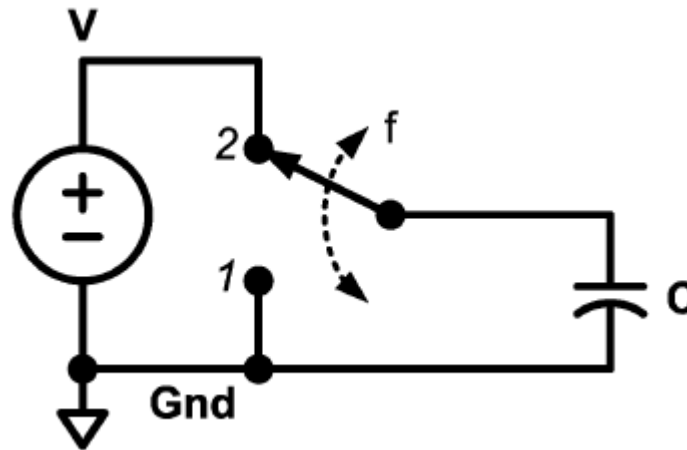


# Κατανάλωση ισχύος ψηφιακών κυκλωμάτων

Γιώργος Δημητρακόπουλος

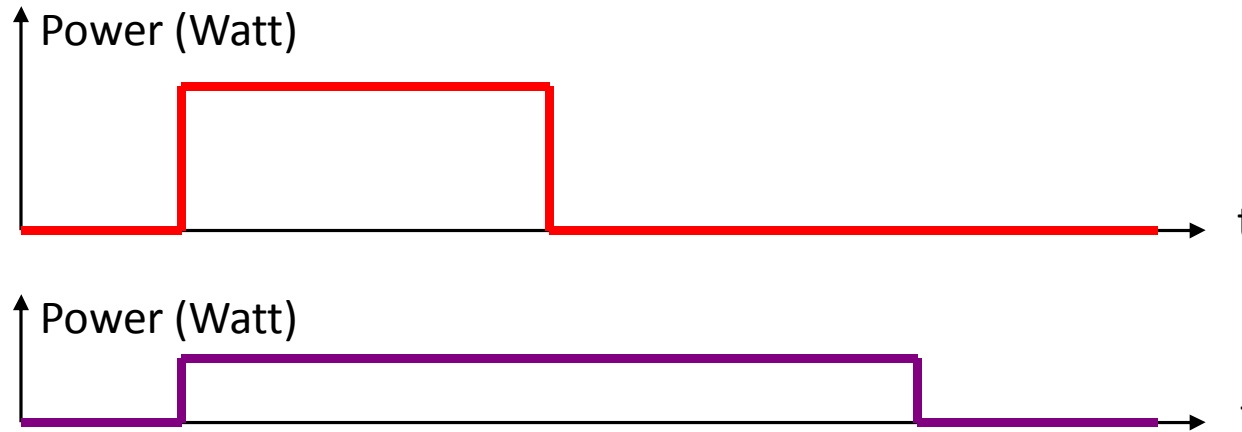
Τμήμα Επιστήμης Υπολογιστών  
Πανεπιστήμιο Κρήτης

# Δυναμική κατανάλωση ισχύος



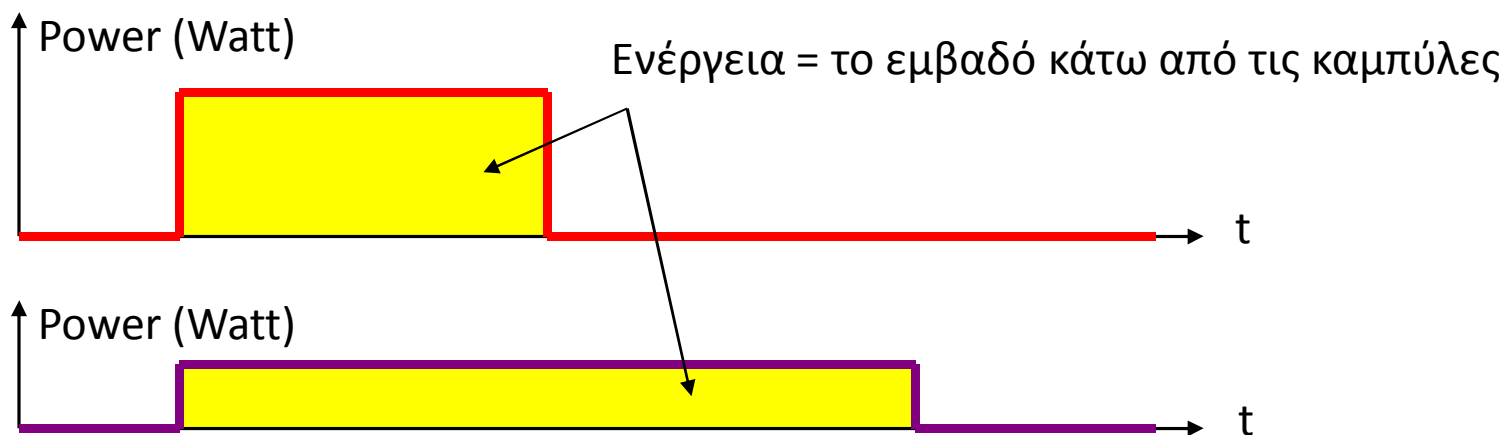
- Σε κάθε κύκλο μετακινούμε το διακόπτη από τη θέση 1 στη θέση 2 και πίσω
  - Φορτίζεις τον πυκνωτή σε τάση  $V$  και στη συνέχεια τον αποφορτίζεις στο Gnd
  - Το φορτίο κινείται από την πηγή τάσης στον πυκνωτή και από εκεί στη γή
  - Έστω ότι το κάνεις αυτό  $f$  φορές το δευτερόλεπτο. Δηλαδή με μία συχνότητα  $f$
- Το ρεύμα πάντα κυλαεί σε κάθε ολοκληρωμένο κύκλο από  $V$  στο Gnd
  - Σε κάθε κύκλο μετακίνησης του διακόπτη  $Q = CV$  coulombs μετακινούνται
  - Έχουμε  $f$  τέτοιους κύκλους το δευτερόλεπτο
  - Η ένταση του ρεύματος σε Ampere είναι  $CVf$  coulombs/second μεταξύ του  $V$  και του Gnd
  - Εφόσον η ισχύς είναι  $\text{Power} = V \times I$  τότε η ισχύς που καταναλώσαμε είναι  $V \times CVf = CV^2f$  Watts

# Ενέργεια - ισχύς



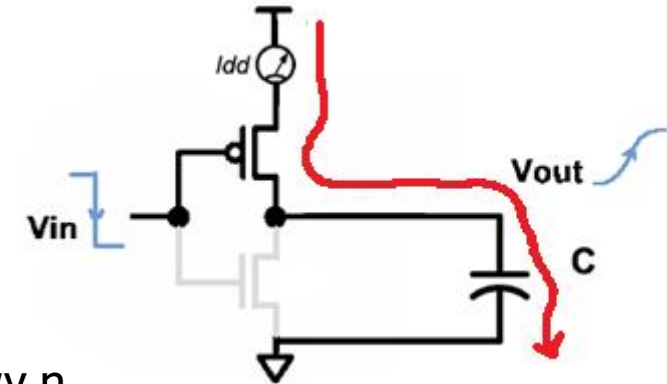
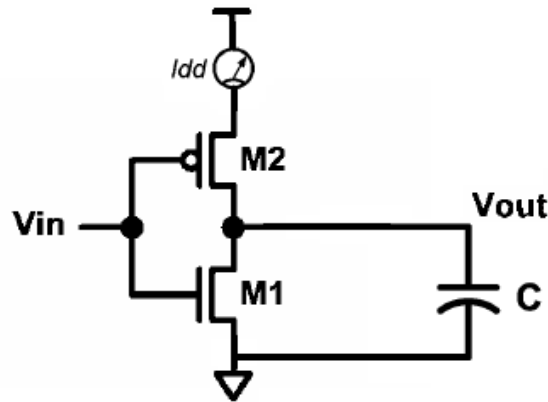
- Η ισχύς δύο εργασιών συναρτήσκει του χρόνου
- Η ισχύς ανάλογη της συχνότητας (ταχύτητας)
- Έτσι η εργασία (b) έχει τη μισή ισχύ από την (α) αλλά θέλει διπλάσιο χρόνο να ολοκληρωθεί

# Ενέργεια - ισχύς



- Η ενέργεια που δαπανούμε και για τις δύο εργασίες ακριβώς η ίδια  
 $Energy = Power \times Time$
- Στην εργασία (β) έχουμε μειώσει τη συχνότητα μειώνοντας την ισχύ αλλά δε μειώνουμε τη ενέργεια που χρειάζεται να δαπανήσουμε
  - Ξοδεύουμε τη μισή ισχύ για διπλάσιο χρόνο → Η ενέργεια ίδια
- Ουσιαστικά Ενέργεια/λειτουργία = Ισχύς x Χρόνος / λειτουργία
- Σε περιπτώσεις με αυστηρούς χρονικούς περιορισμούς θα διαλέγαμε τη λύση (α). Στην αντίθετη περίπτωση τη (β)

