Logical Effort:

Designing for Speed on the Back of an Envelope

David Harris

David_Harris@hmc.edu

Harvey Mudd College

Claremont, CA



Outline

Introduction
Delay in a Logic Gate
Multi-stage Logic Networks
Choosing the Best Number of Stages
Example
Asymmetric & Skewed Logic Gates
Circuit Families
Summary



Logical Effort

David Harris

Introduction

Chip designers face a bewildering array of choices.

- What is the best circuit topology for a function?
- How large should the transistors be?
- How many stages of logic give least delay?

?	?	?
(0	•
0	1	

Logical Effort is a method of answering these questions:

- Uses a very simple model of delay
- Back of the envelope calculations and tractable optimization
- Gives new names to old ideas to emphasize remarkable symmetries

Who cares about logical effort?

- Circuit designers waste too much time simulating and tweaking circuits
- High speed logic designers need to know where time is going in their logic
- CAD engineers need to understand circuits to build better tools



Example

Ben Bitdiddle is the memory designer for the Motoroil 68W86, an embedded processor for automotive applications. Help Ben design the decoder for a register file: $a<3:0>\overline{a}<3:0>|_{4}$ 32 bits

4:16 Decoder

16

Decoder specification:

- 16 word register file
- Each word is 32 bits wide
- Each bit presents a load of 3 unit-sized transistors
- True and complementary inputs of address bits *a*<3:0> are available
- Each input may drive 10 unit-sized transistors

Ben needs to decide:

- How many stages to use?
- How large should each gate be?
 - How fast can the decoder operate?

Register File



16 words

Outline				
 Introduction Delay in a Logic Gate Multi-stage Logic Networks Choosing the Best Number of Stages Example Asymmetric & Skewed Logic Gates Circuit Families Summary 				



Delay in a Logic Gate

Let us express delays in a process-independent unit:



 $\tau \approx 12 \text{ ps}$ in 0.18 µm technology

Effort delay again has two components:

$$= gh$$

electrical effort is sometimes called *"fanout"*

 Logical effort describes relative ability of gate topology to deliver current (defined to be 1 for an inverter)

Electrical effort is the ratio of output to input capacitance





Computing Logical Effort

- DEF: Logical effort is the ratio of the input capacitance of a gate to the input capacitance of an inverter delivering the same output current.
 - Measured from delay *vs.* fanout plots of simulated or measured gates
 - Or estimated, counting capacitance in units of transistor width:



A Catalog of Gates

Table 1: Logical effort of static CMOS gates

Gate type	Number of inputs					
Gale type	1	2	3	4	5	n
inverter	1					
NAND		4/3	5/3	6/3	7/3	(<i>n</i> +2)/3
NOR		5/3	7/3	9/3	11/3	(2 <i>n</i> +1)/3
multiplexer		2	2	2	2	2
XOR, XNOR		4	12	32		

Table 2: Parasitic delay of static CMOS gates

Gate type	Parasitic delay
inverter	P _{inv}
<i>n</i> -input NAND	np _{inv}
<i>n</i> -input NOR	np _{inv}
<i>n</i> -way multiplexer	2np _{inv}
2-input XOR, XNOR	4np _{inv}

 $p_{inv} \approx 1$

parasitic delays depend on diffusion capacitance





Example

Estimate the frequency of an *N*-stage ring oscillator:

∞—••• –⊳

- Logical Effort: g =
- Electrical Effort: h =
- Parasitic Delay: p =
- Stage Delay: d =
- Oscillator Frequency: F =



Example Estimate the delay of a fanout-of-4 (FO4) inverter: l**←**d → l Logical Effort: *g* = Electrical Effort: h =Parasitic Delay: *p* = Stage Delay: d =



Outline				
Introduction Delay in a Logic Gate <u>Multi-stage Logic Networks</u> Choosing the Best Number of Stages Example Asymmetric & Skewed Logic Gates Circuit Families Summary				
Introduction Delay in a Logic Gate <u>Multi-stage Logic Networks</u> Choosing the Best Number of Stages Example Asymmetric & Skewed Logic Gates Circuit Families Summary				



