovic, Dejan M. Markovic, Nikola M. Nedovic

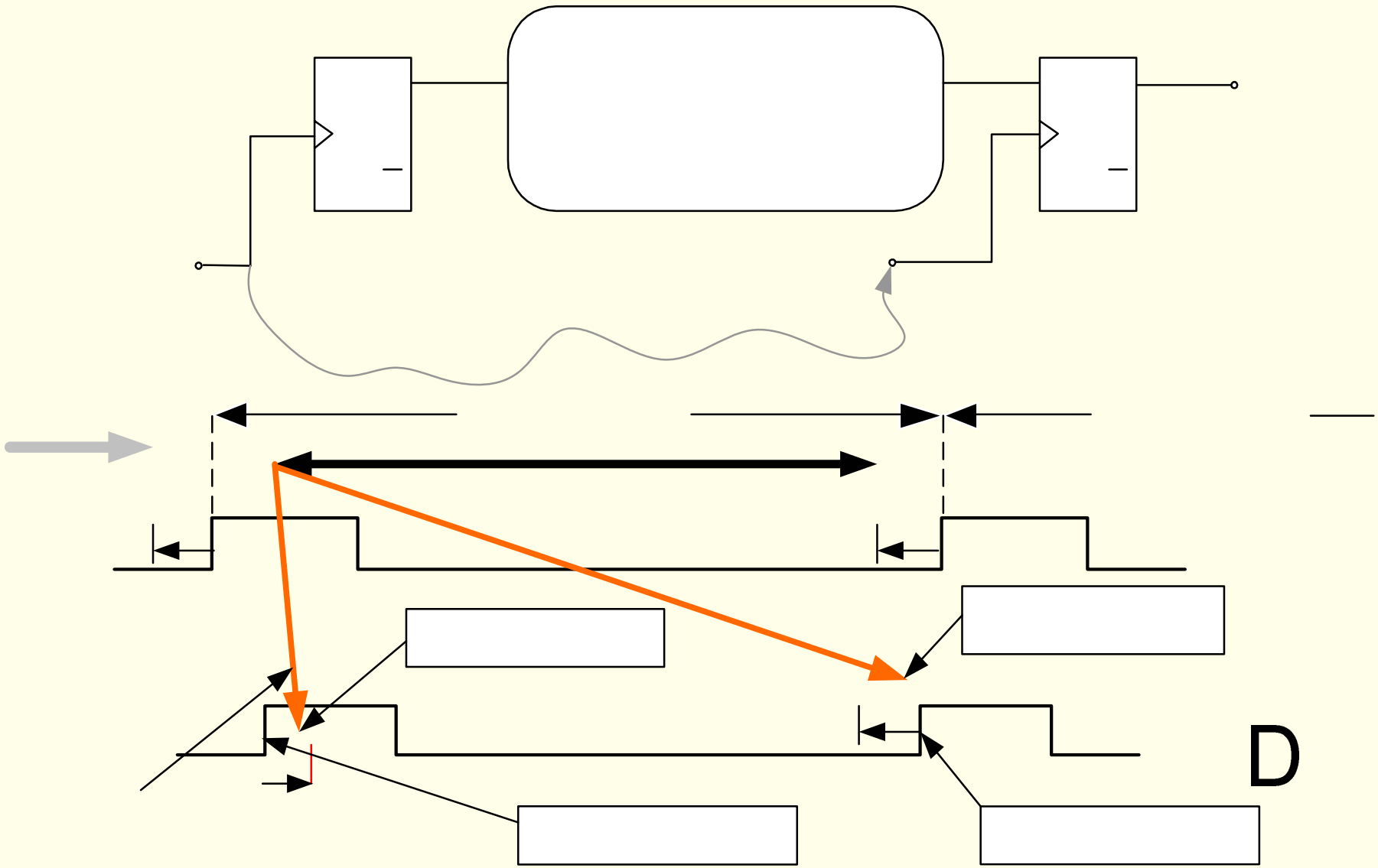
Chapter 4: Pipelining and Timing Analysis



Wiley-Interscience and IEEE Press, January 2003

Timing in a digital system using a single clock and flip-flops

Timing in a digital system using a single clock and flip-flops



Late Data Arrival Analysis
(single clock, FF)

If we set the time reference to *t=0* for the leading edge of the clock

We set max. clock uncertainties to T_L from the nominal time of arrival, T_T (for the trailing edge).

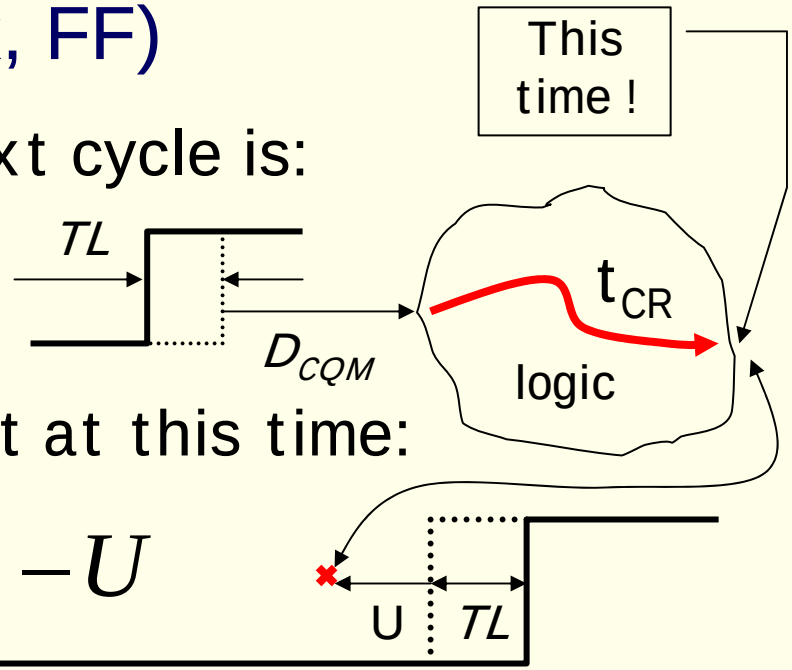
D_{CQm} represents the minimal *Clock-to-Q* (output) delay of the Flip-Flop

D_{Lm} represents minimal delay through the logic (as opposed to index M where D_{CQM} and D_{LM} represent maximal delays).