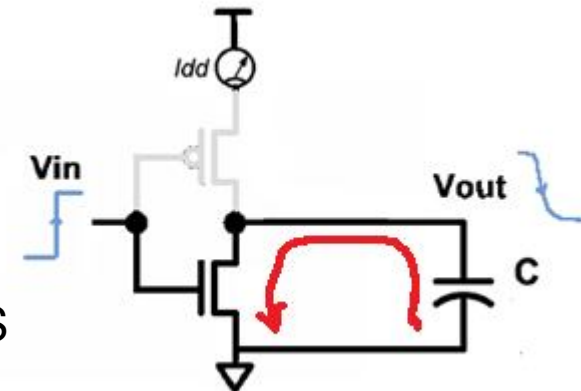
# Η δυναμική κατανάλωση ισχύος ψηφιακών πυλών



- Η δυναμική κατανάλωση ισχύος συμβαίνει μόνο όταν η πύλη αλλάζει κατάσταση
- Σε κάθε κύκλο  $CV_{dd}$  φορτίζει προς τον πυκνωτή
- Σε κάθε κύκλο ξοδεύουμε ενέργεια

$$E_c = CV_{dd}^2$$

- Κατά τη φόρτιση του πυκνωτή αποθηκεύεται σ' αυτόν το  $\frac{1}{2} CV_{dd}^2$ . Το άλλο μισό χάνεται στην «αντίσταση» του PMOS τρανζιστορ
- Όταν η έξοδος επιστρέφει στο 0 αυτή η ενέργεια που είχε ο πυκνωτής χάνεται στην «αντίσταση» του NMOS τρανζιστορ
- Το κύκλωμα καταναλώνει την ισχύ όταν η είσοδος κάνει μετάβαση από το 0 στο 1.



## Μέση δυναμική κατανάλωση ισχύος

- Σε κάθε περίπτωση που η έξοδος μεταβαίνει από 0 σε 1 ξοδεύουμε ενέργεια ίση με

$$E = C_{\text{SWITCHED}} V_{\text{dd}}^2$$

- Η ισχύς είναι η ενέργεια που δαπανούμε ανα μονάδα χρόνου
- Αν κάνουμε υπολογισμούς με συχνότητα  $f_{\text{clk}}$  τότε

$$P_{\text{DYNAMIC}} = f_{\text{clk}} \times C_{\text{SWITCHED}} V_{\text{dd}}^2$$

- Αν από τις  $f_{\text{clk}}$  φορές το δευτερόλεπτο που προσπαθούμε να υπολογίσουμε ένα νέο αποτέλεσμα η έξοδος κάνει μετάβαση από το 0 στο 1 με πιθανότητα  $\alpha$  τότε

$$P_{\text{DYNAMIC}} = \alpha \times f_{\text{clk}} \times C_{\text{SWITCHED}} \times V_{\text{dd}}^2$$

# Παράγοντας μεταβάσεων

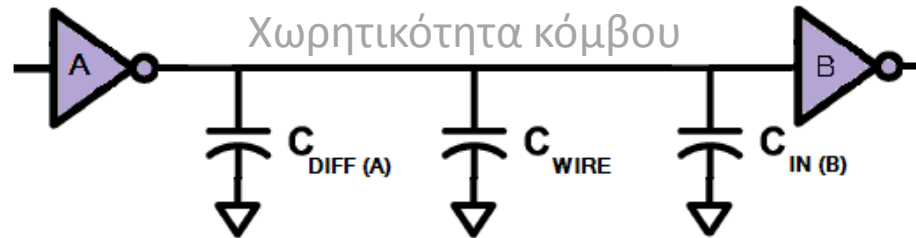
$$P_{dynamic} = \alpha \cdot C_{switched} \cdot V_{DD}^2 \cdot f_{clk}$$

Παράγοντας μεταβάσεων (switching activity)  $\alpha$  = πόσες φορές το σήμα εισόδου κάνει μετάβαση  $0 \rightarrow 1 \rightarrow 0$  ανά περίοδο ρολογιού

- Το σήμα του ρολογιού έχει  $\alpha = 1$  (και για το λόγο αυτό συνεισφέρει σημαντικά στην κατανάλωση ισχύος του κυκλώματος)
- Η έξοδος ενός flip-flop μπορεί να έχει στη χειρότερη περίπτωση  $\alpha = 1/2$
- Δεν αλλάζουν κατάσταση όλοι οι κόμβοι του κυκλώματος σε κάθε κύκλο. Πόσο συχνά συμβαίνει μια αλλαγή εξαρτάται από τις τιμές των δεδομένων. Συνήθως η συνδιαστική λογική των κυκλωμάτων περιγράφεται με  $\alpha = 0.1-0.2$
- Για να έχεις μικρότερη κατανάλωση ισχύος πρέπει να μπορέσεις να μειώσεις τον παράγοντα των μεταβάσεων  $\alpha$

# Χωρητικότητα του κάθε κόμβου

$$P_{dynamic} = \alpha \cdot C_{switched} \cdot V_{DD}^2 \cdot f_{clk}$$



- Η χωρητικότητα  $C_{switched}$  που συμμετέχει στην κατανάλωση ισχύος περιλαμβάνει:
  - Τη χωρητικότητα του ακροδέκτη Gate των τρανζίστορ (χωρητικότητα εισόδου) της επόμενης πύλης  $C_{IN(B)}$
  - Τις παρασιτικές χωρητικότητες των περιοχών διάχυσης των τρανζίστορ που είναι συνδεδεμένες στην έξοδο της πύλης  $C_{DIFF(A)}$
  - Τις χωρητικότητες των ενδιάμεσων καλωδίων  $C_{WIRE}$
  - Για να μειώσεις την κατανάλωση ισχύος πρέπει να μειωθεί η χωρητικότητα  $C_{switched}$

$$C_{switched} = C_{DIFF(A)} + C_{WIRE} + C_{IN(B)}$$



# Δυναμική κατανάλωση ισχύος - τάση τροφοδοσίας/συχνότητα λειτουργίας

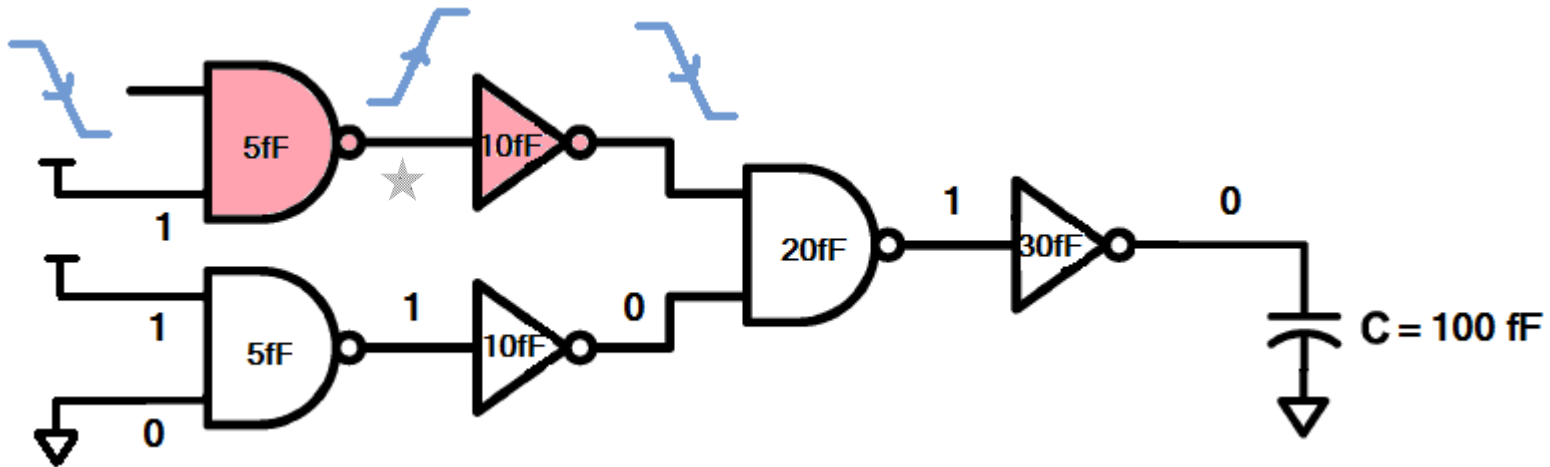
$$P_{dynamic} = \alpha \cdot C_{switched} \cdot V_{DD}^2 \cdot f_{clk}$$