Multi-stage Logic Networks

Logical effort extends to multi-stage networks:

Logical Effort



David Harris

Page 13 of 56

Branching Effort

No! Consider circuits that branch:

$$G = H = H = GH^{2}$$

$$GH = H_{1} = H_{2} = GH^{2}$$





Delay in Multi-stage Networks

We can now compute the delay of a multi-stage network:

- D Path Effort Delay: $D_F = \sum f_i$
- Path Parasitic Delay:

$$P = \sum p_i$$

D Path Delay: $D = \sum d_i = D_F + P$

We can prove that delay is minimized when each stage bears the same effort:

$$\hat{f} = g_i h_i = F^{1/N}$$

Therefore, the minimum delay of an *N*-stage path is:

$$NF^{1/N} + P$$

This is a key result of logical effort. Lowest possible path delay can be found without even calculating the sizes of each gate in the path.



Determining Gate Sizes

Gate sizes can be found by starting at the end of the path and working backward.

At each gate, apply the capacitance transformation:

$$C_{in_i} = \frac{C_{out_i} \bullet g_i}{\hat{f}}$$

Check your work by verifying that the input capacitance specification is satisfied at the beginning of the path.



Page 16 of 56

Example

Select gate sizes *y* and *z* to minimize delay from *A* to *B*

- Logical Effort: G =
- Electrical Effort: H =
- Branching Effort: B =
- Path Effort: F =
- Best Stage Effort: f =
- Delay: D =

Work backward for sizes:

$$Z =$$

y =



Logical Effort

David Harris



Outline Introduction Delay in a Logic Gate Multi-stage Logic Networks **Choosing the Best Number of Stages** Example Asymmetric & Skewed Logic Gates **Circuit Families** Summary







Best Number of Stages (continued)

 $p_{inv} + \rho(1 - ln\rho) = 0$ has no closed form solution.

Neglecting parasitics (*i.e.* $p_{inv} = 0$), we get the familiar result that $\rho = 2.718$ (*e*) For $p_{inv} = 1$, we can solve numerically to obtain $\rho = 3.59$ How sensitive is the delay to using exactly the best number of stages?



Outline

- Introduction
- Delay in a Logic Gate
- Multi-stage Logic Networks
- Choosing the Best Number of Stages
 - **Example**
- Asymmetric & Skewed Logic Gates
 - **Circuit Families**
 - **Summary**





Example

Let's revisit Ben Bitdiddle's decoder problem using logical effort:

Decoder specification:

- 16 word register file
- Each word is 32 bits wide
- Each bit presents a load of 3 unit-sized transistors
- True and complementary inputs of address bits *a*<3:0> are available
- Each input may drive 10 unit-sized transistors

Ben needs to decide:

- How many stages to use?
- How large should each gate be?
- How fast can the decoder operate?



Example: Number of Stages

How many stages should Ben use?

Effort of decoders is dominated by electrical and branching portions

- Electrical Effort: H =
- **D** Branching Effort: B =

If we neglect logical effort (assume G = 1),

 \square Path Effort: F =

Remember that the best stage effort is about $\rho = 4$

Hence, the best number of stages is: N =



Example: Gate Sizes & Delay





Logical Effort

David Harris



Example: Alternative Decoders

Design	Stages	G	Р	D
NAND4; INV	2	2	5	29.8
INV; NAND4; INV	3	2	6	22.1
INV; NAND4; INV; INV	4	2	7	21.1
NAND2; INV; NAND2; INV	4	16/9	6	19.7
INV; NAND2; INV; NAND2; INV	5	16/9	7	20.4
NAND2; INV; NAND2; INV; INV; INV	6	16/9	8	21.6
INV; NAND2; INV; NAND2; INV; INV; INV	7	16/9	9	23.1
NAND2; INV; NAND2; INV; INV; INV; INV; INV	8	16/9	10	24.8

Table 3: Comparison of Decoder Designs

We underestimated the best number of stages by neglecting the logical effort.

