Computer-based Project in VLSI Design
Final Report

Tina Wen
College: Trinity
Email: tw311@cam.ac.uk

June 6, 2007

Introduction

After a detailed top level design of the frequency synthesizer, we go into component level of the chip design and place the IC layout. Capacitance involved in the layout of the components and PCB is calculated and simulated. An estimation of delay is made. The ring oscillator lab is carried out, and the report is attached in the Appendix.

`nor2` stick diagram

A simple stick diagram for the `nor2` gate is constructed and shown in Graph 1 in the Appendix. The positions of the input ports A and B are shaded in green, and the output port Y is shaded in yellow. The position of the positive supply rail is shaded in pink, and its width is measured to be $V_{DD} = 10.5\mu m$. The position of ground is shaded in light blue and its width is measured to be $V_{SS} = 10.5\mu m$. The separation of the supply rails is $70\mu m$. The magnitude of the power supply limits the minimum width of the metal traces, and we want this width to be as small as possible to save space. Gate, source and drain of the four transistors are labeled on the graph. Input and output ports are marked on the graph in orange and brown, respectively. A completed stick diagram is shown in Graph 2 in the Appendix.

Layout-Versus-Schematic using *ICtrace*

Layout versus Schematic of the `nor2t` gate is analyzed using *ICtrace*. LVS runs successfully. The report can be found in the Appendix. There is a minor discrepancy between schematic/layout entities. In the schematic, we have V_{DD} going to the P-MOSFET that has input B first, then to the P-MOSFET that has input A. However, on the layout, we have V_{DD} going to the P-MOSFET that has input A first, then to the P-MOSFET that has input B. This small MOSFET swapping discrepancy would not affect performance of the gate though.

Analysis of parasitic capacitances

Given interconnect specific capacitances provided by Mietec CMOS24 technology, parasitic capacitances due to interconnect can be analyzed and calculated. Table 1 in the Appendix summarizes the result. Capacitances associated with ports are calculated and shown in Table 2 in the Appendix. Areas and perimeters of transistors are also measured and listed in Table 3 in the Appendix. `Polysilicon` is good to be used for conducting gate material, but it introduces big capacitance. To make a high speed chip, we want to reduce capacitance. `Metal2` has the smallest capacitance and is better to be used as contact pins.

Two `nor2` gates are connected in cascade shown in Graph 3 in the Appendix. Assuming the width between the edges of the two `nor2` gates is $3\mu m$, and the width of `metal1` is

$4\mu m$, the area of `metal1` is calculated as: $(51.5 + 6 \times 3) \times 4 = 278\mu m^2$, demonstrated in Graph 4 in the Appendix. The additional interconnect capacitance is $2.3 \times 10^{-2} \times 278 = 6.4fF$. Therefore, the `total capacitance = total input capacitance of next stage + total output capacitance of previous stage + interconnecting capacitance = capacitance from polyA and polyB + capacitance from metalY and polyY + capacitance from 3 metal contacts + additional interconnect capacitance` $= 13.8 + 14.3 + 18.9 + 11.4 + 2.5 + 6.4 = 67.3fF$.

Simulation of the schematic design using *AccuSim*

DC Sweep Analysis

Different forces are put on inputs of the `nor2` gate, and the simulated output is observed and analyzed. Holding input A constant and varying input B linearly produce a plot in Graph 5 in the Appendix. The two cursors mark the transition region. The maximum linear amplification of the gate with respect to input B is $|\frac{\partial V_Y}{\partial V_B}| = \frac{3.6944-1.6185}{2.69-2.6} = 23.07$. The input voltage level is in the range of $[2.6V - 2.69V]$ when the maximum linear gain is achieved. Due to the resolution of the measuring steps, the result may not be satisfactory enough. We can reduce the size of input increment steps to get better resolution. The input switching level is $2.65V$ at which $V_Y = V_B$.