- Για να μειωθεί η κατανάλωση ισχύος πρέπει να μειωθεί η τάση τροφοδοσίας  $V_{DD}$
- Αυτό επηρεάζει άμεσα την ταχύτητα των κυκλωμάτων. Η καθυστέρηση μιας πύλης ακολουθεί χοντρικά τον κανόνα

$$T = K \frac{C_L V_{DD}}{(V_{DD} - V_T)^b}$$

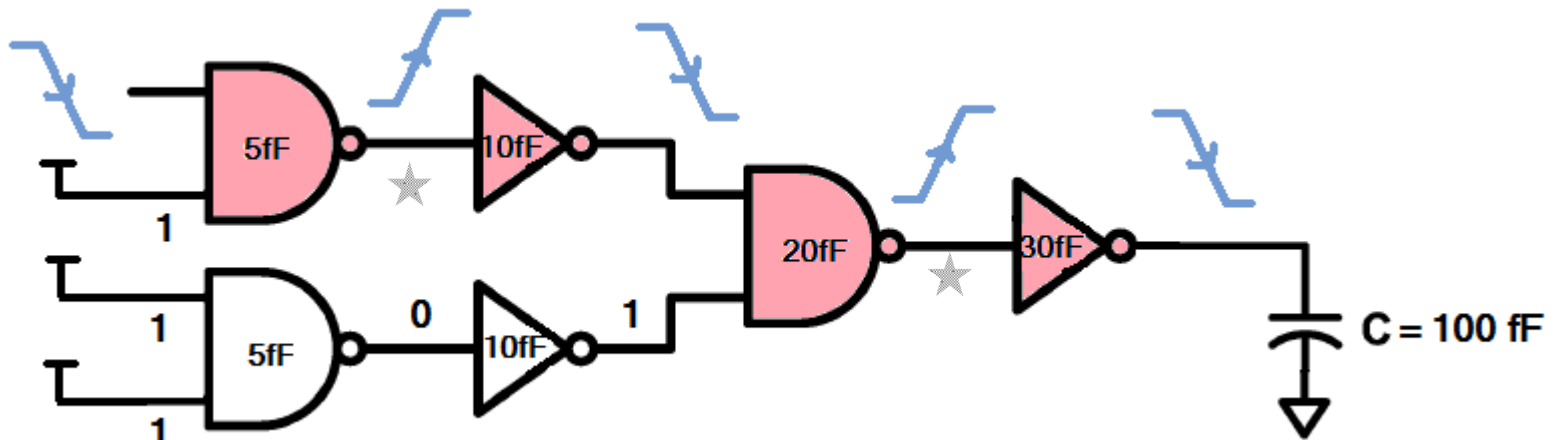
- Επομένως η μείωση του  $V_{DD}$  σημαίνει και αναγκαστική μείωση της συχνότητας λειτουργίας  $f_{clk}$ .
- Σκοπός είναι χωρίς να μεταβάλουμε την καθυστέρηση να μειώσουμε την ισχύ μειώνοντας ουσιαστικά την ενέργεια που απαιτείται ανα λειτουργία
- Επίσης χωρίς να πειράξουμε το  $V_{DD}$  κλείνοντας το ρολόι στις ανενεργές μονάδες μειώνεται το  $\alpha \times f_{clk}$ . Φυσικά αυτή η λύση εισάγει πολλά άλλα προβλήματα τα οποία απαιτούν προσεκτικές λύσεις

# Παράδειγμα υπολογισμού δυναμικής καταναλωσης ισχύος



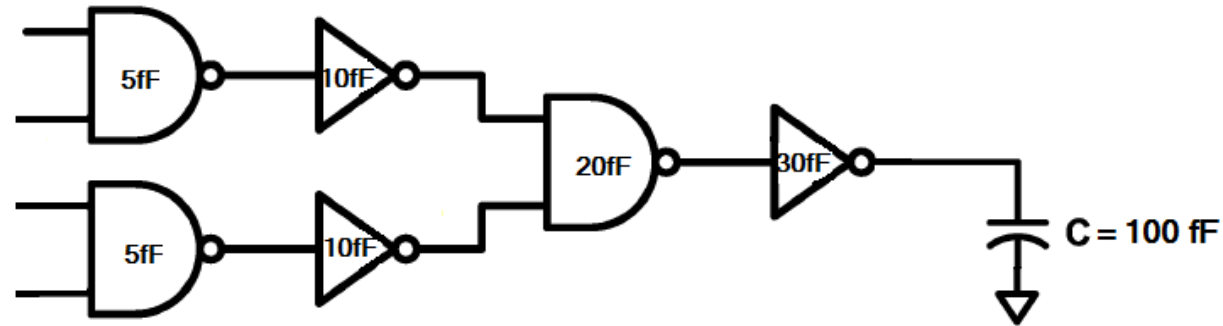
- Εδώ δεν μας απασχολεί η μέση τιμή του παράγοντα μεταβάσεων εφόσον μελετάμε την κατανάλωση ισχύος για μία συγκεκριμένη μετάβαση
- Οι χωρητικότητες που αναγράφονται στις πύλες είναι οι χωρητικότητες της κάθε μιας εισόδου
- Οι μεταβάσεις που καταναλώνουν ενέργεια είναι αυτές που πηγαίνουν από το 0 στο 1 (Κόμβοι με ★)
- Υπάρχει 1 μία μόνο τέτοια μετάβαση η οποία φορτίζει την είσοδο του inverter με κόκκινο χρώμα
- Αν  $V_{DD} = 1.2 \text{ V}$  τότε  $E = 10 \text{ fF} \times V_{DD}^2 = 10 \text{ fF} \times 1.44 = 14.4 \text{ fJ}$ . Αν το κύκλωμα λειτουργεί στα  $100 \text{ MHz}$  τότε  $P = 14.4 \text{ fJ} \times 100 \text{ MHz} = 1.4 \text{ nW}$

# Παράδειγμα υπολογισμού δυναμικής καταναλωσης ισχύος



- Σε αυτή την περίπτωση υπάρχουν πολλοί κομβοί οι οποίοι μεταβαίνουν από το 0 στο 1
- Έτσι  $E = (10\text{fF} + 30\text{fF}) V_{\text{dd}}^2 = 40\text{ fF} \times 1.44 = 57.6\text{ fJ}$  το οποίο για  $f = 100\text{ MHz}$  δίνει  $5.76\text{ nW}$  (5x σε σχέση με πριν)
- Αν αλλάζαμε τη συχνότητα του ίδιου κυκλώματος χωρίς να αλλάξουμε τα μεγέθη των πυλών του σε  $f = 50\text{ MHz}$  η ισχύς θα μειωνόταν στα  $2.88\text{ nW}$ 
  - Η ενέργεια που ξοδεύουμε γι' αυτόν τον υπολογισμό παραμένει η ίδια αφού έχουμε τη μισή ισχύ να δαπανάται για διπλάσιο χρόνο

## Παράδειγμα υπολογισμού μέσης κατανάλωσης ισχύος



- Για το ίδιο κύκλωμα θα μελετήσουμε τη μέση κατανάλωση ισχύος - Χωρίς να γνωρίζουμε δηλαδή τα ακριβή διανύσματα που εμφανίζονται στις εισόδους
- Θα προσπαθήσουμε να εκτιμήσουμε τον παράγοντα μεταβάσεων για κάθε κόμβο του κυκλώματος χρησιμοποιώντας πιθανοτικές μεθόδους.
- Γνωρίζουμε τις πιθανότητες οι εισόδοι να είναι στο 0 ή στο 1  $P(\text{input}=0)$  και  $P(\text{input}=1)$ . Όσο πιο ακριβείς είναι οι εκτιμήσεις για το στατιστικό περιεχόμενο των εισόδων τόσο πιο ρεαλιστικά τα αποτελέσματα
- Γνωρίζοντας αυτές τις πιθανότητες και τη λογική λειτουργία της πύλης μπορούμε να υπολογίσουμε την πιθανότητα η έξοδος της κάθε πύλης να είναι είτε στο 0 είτε στο 1.
- Τότε, η πιθανότητα μια πύλη να κάνει μετάβαση από το 0  $\rightarrow$  1 είναι η πιθανότητα η προηγούμενη τιμή της εξόδου να είναι 0 ενώ η επόμενη να είναι 1:

$$\alpha = P_{\text{SWITCH}} = P(\text{out} = 0) \times P(\text{out} = 1) = (1 - P(\text{out}=1)) \times P(\text{out} = 1)$$