Logical effort facilitates comparing different designs before selecting sizes

- Using more stages also reduces G and P by using multiple 2-input gates
- Our design was about 10% slower than the best





Outline

Introduction

- Delay in a Logic Gate
- Multi-stage Logic Networks
- Choosing the Best Number of Stages
- Example
 - Asymmetric & Skewed Logic Gates
 - Circuit Families
- □ Summary



Asymmetric Gates

Asymmetric logic gates favor one input over another.

Example: suppose input A of a NAND gate is most critical.

- Select sizes so pullup and pulldown still match unit inverter
- Place critical input closest to output



Logical Effort on input A:

Total Logical Effort:



Symmetry Factor

In general, consider gates with arbitrary symmetry factor s:

s = 1/2 in symmetric gate with equal sizes

 \Box s = 1/4 in previous example

Logical effort of inputs:

$$g_A = \frac{\frac{1}{1-s}+2}{3}$$
 $g_B = \frac{\frac{1}{s}+2}{3}$ $g_{tot} = \frac{\frac{1}{s(1-s)}+4}{3}$

Critical input approaches logical effort of inverter = 1 for small s But total logical effort is higher for asymmetric gates

2

а

1/(1-s)

1/s

2

-X





HI- and LO-Skewed Gates

DEF: Logical effort of a skewed gate for a particular transition is the ratio of the input capacitance of that gate to the input capacitance of an unskewed inverter delivering the same output current for the same transition.

Skew gates by reducing size of noncritical transistors.

- ☐ HI-Skewed gates favor rising outputs by downsizing NMOS transistors
- LO-Skewed gates favor falling outputs by downsizing PMOS transistors
- Logical effort is smaller for the favored input due to lower input capacitance
- Logical effort is larger for the other input



Catalog of Skewed Gates



Logical Effort

David Harris

Page 32 of 56



Outline

- Introduction
- Delay in a Logic Gate
- Multi-stage Logic Networks
- Choosing the Best Number of Stages
- Example
 - Asymmetric & Skewed Logic Gates
 - **<u>Circuit Families</u>**

] Summary



Pseudo-NMOS

Pseudo-NMOS gates replace fat PMOS pullups on inputs with a resistive pullup.

- **Control** Resistive pullup must be much weaker than pulldown stack (e.g. 4x)
- Reduces logical effort because inputs must only drive the NMOS transistors
- However, NMOS current reduced by contention with pullup
- Unequal rising and falling efforts
- Quiescent power dissipation when output is low

Example: Pseudo-NMOS inverter

- Logical Effort for falling (down) output: $g_d =$
- Logical Effort for rising (up) output:
 - Average Logical Effort:

 $g_{11} =$ $g_{avg} = (g_u + g_d)/2$





Pseudo-NMOS Gates



Tradeoffs exist between power and effort by varying P/N ratio.



David Harris



Dynamic Logic



- Reduces logical effort because inputs must only drive the NMOS transistors
- Eliminates pseudo-NMOS contention current and power dissipation
- Only the falling ("evaluation") delay is critical
 - Downsize noncritical precharge transistors to reduce clock load and power

Example: Footless dynamic inverter

Logical Effort for falling (down) output: $g_d =$



Robust gates may require keepers and clocked pulldown transistors ("feet").

Feet prevent contention during precharge but increase logical effort Weak keepers prevent floating output at cost of slight contention during eval

Logical Effort

David Harris



Dynamic Gates



Domino Gates

Dynamic gates require monotonically rising inputs.