Late Data Arrival Analysis
(single clock, FF)

The latest data arrival in the next cycle is:

$$t_{DLN} = T_L + D_{CQM} + t_{CR}$$



But it should be there at least at this time:

$$t_{DLN} = P - T_L - U$$

So that:

$$P - T_L - U \geq T_L + D_{CQM} + t_{CR}$$

Giving us:

$$P \geq 2T_L + U + D_{CQM} + t_{CR}$$

$$t_{CR} \leq P - 2T_L + U + D_{CQM}$$

Early Data Arrival Analysis
(single clock, FF)

It is commonly misunderstood that the Flip-Flop provides edge-to-edge timing and is thus easier to use, as compared to the Latch based system, because it does not need to be checked for fast paths in the logic (Hold-time violation).

This is not true, and a simple analysis that follows demonstrates that even with the Flip-Flop design the fast paths can represent a hazard and invalidate the system operation.

Early Data Arrival Analysis
(single clock, FF)

If the clock controlling the Flip-Flop releasing the data is skewed so that it arrives early,

and the clock controlling the Flip-Flop that receives this data arrives late,

a hazard situation exists.

This same hazard situation is present if the data travels through a *fast path* in the logic.

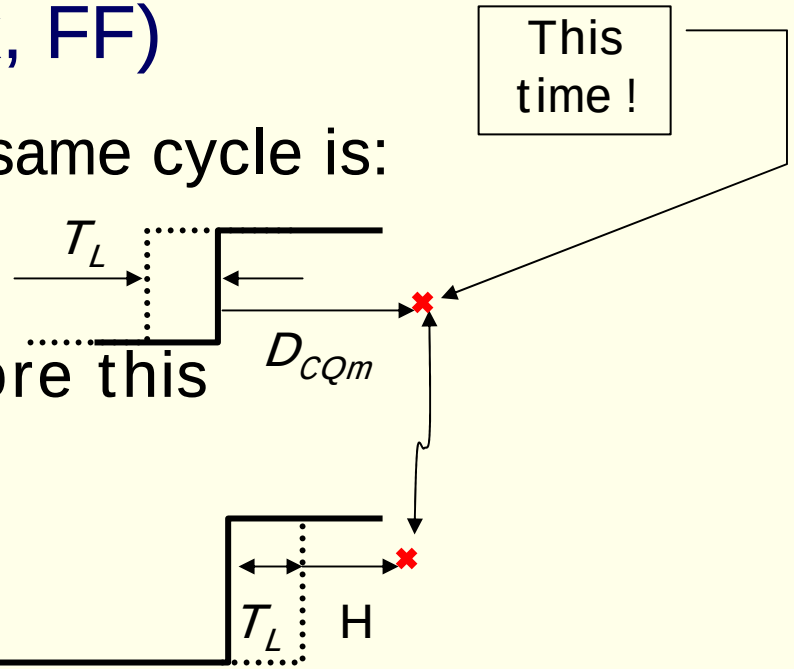
A *fast path* is the path that contains very few logic blocks, or none at all.

This hazard is also referred to as *critical race* (or *race-through*)

Early Data Arrival Analysis
(single clock, FF)

The earlyest data arrival in the same cycle is:

$$t_{DEArr} = -T_L + D_{CQm} + D_{Lm}$$



But it should be there not before this time:

$$t_{DEArrN} = +T_L + H$$

So that:

$$-T_L + D_{CQm} + D_{Lm} > +T_L + H$$

Giving us limits on the fast paths:

$$D_{Lm} > D_{LB} = 2T_L + H - D_{CQm}$$

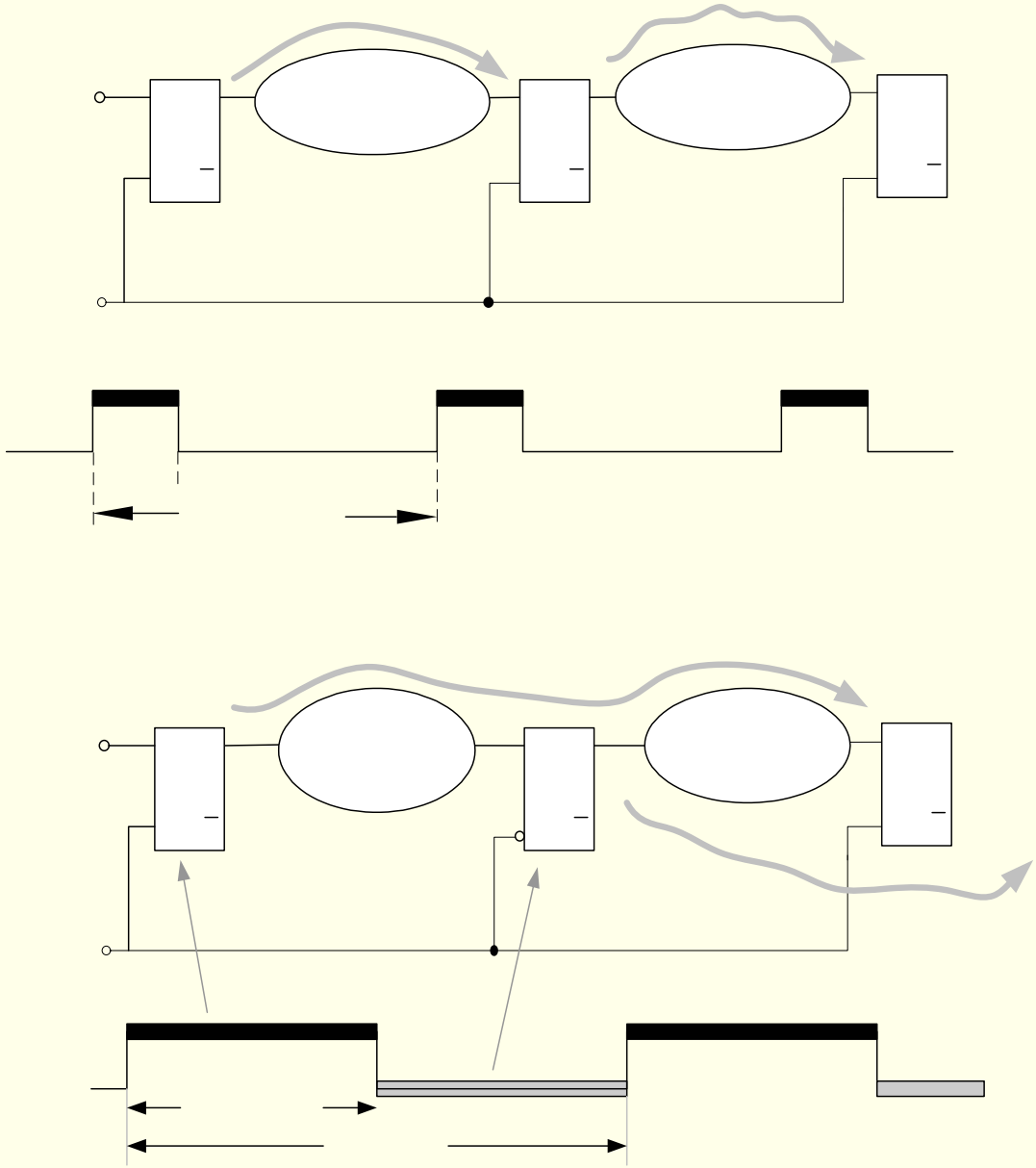
$$W \geq T_L + T_T + t_{CWm}$$

System using a Single Latch

Analysis of a System using a Single Latch

- System using a single Latch is more complex to analyze than Flip-Flop based one.
- Single Latch is transparent while the clock in active and the possibility for the *race-through* exists.
- This analysis is still much simpler than a general analysis of a system using two Latches (Master-Slave Latch based system).
- Use of a single Latch represents a hazard due to the transparency of the Latch, which introduces a possibility of races in the system.
- Therefore, the conditions for the single-latch based system must account for critical race conditions.
- Presence of the CSE delay decreases the “useful time” in the pipeline cycle. Therefore, in spite of the hazards introduced by such design, the additional performance gain may well be worth the risk.

Two ways of using a latch in a single-latch-based system:



Late Data Arrival Analysis

In the case of a Latch, input signal need to arrive at least a Setup Time *U* before the trailing edge.

This edge could arrive earlier. Thus, the latest arrival of data into the latch that assures reliable capture after the period *P* has to be:

$$t_{DLArr} \leq W - T_T - U + P$$

Data captured at the end of the clock period could be a result of two events (whichever later):

a) The data was ready, and clock arrived at the latest possible moment T_L, and the worse case delay of the Latch i.e. D_{CQM} was incurred.

b) The clock was active and data arrived at the last possible moment, which is a setup time *U* and clock skew time T_T before the trailing edge of the clock.

- In both cases the path through the logic was the longest path D_{LM}.

Under the worse scenario data must arrive in time:

$$\max\left\{T_L + D_{CQM}, W - T_T - U + D_{DQM}\right\} + D_{LM} \geq W - T_T - U + P$$

Late Data Arrival Analysis

- This gives a constraint for the clock speed in terms of *P* such as:

$$P \geq \max\left\{T_L + T_T + U + D_{CQM} - W, D_{DQM}\right\} + D_{LM}$$

This inequality breaks down into two inequalities:

$$P \geq D_{LM} + D_{CQM} + T_L + T_T + U - W$$

$$P_m = P \geq D_{LM} + D_{DQM}$$

This shows the minimal bound for *Pm*, which is the time to traverse the loop:

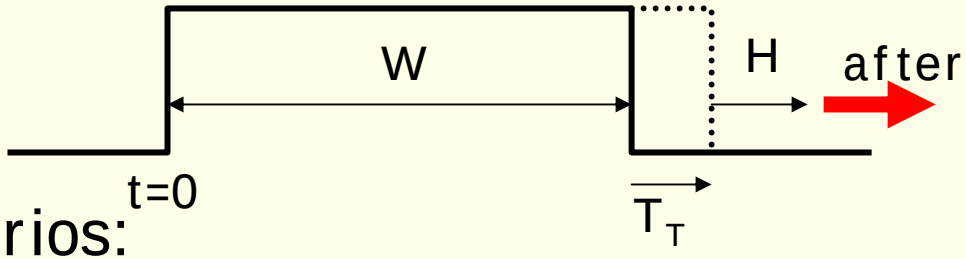
"Starting from the leading edge of a clock pulse, there must be time, under worst case, before the trailing edge of the clock in the next cycle, for a signal to pass through the Latch and the logic block in time to meet the setup time constraint".

The value of *P = Pm* determines the highest frequency of the clock.

Early Signal Arrival Analysis
(Single Latch Based System)

The fastest signal, should arrive at the minimum a hold time after the latest possible arrival of the same clock:

$$t_{DEArrN} \geq W + T_T + H$$

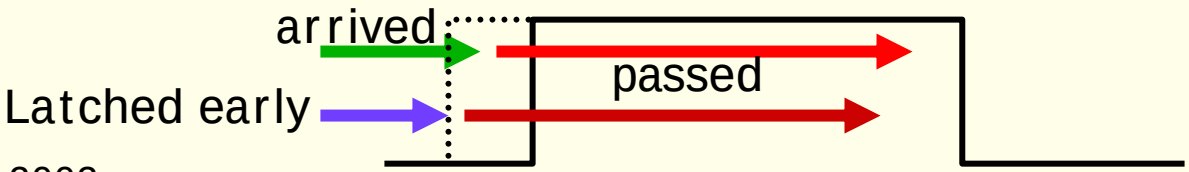


There are two possible scenarios:

(a) signal was latched early and it passed through a fast path in the logic

(b) it arrived early while the Latch was transparent and passed through the fast Latch and fast path in the logic.

$$t_{DEArrN} = \min\left\{t_{CEL} + D_{CQm}, t_{DEArr} + D_{DQm}\right\} + D_{Lm}$$



Early Signal Arrival Analysis
(Single Latch Based System)

The earliest arrival of the clock t_{CEL} happens when the leading edge of the clock is skewed to arrive at $-T_L$. Thus, the condition for preventing race in the system is expressed as:

$$\min\left\{-T_L + D_{CQm}, t_{DEArr} + D_{DQm}\right\} + D_{Lm} \geq W + T_T + H$$

The earliest possible arrival of the clock, plus clock-to-output delay of the Latch has to occur earlier in time than early arrival of the data, thus:

$$-T_L + D_{CQm} + D_{Lm} \geq W + T_T + H$$

which gives us a lower bound on the signal delay in the logic:

$$D_{Lm} > D_{LmB} \geq W + T_T + T_L + H - D_{CQm}$$

Early Signal Arrival Analysis
(Single Latch Based System)

The conditions for reliable operation of a system using a single Latch are:

$$P_m = P \geq D_{LM} + D_{CQM} + T_L + T_T + U - W$$

$$P \geq D_{LM} + D_{DQM}$$

$$D_{Lm} > D_{LmB} \geq W + T_T + T_L + H - D_{CQm}$$

the increase of the clock width *W* may be beneficial for speed, but it increases the minimal bound for the fast paths

Early Signal Arrival Analysis
(Single Latch Based System)

Maximum useful value for *W* is obtained when the period *P* is minimal:

$$W^{opt} = T_L + T_T + U + D_{CQM} - D_{DQM}$$

Substitute the optimal clock width W^{opt} we obtain the values for the maximal speed and minimal signal delay in the logic which has to be maintained in order to satisfy the conditions for optimal single-latch system clocking:

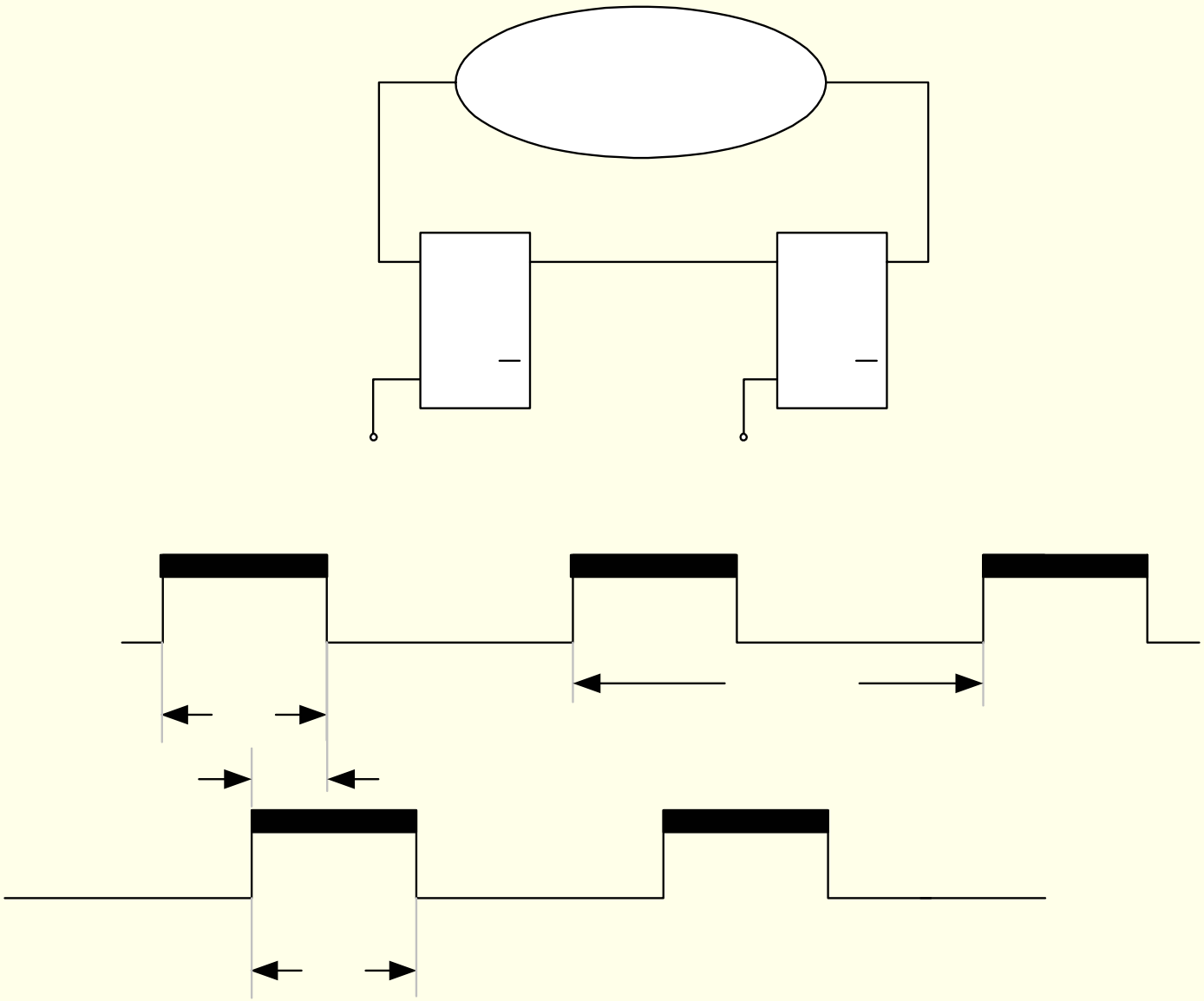
$$P \geq D_{LM} + D_{DQM}$$

$$D_{LmB} = 2(T_T + T_L) + H + U + D_{CQM} - D_{CQm} - D_{DQM}$$

In a single Latch system, it is possible to make the clock period *P* as small as the sum of the delays in the signal path: Latch and critical path delay in the logic. This can be achieved by adjusting the clock width *W*, while taking care of D_{LmB}