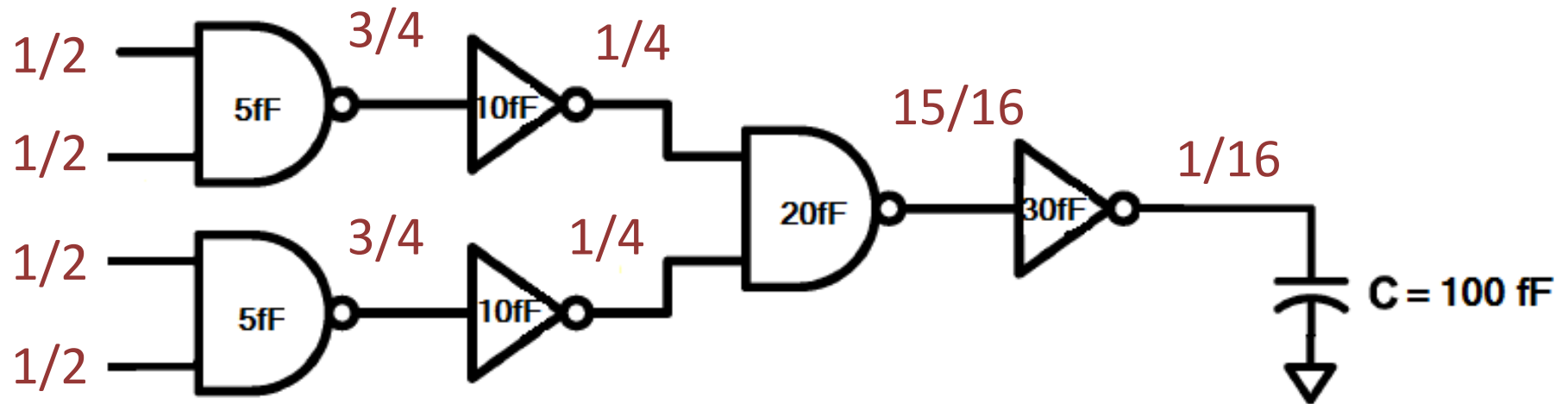
# Υπολογισμός παράγοντα μεταβάσεων



- Γνωρίζουμε την πιθανότητα οι είσοδοι να είναι 1
  - Επομένως η πιθανότητα η είσοδος 0 είναι  
 $P(\text{input} = 0) = 1 - P(\text{input}=1)$
- Η έξοδος της NAND θα είναι 1 όταν τουλάχιστον μία από τις εισόδους 0  
$$P(Z=1) = P(A=0) \cdot P(B=0) + P(A=0) \cdot P(B=1) + P(A=1) \cdot P(B=0)$$
- Η έξοδος της NAND θα είναι 0 όταν και οι δύο είσοδοι 1 (απλούστερο)  
$$P(Z=0) = P(A=1) \cdot P(B=1)$$
 έτσι ισοδύναμα  
$$P(Z=1) = 1 - P(Z=0) = 1 - P(A=1) \cdot P(B=1)$$

# Επιστροφή στο παράδειγμα

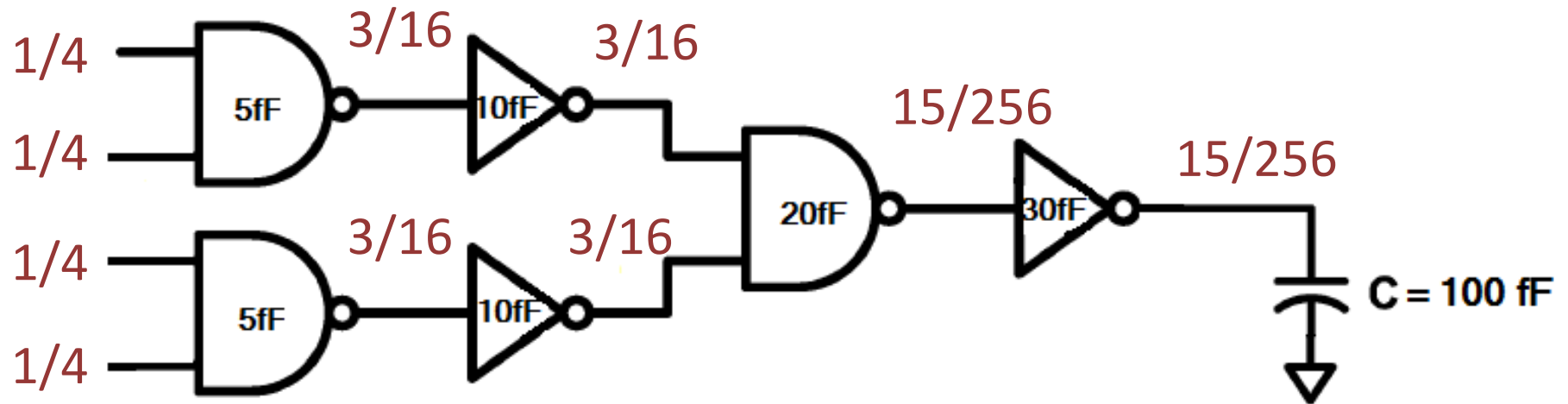
Πιθανότητες ο κάθε κόμβος να είναι ίσος με 1



- Οι πιθανότητες ο κάθε κόμβος να έχει την τιμή 1 υπολογίζεται όπως και πριν με βάση τις πιθανότητες των εισόδων ( $1/2$  ισοπίθανα στο 0 ή στο 1) και τη λογική λειτουργία των πυλών
- Οι πιθανότητες ο κόμβος να είναι 0 μπορούν απ'ευθείας να υπολογιστούν
- Αυτό που μένει είναι να υπολογίσουμε τον παράγοντα μετάβασης  $\alpha$  για κάθε κόμβο  $= (1 - P(x=1)) \times P(x = 1)$

# Ολοκλήρωση παραδείγματος

Παράγοντες μεταβάσεων  $\alpha$  ( $0 \rightarrow 1$ ) - switching activities



- Κάθε κόμβος φορτίζεται από το 0 στο 1 με πιθανότητα ίση με αυτή που αναγράφεται σε κάθε κόμβο
  - Η ενέργεια του κάθε κόμβου  $E = \alpha \times C_{\text{switched}} \times V_{DD}^2$
- Επομένως η εκτιμώμενη συνολική ενέργεια
  - $(4 * \frac{1}{4} * 5\text{fF} + 2 * \frac{3}{16} * 10\text{fF} + 2 * \frac{3}{16} * 20\text{fF} + 15/256 * 30\text{fF} + 15/256 * 100\text{fF}) * V_{DD}^2 = 30,8\text{fF} * V_{DD}^2$
- Η συνολική χωρητικότητα είναι 180 fF. Επομένως το effective switching activity είναι περίπου  $30,8/180 \approx 1/6 = 0.167$ .
  - Δηλαδή αναμένουμε πως όταν οι είσοδοι είναι ισοπίθانا στο 0 στο 1 τότε μόνο το 1/6 της συνολικής χωρητικότητας να υποστεί μετάβαση από το 0 στο 1. Προσέξτε πως μας ενδιαφέρει να έχουμε μικρό παράγοντα μεταβάσεων σε κόμβους με μεγάλη χωρητικότητας (όπως οι 2 πιο κοντά στην έξοδο του παραδείγματος)

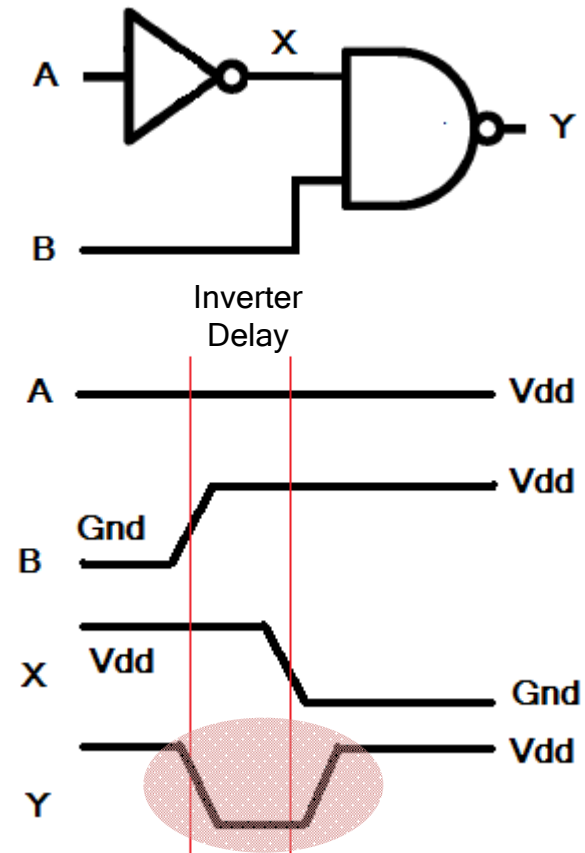
# Σχολιασμός

- Για το ίδιο κύκλωμα υπολογίσαμε πως η ισχύς του στα 100 MHz παίρνει τιμές
  - Για μια μετάβαση με ένα κόμβο από το 0 στο 1 = 1.4nW
  - Για μια μετάβαση με περισσότερους κόμβους από το 0 στο 1 = 5.6nW
  - Η μέση κατανάλωση ισχύος που περιμένου αν είσοδοι ισοπίθανα στο 0 ή στο 1 = 3.8nW
- Η στιγμιαία μέγιστη ισχύς μπορεί να διαφέρει πολύ από τη μέση κατανάλωση ανάλογα με τη μετάβαση που συμβαίνει στο κύκλωμα
  - Στην πράξη και οι δύο τιμές είναι σημαντικές
- Η ακριβής εκτίμηση γίνεται πολύ δύσκολη στην περίπτωση που τα σήματα δεν είναι ανεξάρτητα πιθανοτικά
  - Εμείς θεωρήσαμε πως η μετάβαση κάθε σήματος στην είσοδο είναι ένα ανεξάρτητο γεγονός
- Η ισχύς μεταβάλλεται ανάλογα με τη συχνότητα που λειτουργούμε το κύκλωμα. Φυσικά και η συχνότητα έχει ένα πάνω όριο που εξαρτάται από την ελάχιστη καθυστέρηση του κυκλώματος