- However, they generate monotonically falling outputs
- Alternate dynamic gates with HI-skew inverting static gates
- Dynamic / static pair is called a domino gate

Example: Domino Buffer

- Constraints: maximum input capacitance = 3, load = 54
- Logical Effort: G =
- Branching Effort: B =
- Electrical Effort: H =
- Path Effort: F =
- Stage Effort: f =
- HI-Skew Inverter: size =
- Transistor Sizes: n = p =



Logical Effort

David Harris



Comparison of Circuit Families

Assumptions:

PMOS transistors have half the drive of NMOS transistors

3 Skewed gates downsize noncritical transistors by factor of two

D Pseudo-NMOS gates have 1/4 strength pullups

Circuit Style	Inverter g		n-input NAND g		n-input NOR g	
Circuit Style	9 _u	9 _d	9 _u	9 _d	9 _u	9 _d
Static CMOS	1		(n+2)/3		(2n+1)/3	
HI-Skew	5/6	5/3	(n/2+2)/3	(n+4)/3	(2n+.5)/3	(4n+1)/3
LO-Skew	4/3	2/3	2(n+1)/3	(n+1)/3	2(n+1)/3	(n+1)/3
Pseudo-NMOS	4/3	4/9	4n/3	4n/9	4/3	4/9
Footed Dynamic	2/	/3	(n+	1)/3	2/	/3
Footless Dynamic 1/3		n/3		1/3		

Table 4: Summary of Logical Efforts

Adjust these numbers as you change your assumptions.





Outline

- Introduction
- Delay in a Logic Gate
- Multi-stage Logic Networks
- Choosing the Best Number of Stages
- Example
 - Asymmetric & Skewed Logic Gates
 - Circuit Families
 - <u>Summary</u>



Summary

Table 5: Key Definitions of Logical Effort

Term	Stage expression	Path expression
Logical effort	$oldsymbol{g}$ (seeTable 1)	$G = \prod g_i$
Electrical effort	$h = \frac{C_{out}}{C_{in}}$	$H = rac{C_{out (path)}}{C_{in (path)}}$
Branching effort	n/a	$B = \prod b_i$
Effort	f = gh	F = GBH
Effort delay	f	$D_F = \sum f_i$
Number of stages	1	N
Parasitic delay	$oldsymbol{ ho}$ (seeTable 2)	$P = \sum p_i$
Delay	d = f + p	$D = D_F + P$

Logical Effort

David Harris



Method of Logical Effort

Logical effort helps you find the best number of stages, the best size of each gate, and the minimum delay of a circuit with the following procedure:

Compute the path effort: F = GBH

I Estimate the best number of stages:

Estimate the minimum delay:

3 Sketch your path using the number of stages computed above

- Compute the stage effort: $\hat{f} = F^{1/N}$
 - Starting at the end, work backward to find transistor sizes:

$$C_{in_i} = \frac{C_{out_i} \bullet g_i}{\hat{f}}$$

 $N \approx \log_{a} F$

 $D - \hat{N}F^{1/\hat{N}} + P$



Limitations of Logical Effort Logical effort is not a panacea. Some limitations include: Chicken & egg problem how to estimate G and best number of stages before the path is designed Simplistic delay model neglects effects of input slopes Interconnect iteration required in designs with branching and non-negligible wire C or RC same convergence difficulties as in synthesis / placement problem . . Maximum speed only optimizes circuits for speed, not area or power under a fixed speed constraint



Conclusion

Logical effort is a useful concept for thinking about delay in circuits:

- ☐ Facilitates comparison of different circuit topologies
- Easily select gate sizes for minimum delay
- Circuits are fastest when effort delays of each stage are equal and about 4
- □ Path delay is insensitive to modest deviations from optimal sizes
- Logic gates can be skewed to favor one input or edge at the cost of another
- Logical effort can be applied to domino, pseudo-NMOS, and other logic families

Logical effort provides a language for engineers to discuss why circuits are fast.

Like any language, requires practice to master

A book on Logical Effort is available from Morgan Kaufmann Publishers

http://www.mkp.com/Logical_Effort

Discusses P/N ratios, gate characterization, pass gate logic, forks, wires, etc.