Holding input B constant and varying input A linearly produce a similar plot. The input A transfer function gain is $|\frac{\partial V_Y}{\partial V_A}| = \frac{3.3704-1.3887}{2.59-2.49} = 19.8$. Input voltage level is in the range of $[2.49V - 2.59V]$ when the maximum linear gain is achieved. The switching level for input A is $2.53V$.
$NM_H = V_{OH} - V_{IH} = 4.6944 - 3.15 = 1.54V$
$NM_L = V_{IL} - V_{OL} = 1.97 - 284.61 \times 10^{-3} = 1.69V$
There is no way that we can vary two inputs at the same time in *Accusim*. Instead, we can vary one input and keep the other fixed, and take the measurement. Then do the same thing for the other input. Finally, we can use *MATLAB* to make a 3-D plot to see the effect of varying two inputs at the same time.

Transient response with internal circuit parasitics

According to the calculation in Table 2, the corresponding capacitances are added between input/output ports and ground as shown in Graph 6 in the Appendix. With input A being a pulse with delay $10ns$ and width $50ns$, the output rise and fall times is measured and recorded in Table 4 in the Appendix. In this setup, port A has capacitance $14.6pF$, port B has capacitance $15.1pF$, port Y has capacitance $31.1pF$.

Transient response with internal/external parasitics

As the total capacitive load driven by the original gate being $67.3pF$ calculated above, the simulation is run again with this external capacitance. The result is shown in Table

5 in the Appendix. All the rise/fall and delay times is increased in this table compared to Table 2 due to the fact that the capacitive loading on output Y is bigger in this case. A bigger capacitance leads to longer time to charge/discharge.

Transient response with reduced supply voltage

Having the supply voltage set at $3V$ and external capacitance fixed at $67.3pF$, the simulation is run and result is shown in Table 6 in the Appendix. All the rise/fall and delay times is increased in this table compared to Table 5. A smaller supply voltage leads to a smaller charging/discharging current, and thus leads to a longer rise/fall and delay times. Given the value in Table 10 measured in the lab ring oscillator, the gate delay with supply at $3V$ is $\frac{\frac{1}{1910k}}{115\times 2} = 2.28ns$. This measured value is comparable to our simulation result in Table 6.

Transient response with doubled $\frac{W}{L}$ ratio

The width of two P-MOSFETS is changed from $37\mu m$ to $74\mu m$, and the length remains as $3\mu m$. The $\frac{W}{L}$ ratio is essentially doubled. With the power supply set back to $5V$, the simulation is run and result is recorded in Table 7 in the Appendix. Compare the numbers with Table 5 we see that the output rise time is decreased. As the aspect ratio of P-MOSFET is doubled, the conductance of P-MOSFET is doubled, so the output rising time reduces. However, output falling time is not affected, because it only depends on N-MOSFETs, whose aspect ratio remains the same. In fact, as the aspect ratio of P-MOSFET increases, the capacitance associated with the MOSFET should increase as well, which should increase rise/fall and delay times. These two effects cancel out each other to a certain extent. However, in this case, the change of MOSFET capacitance is not modeled in the simulation.

Use results from *AccuSim* to model the behavior of design

According to our calculation above, the rise and fall times of NOR gate was not modeled correctly before. Delay time should be read using the values in Table 5, because it takes into all the relevant capacitance into account. T_{rise} is the delay time for the output to rise. We change this value to $0.67ns$, which is the output rise time as both inputs fall. T_{fall} is the delay time for the output to fall. We change this value to $0.35ns$, which is the output fall time as both inputs rise. This is the correct modeling for most NOR gates in our ring oscillator. However, this is not the case for 3 out of all 29 NOR gates. For the 2 NOR gates right before `out1` and `out2`, the delay time should be longer due to additional output capacitive loading. Also for the first NOR gate with an input `ring_enb`, the delay time should be read as the rise/fall time as one input falls/rises rather than both inputs. Therefore, the delay time should be smaller for this single gate. Since 3 is a small number compared to 29, we can approximate delay time this way. Simulation is

run again and result is shown in Graph 7 in the Appendix. Timing specifications of NOR gate is measured again and result is shown in Table 8. As can be observed, all the delay time is smaller compared to the values we got from week 1. This is due to the reduction of our capacitance modeling. Smaller capacitance leads to shorter delay time.