# Παράλειψη glitches

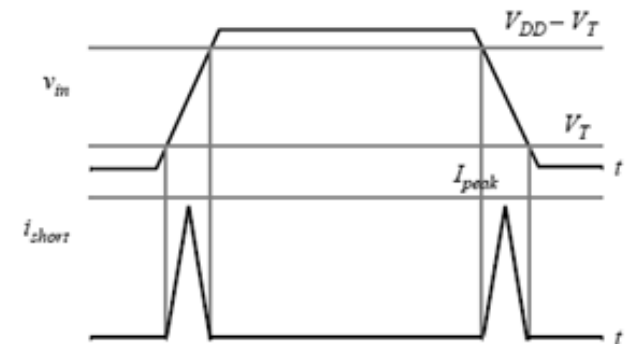
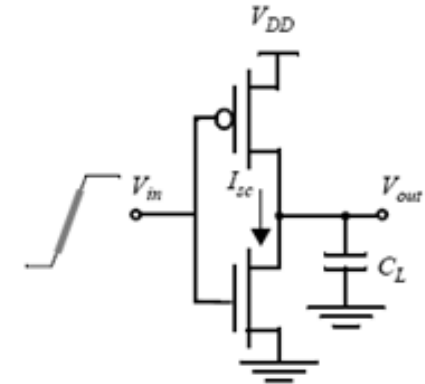
- Οι αναλυτικές μέθοδοι μπορούν να πλησιάσουν πολύ κοντά στην πραγματικότητα όταν είναι γνωστή η στατιστική πληροφορία των εισόδων
- Στην εκτίμηση της δυναμικής κατανάλωσης ενέργειας δε λάβαμε υπόψιν τις στιγμιαίες λανθασμένες μεταβάσεις των εξόδων (glitches) οι οποίες προέρχονται από τις χρονικές διαφορές στη διάδοση των σημάτων από μονοπάτια διαφορετικής καθυστέρησης
- Για την εκτίμηση της μέγιστης κατανάλωσης ισχύος πρέπει να αναζητήσουμε εκείνα τα ζευγάρια των εισόδων που προκαλούν τις περισσότερες μεταβάσεις στο κύκλωμα. Προσοχή!!! Μας απασχολεί να δούμε αυτές τις μεταβάσεις στους κόμβους με τη μεγαλύτερη χωρητικότητα
- Η μόνη έγκυρη μέθοδος για την εκτίμηση της κατανάλωσης ισχύος είναι η εξομοίωση του κυκλώματος χρησιμοποιώντας ταυτόχρονα πραγματική χρονική πληροφορία για τις πύλες και τα καλώδια (timing)



Μη επιθυμητή προσωρινή μετάβαση (glitch) η οποία αποφορτίζει στιγμιαία τη χωρητικότητα του κόμβου Y

# Ρεύμα βραχυκυκλώματος

- Κατά τη διάρκεια της μετάβασης σε μια στατική πύλη CMOS τα PMOS και τα NMOS τρανζίστορ είναι ταυτόχρονα ενεργοποιημένα για ένα σύντομο χρονικό διάστημα.
- Στο διάστημα αυτό σχηματίζεται ένα απευθείας μονοπάτι ρεύματος μεταξύ του Vdd και το Gnd. Η ισχύς του ρεύματος αυτού καταναλίσκεται στα τρανζίστορ χωρίς να συνεισφέρει στη φόρτιση της χωρητικότητας εξόδου
- Η διάρκεια αυτού του χρονικού διαστήματος εξαρτάται από το χρόνο ανόδου ή καθόδου των εισόδων και του σήματος της εξόδου
- Σε ένα καλά σχεδιασμένο κύκλωμα (χρόνοι ανόδου/καθόδου - μέτρηση 10%-90% μεταξύ 1 - 1.5 FO4) η κατανάλωση ισχύος λόγω του ρεύματος βραχυκυκλώματος είναι το 5 - 10% της συνολικής δυναμικής κατανάλωσης ισχύος
- Αυτή η απώλεια ισχύος πραγματοποιείται μόνο όταν η πύλη αλλάζει κατάσταση. Ουσιαστικά είναι μέρος της δυναμικής κατανάλωσης ισχύος απλά αντιμετωπίζεται χωριστά
- Το ρεύμα βραχυκυκλώματος εξαρτάται από την τιμή του Vdd. Αν  $V_{dd} < V_{TN} + |V_{TP}|$  τότε τα τρανζίστορ δε είναι ποτέ ενεργοποιημένα ταυτόχρονα και έτσι το φαινόμενο εξαλείφεται



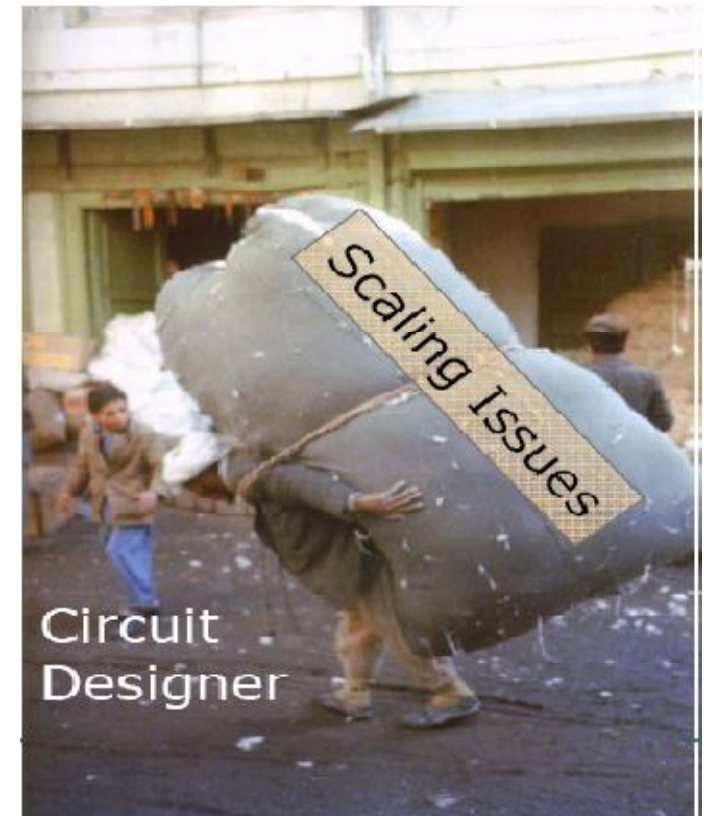
# Ρεύματα διαρροής

- Δυστυχώς ακόμη και όταν  $V_{GS} < V_T$  το τρανζίστορ διαρρέεται από μια ποσότητα ρεύματος η οποία δυστυχώς με τη σμίκρυνση της τεχνολογίας συνεχώς αυξάνει
- Επομένως ουσιαστικά το κύκλωμα παρουσιάζει κατανάλωση ισχύος ακόμη και όταν δεν έχουμε μεταβάσεις στην έξοδο της πύλης
- Τα ρεύματα διαρροής αυξάνουν
  - Με την άνοδο της θερμοκρασίας
  - Με τη μείωση της τάσης κατωφλίου των τρανζίστορ και μάλιστα εκθετικά (τη μείωση αυτή τη χρειαζόμαστε για την αύξηση της ταχύτητας του κυκλώματος γιατί χαμηλό κατώφλι → μεγάλο ρεύμα → γρήγορη φόρτιση/εκφόρτιση χωρητικότητας εξόδου)
- Τα ρεύματα διαρροής εξαρτώνται σημαντικά από την τάση τροφοδοσίας του κυκλώματος
- Ενώ ξεκίνησε σαν ένα αμελητέο ποσό της συνολικής κατανάλωσης ισχύος στις μέρες μας η κατανάλωση λόγω των ρευμάτων διαρροής μπορεί να αγγίξει έως και το 30% της συνολικής κατανάλωσης ισχύος
- Προς το παρόν οι καινούριες τεχνολογίες μαζί με τις σχεδιαστικές τεχνικές που προέκυψαν έχουν ανταποκριθεί και κρατούν τα ρεύματα διαρροής σε ανεκτά επίπεδα. Βέβαια νέες σχεδιαστικές λύσεις συνεχώς θα απαιτούνται

# Συνολική κατανάλωση ισχύος

$$P \approx \underbrace{C_{TOT} \cdot \alpha \cdot F \cdot V_{dd}^2}_{\text{Dynamic Power}} + \underbrace{N_{TOT} \cdot \alpha \cdot F \cdot V_{dd} \cdot I_{sc}}_{\text{Short-Circuit Power}} + \underbrace{N_{ON} \cdot I_{LEAK} \cdot V_{dd}}_{\text{Leakage Power}}$$