Analysis of a System with two-phase Clock and two Latches in an M-S arrangement

System using two-phase clock and two latches in M-S arrangement



System using two-phase clock and two latches in M-S arrangement

From the latest signal arrival analysis, several conditions can be derived. First, we need to assure an orderly transfer into L_2 Latch (*Slave*) from the L_1 Latch (*Master*), even if the signal arrived late (in the last possible moment) into the (*Master*) L_1 Latch. This analysis yields the following conditions:

$$W_2 \geq V + U_2 - U_1 + D_{1DQM} + T_{2T} + T_{1L}$$

$$W_1 + W_2 \geq V + U_2 + D_{1CQM} + T_{1T} + T_{2T}$$

These conditions assure timely arrival of the signal into the *L2* Latch, thus an orderly *L1-L2* transfer (from *Master* to *Slave*)

System using two-phase clock and two latches in M-S arrangement

The analysis of the latest arrival of the signal into *L1* Latch in the next cycle (critical path analysis) yields to the equations:

$$P \geq D_{1DQM} + D_{2DQM} + D_{LM}$$

$$W_1 \geq -P + D_{1CQM} + D_{2DQM} + U_1 + D_{LM} + T_{1L} + T_{1T}$$

$$P \geq -V + D_{2CQM} + U_1 + D_{LM} + T_{1T} + T_{2L}$$

This conditions assure timely arrival of the signal that starts on the leading edge of $\phi 1$, traverses the path through *L2*, the longest path in the logic and arrives before the trailing edge of $\phi 1$, in time to be captured.

The last equation shows that the amount of overlap *V* between the clocks $\phi 1$ and $\phi 2$ allows the system to run at greater speed.

System using two-phase clock and two latches in M-S arrangement

If we increase V we can tolerate longer "*critical path*" D_{LM}. However, the increase of the clock introduce a possibility of race conditions, thus requiring a *fast path* analysis. High-performance systems are designed with the objective of maximizing performance. Therefore, overlapping of the clocks is commonly employed, leading to the constraint of the minimal signal delay in the logic D_{LmB}:

$$D_{Lm} > D_{LmB} = V + H_1 + T_{1T} + T_{2L} - D_{2CQm}$$

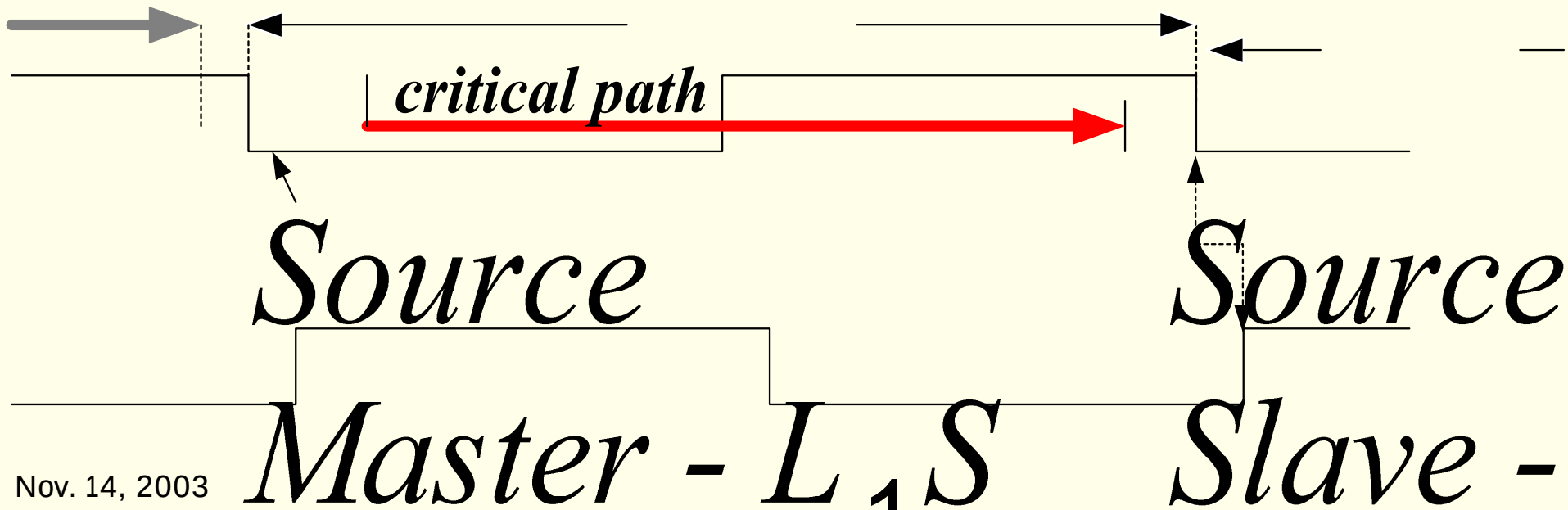
The maximal amount of overlap V that can be used is:

$$V_{\max} = T_{1T} + T_{2L} + D_{2CQM} + U_1 - D_{1DQM} - D_{2DQM}$$

For maximal performance, it is possible to adjust the clock overlap V so that the system runs at the maximal frequency.