Creation of design viewpoint and library

The new design viewpoint is set up successfully and the result is shown in the Appendix. A list of components are added to the ring oscillator library also shown in the Appendix.

Creating a flattened layout

Routing is completed automatically. The placement and routing of the automatic result is relatively efficient. However, there are some cases that shorter paths can be taken. For example, A, B and Y are all connected vertically across the device. They could be all connected across horizontally to save space and increase efficiency. At this stage, the dimension of the chip is noted to be $[2801.0, 2608.5]$. There are 309 via objects and 118 nets generated by this operation.

Then we attempt to ask the software to reduce the number of via. The operation from `metal1` to `metal2` reduces the via number to 219. A further operation from `metal2` to `metal1` reduces the via number to 218.

Finally, we ask the software to compact the design. After performing the compact command in the `down` direction, the dimension of the chip does not seem to decrease. It remains to be $[2801.0, 2608.5]$. After performing the compact command in the `right` direction, the dimension of the chip is reduced to $[2602.5, 2600]$. This is indeed better than before.

Both black and white and colored prints are generated shown in Graph 8 and 9 in the Appendix.

Conclusion

Semi-custom design is investigated in this final part of the project. Placement and floor planning of gates and pads, automatic routing and optimization are performed. Capacitance is estimated and the estimation is used to run simulation and compared with experimental values. This project gave me an insight of how chips are designed and manufactured in industry. It also taught me important debugging skills that are essential for an electrical engineer.

Appendix

Ring oscillator lab

Introduction

The gate delay of an inverter is measured in several different ways. The relative accuracy of each is studied and compared.

Testing the single device

After testing that all the pins are connected correctly, the chip with serial number $21-13$ is put in Box A. A square wave with amplitude about $3V$, frequency about $1MHz$ is fed into `INPUT2`. `OUTPUT5` and `OUTPUT2` are seen on the oscilloscope screen, and sketched in Graph 1.1 in the Appendix. Clearly the relationship between `OUTPUT5` and `INPUT2` is $\texttt{OUTPUT5} = \overline{\texttt{INPUT2}}$. The single NOR gate inverts the input signal as the other input is held low. The gate delay can be estimated by reading from the oscilloscope screen. The delay is estimated to be $0.05\mu s$. The reading error of this direct measuring method can be very big, and the estimation can be quite far from the actual value.

The input and output pads are both inverting, so the $0.05\mu s$ is the delay for two inverters plus a NOR gate, the delay for 3 inverters. The load for the output inverter could be very big due to wire and scope probe loading. Therefore, the measured delay should be greater than the actual delay.

Testing ring-115

A square wave of frequency $100kHz$ and a duty cycle of 50% is fed to `INPUT2` of ring-115. `INPUT2` and `OUTPUT10` are observed on the oscilloscope. The waveforms are sketched in Graph 1.2 in the Appendix.

When `INPUT2` is high, output of the first NOR gate is always low. Therefore, the ring cannot oscillate. When `INPUT2` is low, output of the first NOR gate inverts the signal transmitted from the previous stage, and the ring oscillates. The ring oscillates 10 times per time when `INPUT2` is low. This means the frequency of the ring oscillator is 20 times as big as the square wave. Therefore, the ring oscillator frequency is $20 \times 100kHz = 2MHz$.

Determining the gate delay by measurements on ring-115 and ring-113

From last section, period of oscillation of the ring oscillator is found to be $\frac{1}{20\times100} = 0.5\times10^{-6}s$. The accuracy of this frequency measurement is limited by minimum time scale of the oscilloscope. A more accurate measurement can be done by using the frequency counter. The period is measured to be $0.54\mu s$. Each gate is flipped twice during one cycle of oscillation. Therefore, the gate delay is: $\frac{0.54\times10^{-6}}{115\times2} = 2.35 \times 10^{-9}s$.