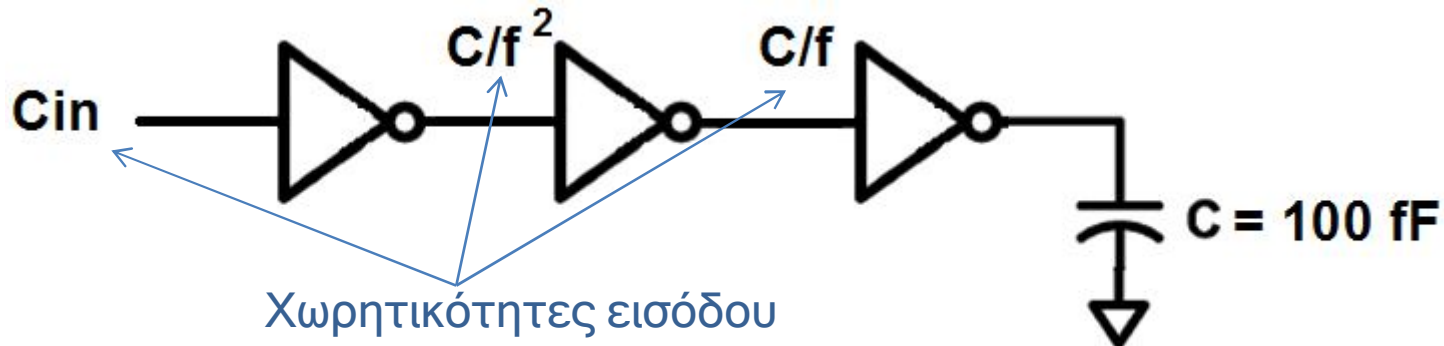
- Η συνολική κατανάλωση ισχύος αποτελείται από τρία μέρη
  - Τη δυναμική κατανάλωση ισχύος
  - Τη συνεισφορά του ρεύματος βραχυκυκλώματος
  - Τα ρεύματα διαρροής
- Οι μικρότερου μεγέθους τεχνολογίες δυσκολεύουν ολοένα και περισσότερο το πρόβλημα
- Οι επιδόσεις των κυκλωμάτων περιορίζονται πια από την κατανάλωση ισχύος τους
  - Προβλήματα θερμοκρασίας
  - Πακεταρίσματος του ολοκληρωμένου
  - Χρόνος αντοχής μπαταρίας για φορητές συσκευές



## Συσχέτιση ενέργειας καθυστέρησης

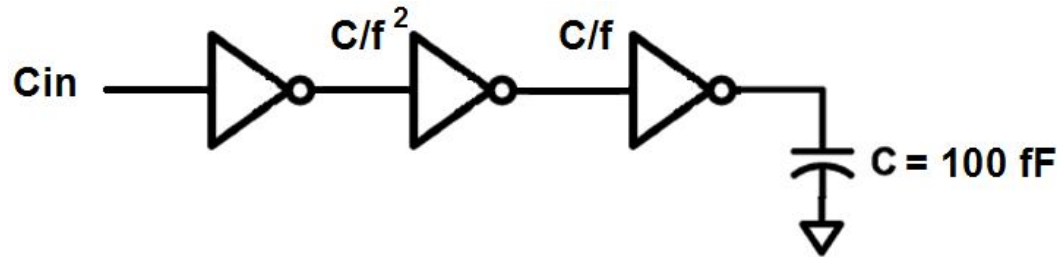
- Η ελάχιστη καθυστέρηση την οποία μπορεί να επιτύχει ένα κύκλωμα για ένα δεδομένο  $V_{DD}$  συνδέεται άμεσα με την κατανάλωση ενέργειας γιατί από αυτό εξαρτάται η τιμή του  $C_{SWITCHED}$ 
  - Το  $\alpha$  εξαρτάται από την τοπολογία του κυκλώματος και τη λογική συνάρτηση που υλοποιεί
- Έτσι ανάλογα με την καθυστέρηση που στοχεύουμε αλλάζει αυτόματα και η δυναμική ενέργεια (και κατανάλωση ισχύος για μια δεδομένη συχνότητα λειτουργίας)

# Παράδειγμα

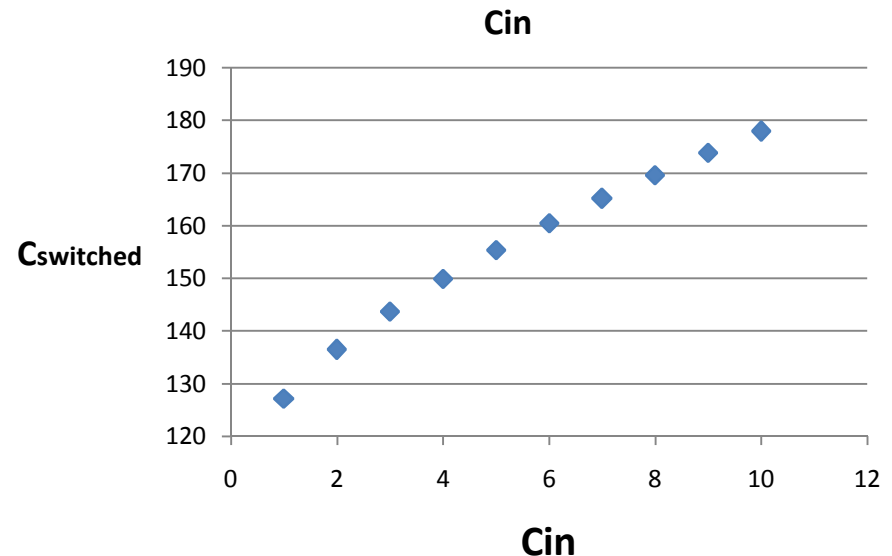
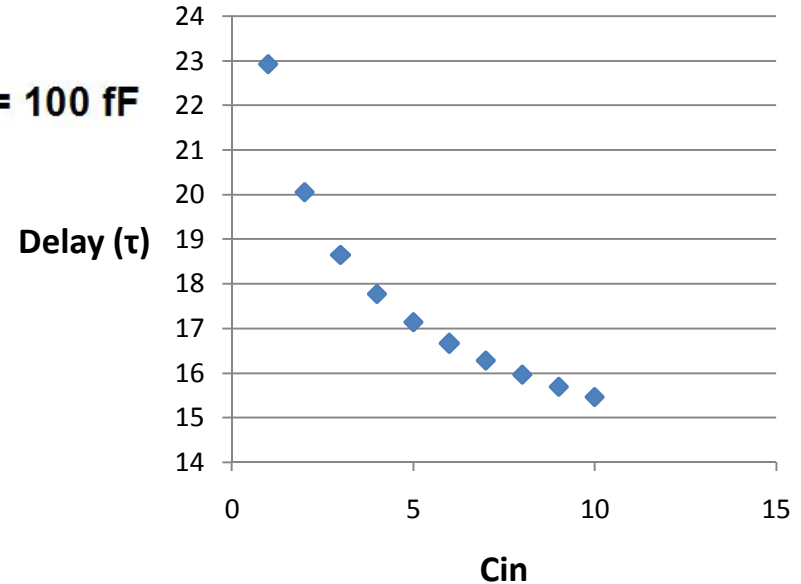


- Επιλέγουμε τα μεγέθη των αντιστροφέντων κυνηγώντας την ελάχιστη καθυστέρηση για σταθερό  $V_{DD}$
- Έτσι  $f = (100/C_{in})^{1/3}$
- $\text{Delay} = 3\tau_{INV}(f + p_{INV})$
- Η χωρητικότητα εισόδου της κάθε πύλης για την ελάχιστη καθυστέρηση  $C/f^j$ . Η τιμή του  $j$  ανάλογα με τη θέση στο μονοπάτι. Αυτή που οδηγεί την έξοδο  $j=1$  η προ-τελευταία  $j=N-1$
- $\text{Energy} = \alpha C_{SWITCHED} V_{DD}^2 = \alpha (C_{IN} + C/f^2 + C/f + C)V_{DD}^2$
- Ανάλογα με την τιμή του  $C_{IN}$  πετυχαίνουμε και άλλη καθυστέρηση ενώ αλλάζει αντίστοιχα και η συνολική δυναμική ενέργεια που χρειάζεται να δαπανήσουμε

# Βελτιστοποίηση ενέργειας-καθυστέρησης

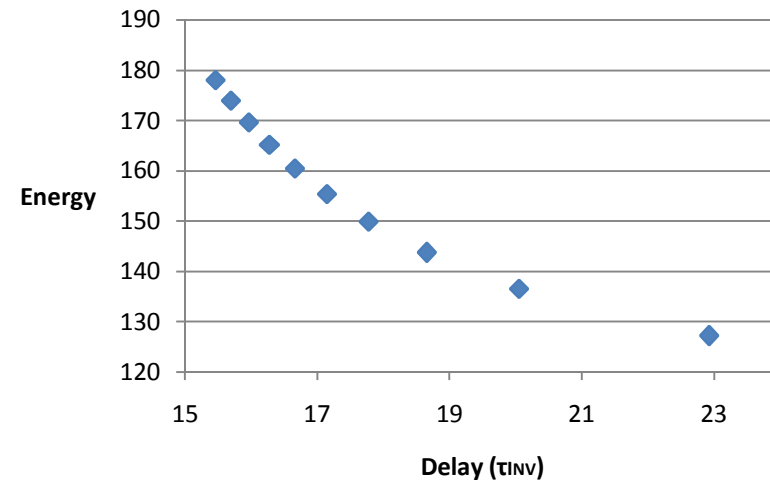
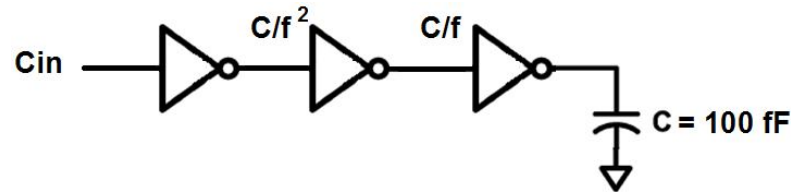