M-S (L1-L2 latch) with non-overlapping clocks Φ_1 and Φ_2 obtained by locally generating clock Φ_2

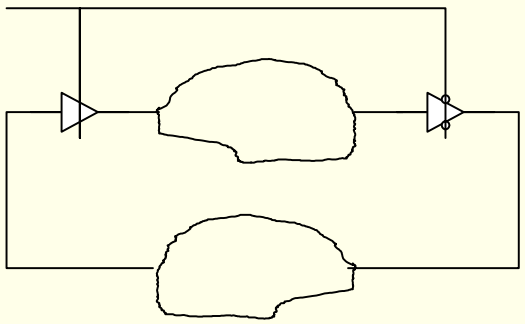
(This arrangement is also commonly referred to as flip-flop)

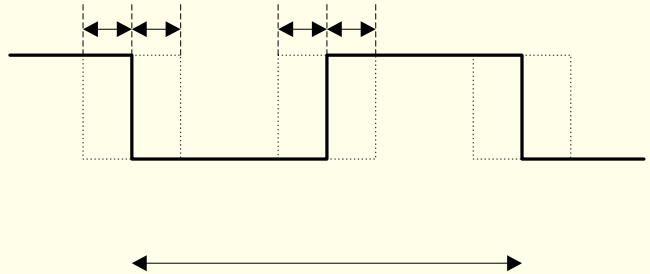


Example: Clocking in the first generation of Alpha Processor (WD21064)

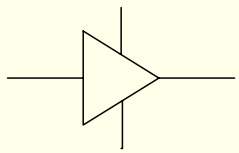
Timing arrangement and Latches used in the first-generation Alpha processor

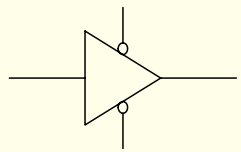
(a)

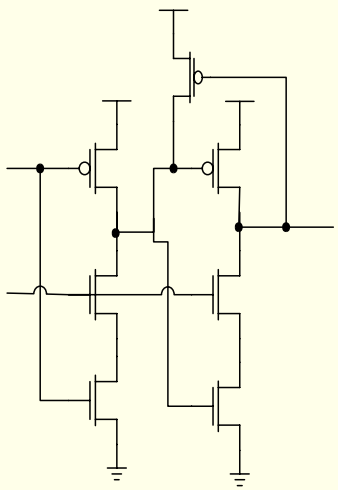


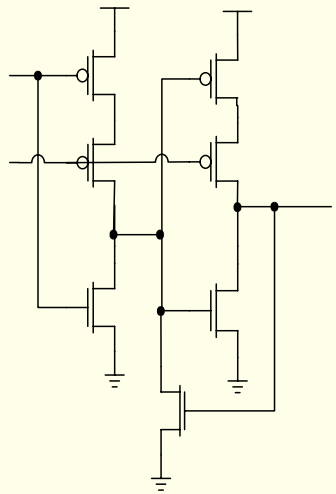


(b)









Clk

Example: Clocking the Alpha Processor

Clock skew: $T_L = T_T = 20ps$, for both edges of the clock.

Latch L1 parameters are: clock to Q delay $D_{CQM} = 50ps$, $D_{CQm} = 30ps$, D to Q delay $D_{DQM} = 60ps$, setup time U = 20ps, hold time H = 30ps.

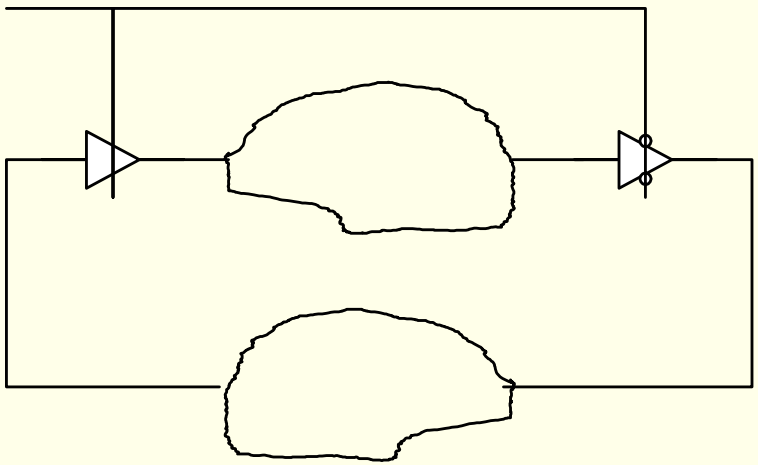
Latch L2 parameters are: $D_{CQM} = 60ps$, $D_{CQm} = 40ps$, $D_{DQM} = 70ps$, U = 30ps, H = 40ps.

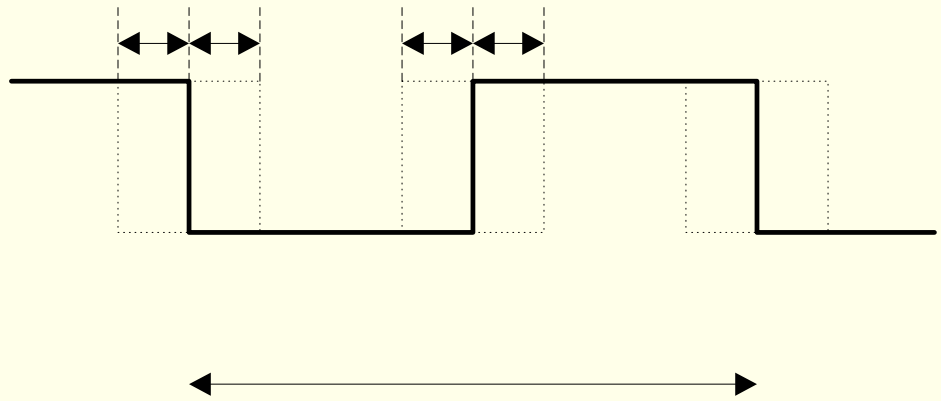
The critical paths in the logic sections 1 and 2 are:

D_{L1M}=200pS and D_{L2M}=170ps

Example: Clocking the Alpha Processor

For the given clock setup: $V_t = 0$ and clearly $P = W1 + W2$





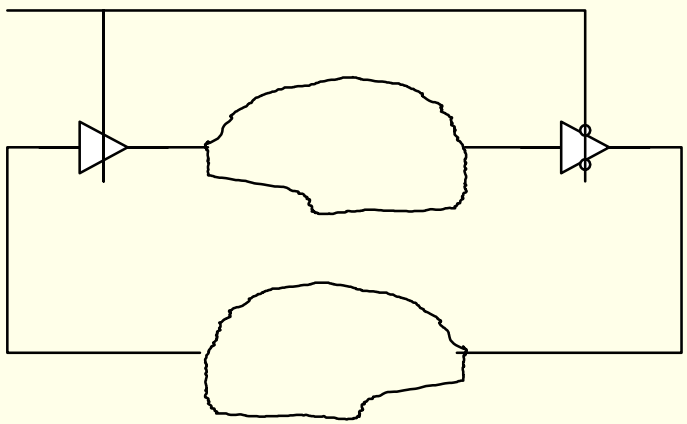
Example: Clocking the Alpha Processor

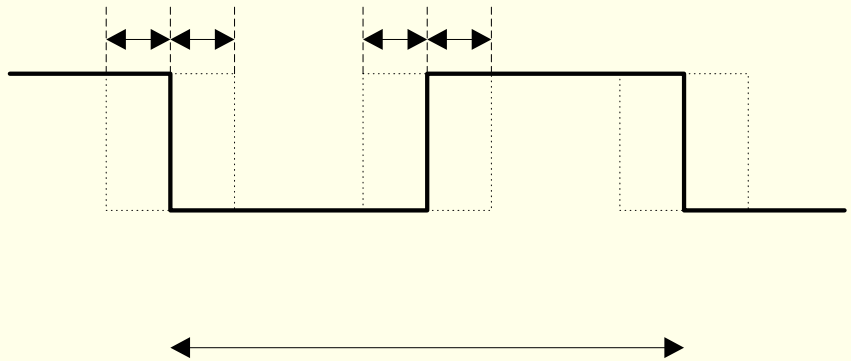
Nominal time, t=0 is set at the leading edge of the clock. The latest allowed data arrival times into latches L1 and L2, respectively:

$$t_{D1LArr} \le W_1 - T_T - U_1 \qquad t_{D2LArr} \le P - T_L - U_2$$

The latest arrival time of data into latch L2 is limited by the time at which latch L1 releases the data into the logic stage Logic1:

$$t_{D2LArr} = \max\left\{t_{D1LArr} + D_{1DQM}, T_L + D_{1CQM}\right\} + D_{L1M}$$





Example: Clocking the Alpha Processor

With the nominal time, t=0 set at the leading edge of the clock we obtain the latest allowed data arrival times into latches L1 and L2, respectively:

$$t_{D1LArr} \le W_1 - T_T - U_1 \qquad (4.34)$$

$$t_{D2LArr} \le P - T_L - U_2 \qquad (4.35)$$

The latest arrival time of data into latch L2 is limited by the time at which latch L1 releases the data into the logic stage Logic1:

$$t_{D2LArr} = \max\left\{t_{D1LArr} + D_{1DQM}, T_L + D_{1CQM}\right\} + D_{L1M}\,. \qquad (4.36)$$

Combining Eq. (4.35) and Eq. (4.36) we obtain:

$$W_1 - T_T - U_1 + D_{1DQM} + D_{L1M} \le P - T_L - U_2 \qquad (4.37a)$$

$$T_L + D_{1CQM} + D_{L1M} \le P - T_L - U_2\,. \qquad (4.37b)$$

Example: Clocking the Alpha Processor

Rearanging Eq. (4.37a,b) we obtain a set of bounds for W_2 and P:

$$W_2 \geq U_2 - U_1 + D_{1DQM} + T_L - T_T + D_{L1M} \qquad (4.38a)$$

$$P \geq U_2 + D_{1CQM} + 2T_L + D_{L1M} \,. \qquad (4.38b)$$

Due to the symmetry of the clocking scheme, moving the reference point from the leading edge of the clock to the trailing edge of the clock will give us the same equations with indexes interchanged. To check this, start from the equation dual to Eq. (4.36):

$$t_{D1LArr} = \max\left\{t_{D2LArr} + D_{2DQM}, -W_2 + T_T + D_{2CQM}\right\} + D_{L2M} \,. \qquad (4.39)$$

Example: Clocking the Alpha Processor

Combining Eq. (4.34) and Eq. (4.39) and rearranging, we obtain a set of bounds for W_1 and P:

$$W_1 \geq U_1 - U_2 + D_{2DQM} + T_T - T_L + D_{L2M} \qquad (4.40a)$$

$$P \geq U_1 + D_{2CQM} + 2T_T + D_{L2M}\,. \qquad (4.40b)$$

Combining Eq. (4.38a) and Eq. (4.40a) we obtain a third and often the most critical bound for the clock period:

$$P = W_1 + W_2 \geq D_{1DQM} + D_{L1M} + D_{2DQM} + D_{L2M}\,. \qquad (4.41)$$

Example: Clocking the Alpha Processor

Substituting the values into Eqs. (4.38,4.40 and 4.41) we obtain:

$$W_2 \geq 270\,ps$$
$$P \geq 320\,ps$$
$$W_1 \geq 230\,ps$$
$$P \geq 270\,ps\ ,$$