Example Estimate the frequency of an *N*-stage ring oscillator: - ••• -|> $g \equiv 1$ Logical Effort: Electrical Effort: $h = \frac{C_{out}}{C_{in}} = 1$ p = p_{inv}≈ 1 Parasitic Delay: d = gh + p = 2Stage Delay: A 31 stage ring Oscillator Frequency: $F = \frac{1}{2Nd_{abs}} = \frac{1}{4N\tau}$ oscillator in a $0.18 \,\mu m$ process oscillates at about 670 MHz.



Example Estimate the delay of a fanout-of-4 (FO4) inverter: $g \equiv 1$ Logical Effort: $h = \frac{C_{out}}{C_{in}} = 4$ The FO4 inverter Electrical Effort: delay is a useful metric to characterize process performance. p = p_{inv}≈ 1 Parasitic Delay: 1 FO4 delay = 5τ d = gh + p = 5Stage Delay: This is about 60 ps in a 0.18 μ m process. **Logical Effort David Harris** Page 47 of 56

Branching Effort

No! Consider circuits that branch:



Introduce new kind of effort to account for branching within a network:



Example

Select gate sizes *y* and *z* to minimize delay from A to B

Logical Effort:

Electrical Effort:

Path Effort:

Best Stage Effort: $\hat{f} = F^{1/3} \approx 5$

Delay:

 $G = (4/3)^3$ $H = \frac{C_{out}}{C_{in}} = 9$ Branching Effort: $B = 2 \bullet 3 = 6$

F = GHB = 128

 $D = 3 \bullet 5 + 3 \bullet 2 = 21$



Work backward for sizes:

$$z = \frac{9C \cdot (4/3)}{5} = 2.4C$$
$$y = \frac{3z \cdot (4/3)}{5} = 1.92C$$



Example: Number of Stages

How many stages should Ben use?

Effort of decoders is dominated by electrical and branching portions

- Electrical Effort: $H = \frac{32 \cdot 3}{10} = 9.6$. .

Branching Effort: B = 8 because each address input controls half the outputs

If we neglect logical effort,

Path Effort: $F = GBH = 8 \bullet 9.6 = 76.8$

Remember that the best stage effort is about $\rho = 4$

- Hence, the best number of stages is: $N = \log_{4} 76.8 = 3.1$
 - Let's try a 3-stage design

Logical Effort

David Harris



Example: Gate Sizes & Delay

Lets try a 3-stage design using 16 4-input NAND gates with $G = 1 \bullet 2 \bullet 1 = 2$



Logical Effort

Page 51 of 56





Skewed Gates

Skewed gates favor one edge over the other.

Example: suppose rising output of inverter is most important.

Downsize noncritical NMOS transistor to reduce total input capacitance





Pseudo-NMOS

Pseudo-NMOS gates replace fat PMOS pullups on inputs with a resistive pullup.

- Resistive pullup must be much weaker than pulldown stack (e.g. 4x)
- Reduces logical effort because inputs must only drive the NMOS transistors
- However, NMOS current reduced by contention with pullup
- Unequal rising and falling efforts
 - Logical effort can be applied to domino, pseudo-NMOS, and other logic families

Example: Pseudo-NMOS inverter

- Logical Effort for falling (down) output:
- Logical Effort for rising (up) output:
 - Average Logical Effort:

 $g_{d} = 4/9$ $g_u = 4/3$ $g_{avg} = (g_u + g_d)/2 = 8/9$



Page 54 of 56

Dynamic Logic





Domino Gates

Dynamic gates require monotonically rising inputs.

- However, they generate monotonically falling outputs
- Alternate dynamic gates with HI-skew inverting static gates
- Dynamic / static pair is called a domino gate

Example: Domino Buffer

Constraints: maximum input capacitance = 3, load = 54

Logical Effort: G = (1/3) * (5/6) = 5/18

Branching Effort: B = 1

Electrical Effort: H = 54/3 = 18

- Path Effort: F = (5/18) * 1 * 18 = 5
- **3** Stage Effort: $f = \sqrt{5} = 2.2$

HI-Skew Inverter: size =54 * (5/6) / 2.2 = 20

Transistor Sizes: n = 4 p = 16