- Αν θεωρήσουμε πως  $C_{in} < \text{Limit}$  ( $=10 \text{ fF}$ ) ώστε να μην επιβραδύνουμε σημαντικά το προηγούμενο κύκλωμα μπορούμε να διαλέξουμε μια πληθώρα λύσεων
- Κάθε λύση έχει το δικό της βέλτιστο stage effort  $f$
- Για κάθε  $C_{in}$  (μεταβλητής προβλήματος βελτιστοποίησης) που διαλέγουμε πετυχαίνουμε την ελάχιστη δυνατή καθυστέρηση. Καθε επιλογή συνεπάγεται μια κατανάλωση ενέργειας
- Όσο αυξάνουμε το  $C_{in}$  η καθυστέρηση μειώνεται ενώ αυξάνεται η τιμή του  $C_{switched}$  όπως αναμενόταν





# Βέλτιστες καμπύλες ενέργειας/καθυστέρησης

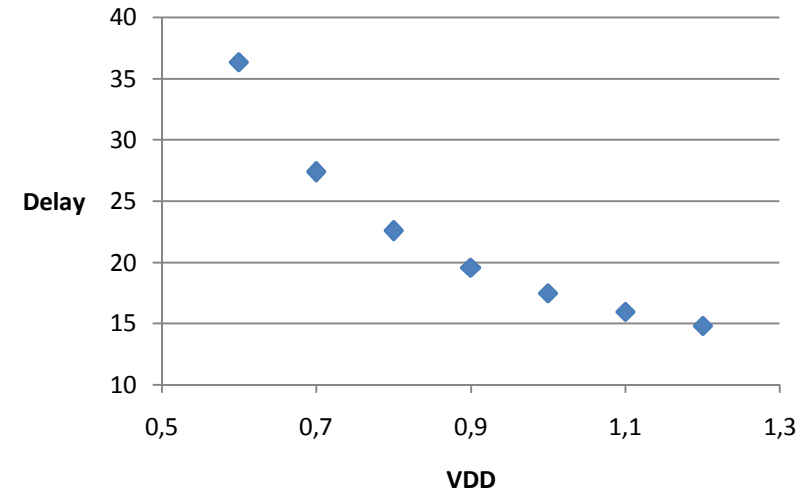


- Τοποθετούμε όλα τα σημεία που προέκυψαν από τη βελτιστοποίηση ως προς το  $C_{in}$  στο ίδιο διάγραμμα με άξονες την καθυστέρηση και την ενέργεια
- Όλα τα σημεία της καμπύλης είναι βέλτιστα ως προς την ενέργεια και την καθυστέρηση: Ο λόγος είναι ότι δε μπορούμε να μειώσουμε την καθυστέρηση περαιτέρω χωρίς να αύξησουμε την ενέργεια που δαπανούμε ενώ αντίθετα δεν μπορούμε να μειώσουμε την ενέργεια χωρίς να αυξήσουμε την καθυστέρηση
  - Ο λόγος είναι ότι χρησιμοποιούμε τη μέθοδο του Logical effort που εξασφαλίζει πως για κάθε  $C_{in}$  η λύση που διαλέγουμε πετυχαίνουμε τη βέλτιστη καθυστέρηση (Διαβάστε επιπλέον κείμενο για βέλτιστες καμπύλες ενέργειας - καθυστέρησης)
  - Αν θέλουμε μια καλύτερη καμπύλη (πιο αποδοτικά σημεία) πρέπει να εισάγουμε στο πρόβλημα μας και άλλες μεταβλητές βελτιστοποίησης όπως το πλήθος των σταδίων  $N$  και όχι μόνο το  $C_{in}$
- Αν αλλάξουμε  $C_{out}$  ή αριθμό από αντιστροφείς πρέπει να δημιουργήσουμε εκ νέου μια νέα καμπύλη



# Η συνεισφορά της τάσης τροφοδοσίας

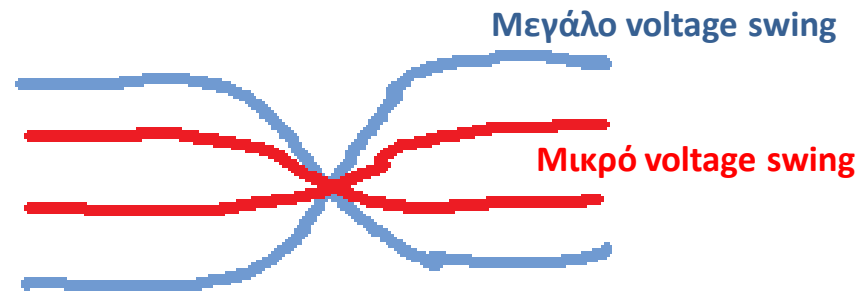
- Ανάλογη βελτιστοποίηση μπορούμε να επιτύχουμε αν μεταβάλλουμε την τάση τροφοδοσίας του κυκλώματος
- Με τη μείωση του VDD κερδίζουμε «τετραγωνικά» την απαιτούμενη ενέργεια αλλά χάνουμε επίσης σημαντικά στην καθυστέρηση των πυλών
- Δεν επαναυπολογίζουμε τα μεγέθη των τρανζίστορ απλά μεταβάλλουμε το διακόπτη που λέγεται τάση τροφοδοσίας
- Μπορούμε ουσιαστικά να επιτύχουμε την ελάχιστη καθυστέρηση για μια ζητούμενη ταχύτητα ή το αντίθετο
- Παράλληλα με την κατάλληλη επιλογή των μεγεθών των τρανζίστορ κατά τη φάση τη σχεδίασης μπορούμε να επιτύχουμε τη ζητούμενη καθυστέρηση ελαχιστοποιώντας το όσο το δυνατόν περισσότερο την απαιτούμενη ενέργεια πειράζοντας ταυτόχρονα την τάση VDD



$$Delay = Kconst \frac{C_L V_{DD}}{(V_{DD} - V_T)^b}$$

# Η συνεισφορά της τάσης τροφοδοσίας

- Σε άλλες περιπτώσεις η μείωση της τάσης VDD μειώνει ταυτόχρονα την καθυστέρηση και την κατανάλωση ισχύος
- Δυσκολεύει βέβαια τη σχεδίαση - απομακρυνόμαστε από την αμιγώς ψηφιακή σχεδίαση
- Οδήγηση μεγάλων διαδρόμων δεδομένων με μικρή διαφορά δυναμικού
- Παράδειγμα
  - 10pF bus, 64 bits
  - 500MHz clock
  - toggles at f/4 (125MHz)
  - V=3.3V
    - $E = 640\text{pF} \times 3.3^2 = 7\text{nJ}$
    - $P = 7\text{nJ} \times 125\text{MHz} = 0.875\text{W}$
  - at 100mV
    - $E = 0.2\text{nJ}$ ,  $P = 27\text{mW}$
- Η καθυστέρηση επίσης μπορεί να μειωθεί γιατί το λογικό-1 από το λογικό-0 απέχουν λιγότερο μεταξύ τους οπότε μπορούμε ταχύτερα να τα υπολογίσουμε
- Χρησιμοποιείται ευρέως στους διαδρόμους δεδομένων του συστήματος και στις διεπαφές των RAMs