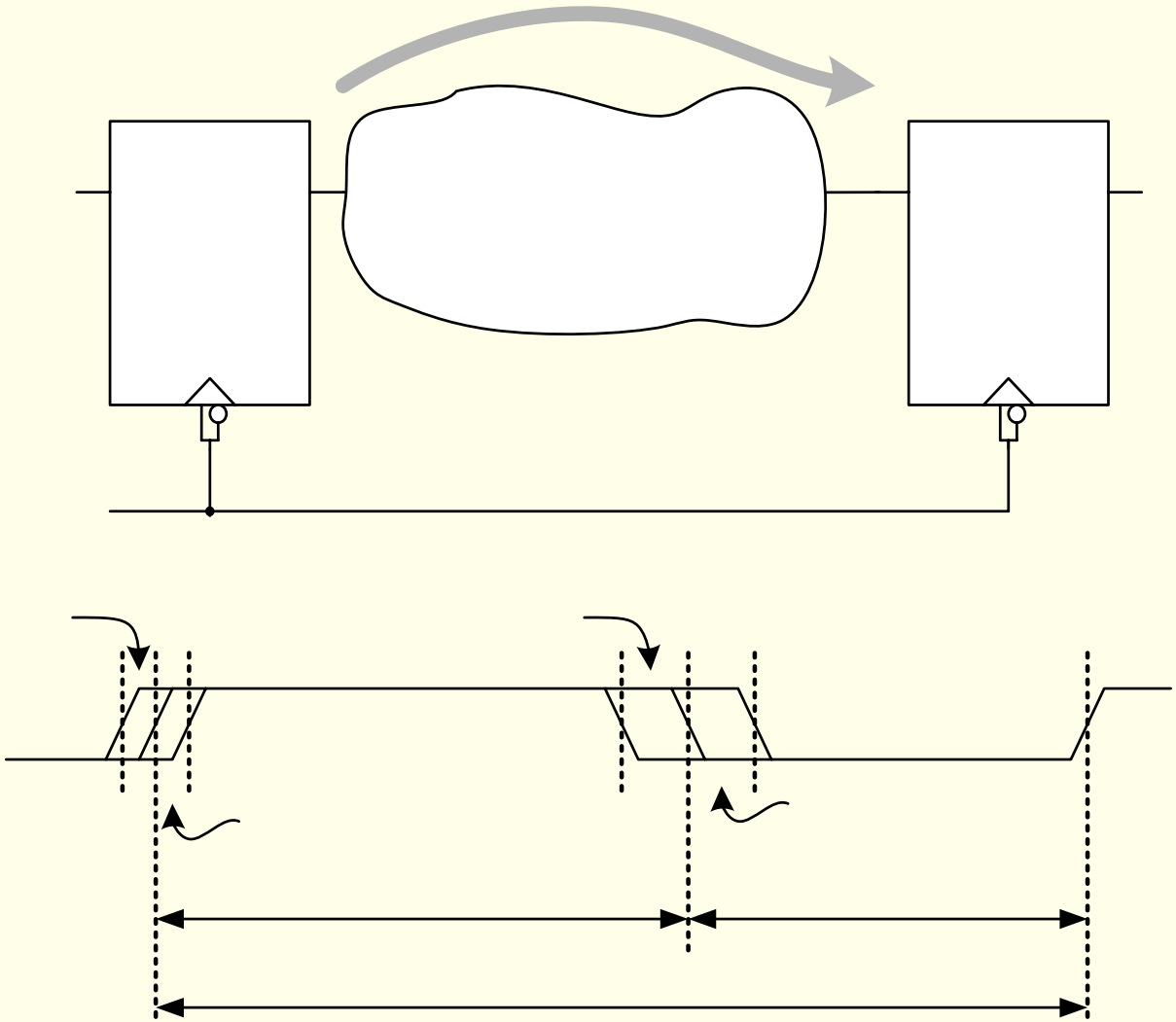
and the most critical bound for P,

$$P = W_1 + W_2 \geq 500\,ps\ .$$

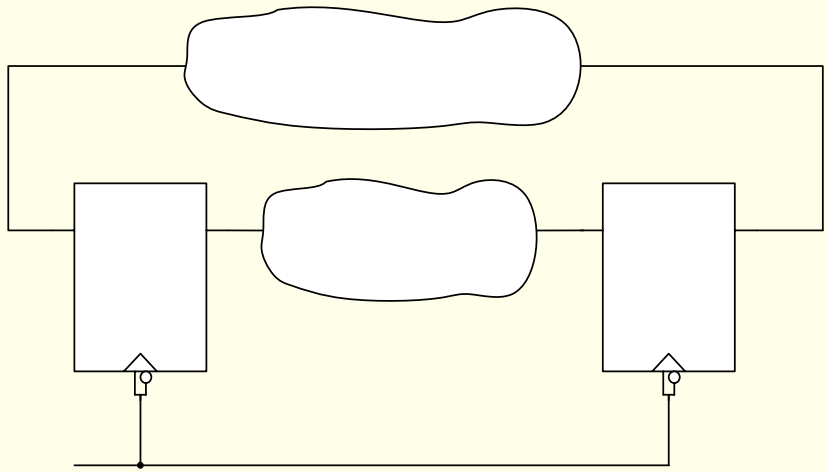
Thus the minimal clock period is P_{min}=500ps, and maximal frequency at which this system can run is f_{max}=2GHz.

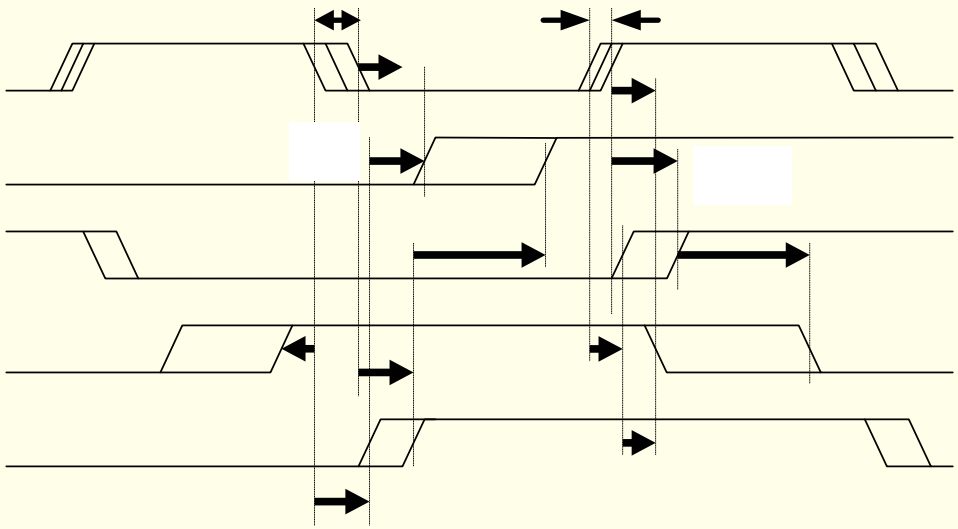
Digital system using a single-phase clock and dual-edge triggered storage elements

Digital system using a single-phase clock and dual-edge triggered storage elements



Two-stage dual-edge-triggered system





Allowed clock period as a function of the clock duty cycle in the dual-edge-triggered system