## Γενικές στρατηγικές μείωσης της κατανάλωσης ισχύος

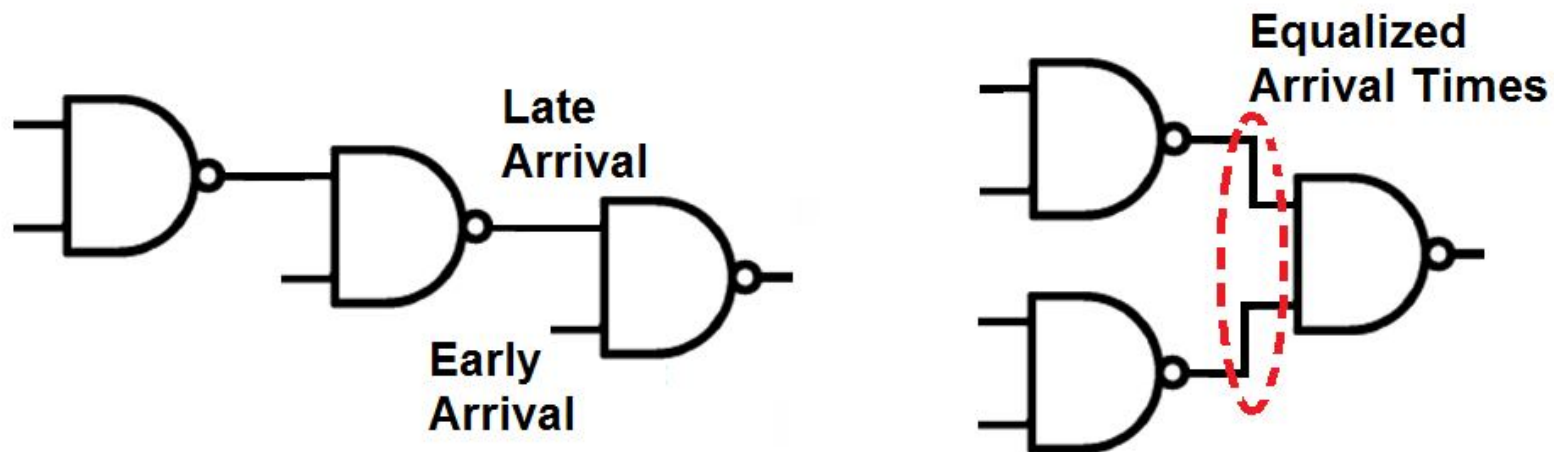
- Κλείσε τις ανενεργές περιοχές του κυκλώματος
- Μείωσε τον αριθμό των μεταβάσεων
- Παραλληλισμός για μείωση της κατανάλωσης και όχι της απόδοσης
- Επαναπροσδιορισμός του προβλήματος

# Κλειστές οι ανενεργές περιοχές του κυκλώματος

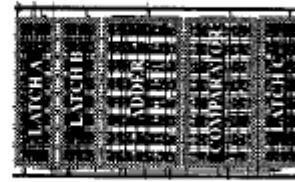
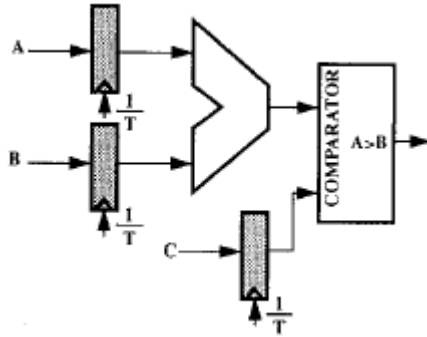
- Το μεγαλύτερο μέρος του κυκλώματος είναι ανενεργό την περισσότερη ώρα
- Χρησιμοποίησε μόνο εκείνα τα κομμάτια που απαιτούνται για το τρέχοντα υπολογισμό
  - Οι Floating point μονάδες δε χρειάζονται κατά την επεξεργασία κειμένου
  - Δε μιλάμε στο κινητό την ώρα που ακούμε (συνήθως)
- Η λογική CMOS δαπανά την περισσότερη ενέργεια όταν οι πύλες αλλάζουν κατάσταση
- Γι' αυτό πρέπει να σιγουρευτούμε πως οι ανενεργές μονάδες του κυκλώματος δεν παρουσιάζουν μεταβάσεις
- Σταμάτησε τα ρολόγια αυτών των μονάδων - 'Έτσι τα σήματα θα πάψουν ν' αλλάζουν (clock gating)
  - Η δέσμευση του ρολογιού θέλει προσοχή όσον αφορά το χρονισμό (εισάγει επιπλέον clock skew)
  - Επίσης θέλει προσοχή το αναβόσβημα του ρολογιού από ενεργό σε ανενεργό
- Στις νέες τεχνολογίες όπου τα ρεύματα διαρροής δημιουργούν προβλήματα απαιτείται επίσης το κλείδωμα της τροφοδοσίας των κυκλωμάτων (power gating)

# Αποφυγή glitches

- Όπως δείξαμε τα κυκλώματα παρουσιάζουν «περίεργες» μεταβάσεις κατά τον υπολογισμό
- Η αιτία είναι οι διαφορετικοί χρονικοί άφιξης των εισόδων μιας πύλης λόγω των διαφορετικών μονοπατιών από τα οποία διέρχονται τα σήματα
- Το πρόβλημα αυτό διορθώνεται εν μέρει ισορροπώντας την καθυστέρηση των διαφορετικών μονοπατιών
- Πολλές φορές αυτό βοηθάει στη μείωση της συνολικής καθυστέρησης του κυκλώματος!!



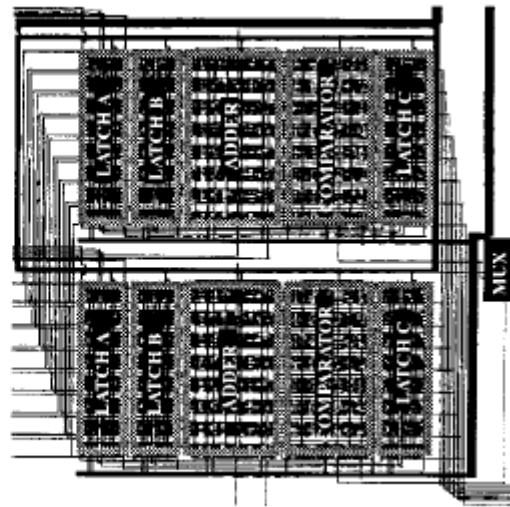
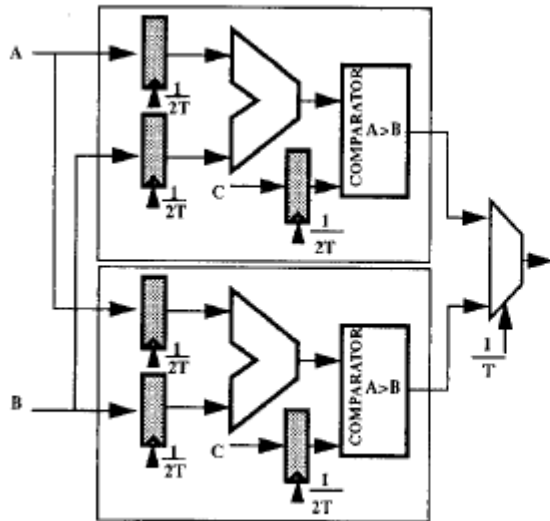
# Παράλληλος υπολογισμός



Area ■ 636 x 833 μ<sup>2</sup>

Αρχικός σχεδιασμός

$$P_{\text{ref}} = C_{\text{ref}} V_{\text{ref}}^2 f_{\text{ref}}$$



Area ■ 1476 x 1219 μ<sup>2</sup>

$$P_{\text{par}} = C_{\text{par}} V_{\text{par}}^2 f_{\text{par}} = (2.15C_{\text{ref}}) (0.58V_{\text{ref}})^2 \left(\frac{f_{\text{ref}}}{2}\right) \approx 0.36 P_{\text{ref}}$$

- Η χρήση του παράλληλου υπολογισμού πάντα μας επιτρέπει να κερδίζουμε σχεδόν για όλες τις κατηγορίες των κυκλωμάτων
- Βελτιώνεις το throughput τοποθετώντας παράλληλες μονάδες
- Αν θέλεις το ίδιο throughput με αρχικό σχεδιασμό μπορείς να κάνεις την παράλληλη εκδοχή αργότερη μειώνοντας το VDD και τη συχνότητα
- Παρόλη την επιβάρυνση από την επιπλέον καλωδίωση το κέρδος σε ισχύ είναι σημαντικό

# Επαναπροσδιορισμός του προβλήματος

- Αν μπορείς για μια δεδομένη εργασία να κάνεις λιγότερη δουλειά για το ίδιο αποτέλεσμα πάντα σου επιτρέπει να μειώσεις την κατανάλωση ισχύος
- Λιγότερη δουλειά σημαίνει πως η νέα λύση που έχεις επιλέξει απαιτεί λιγότερες λειτουργίες
- Για παράδειγμα θεωρήστε πως η λύση ενός προβλήματος απαιτεί 100 λειτουργίες - εντολές (instructions)
  - Energy = 100  $E_{inst}$
  - Delay = 100  $T_{inst}$
- Αν ένας άλλος αλγόριθμος απαιτεί 50 λειτουργίες τότε:
  - Energy =  $\frac{1}{2}$  Energy(old) και Delay =  $\frac{1}{2}$  Delay(old)
- Αν η προηγούμενη λύση ήταν επαρκής τότε
  - Μπορούμε να ανταλλάξουμε την επιπλέον απόδοση που έχουμε ώστε να μειώσουμε περαιτέρω την ενέργεια ανά λειτουργία μειώνοντας πχ το VDD
- Η κάθε λειτουργία από τη νέα λύση πρέπει να απαιτεί περίπου την ίδια ενέργεια - Δεν είμαι σίγουρος πως η αντικατάσταση 10 προσθέσεων με 5 πολλαπλασιασμούς είναι πιο αποδοτική λύση