Another way is used to estimate the gate delay by measuring the delay of a string of gates, and then dividing the delay by the number of gates in between. There are 37 gates

connected from OUTPUT6 to OUTPUT7. The measured delay from OUTPUT6 to OUTPUT7 is $0.1\mu s$. Therefore, the delay of one gate is $\frac{0.1\mu s}{37} = 2.7 \times 10^{-9} s$. Note that the assumption is that OUTPUT6 comes before OUTPUT7. If we look at OUTPUT6 to be after OUTPUT7. There are 76 gates in between. The delay time from OUTPUT7 to OUTPUT6 is $0.177\mu s$. Therefore, the delay of one gate is $\frac{0.177\times10^{-6}}{76} = 2.33 \times 10^{-9} s$. Either way, the result is pretty close to the gate delay measured by getting the period of the ring oscillator. This method is good in two ways. First, by measuring a large number of gates, it offsets minor manufacturing differences between each gate, and gives us a more accurate delay time for this particular kind of gate. Second, we do not have to worry about output loading. The oscillator probe as a load, in this case, only introduces a phase shift, but not a change of period.

Measure the effect of varying the power supply voltage

By setting PIN 3 and 4 high, and PIN 2 low, ring-113 free-runs. The minimum supply voltage required to establish full-amplitude oscillations is found to be $0.65V$, corresponding to the lowest frequency being $\frac{1}{5.5ms} = 181.8Hz$. When the supply voltage is below $0.65V$, the gate threshold of MOSFETS cannot be reached, and thus there is no full-amplitude oscillation.

The frequency of the free-running ring-113 is measured using the frequency counter as the power supply voltage is varied. The result is shown in Table 9, and plotted in Figure 1 in the Appendix.

The delay time of each MOSFET can be approximated by

$$\tau = \frac{3C}{G} = \frac{3C}{\mu C_{ox} V_{dd} \frac{W}{L}} \infty \frac{1}{V_{dd}} \tag{1}$$

Therefore, as the supply voltage increases, the delay τ of each gate decreases linearly, and thus oscillation frequency increases. However, when the supply voltage is too low, MOSFETs are just turned on, so the relationship between the supply voltage and the frequency is not quite linear. In the case when the supply voltage is very high, the devices consume a great deal of current. This increases heat dissipation, and causes devices to diverge from normal operation. Therefore, the supply voltage versus frequency curve diverges from linear at high supply voltages.

Performance comparison with different transistors in the ring

The chip with serial number $28-45$ has p-channel transistors three times as wide as those in the chip with serial number $21-13$ used in previous sections. The performance of the two chips are compared and the result is shown in Table 10.

From Equation (1), we know that as the width W of the transistor increases, the delay time should decrease. However, the delay does not decrease 3 times as much. This is due to two reasons. First, only the p-channel transistors are 3 times wider in the $28-45$ chip, but the n-channel transistors are the same dimension. This decreases the rise time only, but does not affect the fall time. The other reason is that the capacitance of the device increases as the width increases, so the loading capacitance of each stage increases, and

this increases the delay time. When V_{dd} is $5V$, the improvement seen from Equation (1) becomes less noticeable, and the delay is increased because of the two reasons above, thus the device actually runs slower. The performance of $28 - 45$ becomes worse.

Simulation of the device performance

Input A linearly transitions to high, stays high for about $2ns$, and then transitions linearly back done to low. The signal is passed through a number of inverting NOR gates, leading to the arise of signals B-E. Signal E is produced after signal C passing through two inverting gates, and thus they look similar. However, they are different in detail from waveform A because *Accusim* takes into account of the MOSFET channel resistances and all parasitic capacitances to substrate. The rise time for C and E is longer due to the RC lowpass filter effect. The ringing effect on signal C and E is due to inductance of the circuit.

The delay time can be measured by observing the delay between signal A and E. The delay is $2.2ns$. There are 4 gates in between A and E. Therefore, each gate delay is: $\frac{2.2ns}{4} = 0.55ns$.

Having the power supply set to $5V$, the measured period of oscillation is $\frac{1}{3720k} = 268.8ns$. Therefore, each gate delay is: $\frac{268.8\times10^{-9}}{115\times2} = 1.17ns$.

The simulated value is less than half of the experimental value. This is due to the fact that *AccuSim* does not model resistances like polysilicon lines interconnecting the devices. Therefore, the RC constant is modeled smaller than practice, resulting in a shorter delay.

Stroboscopic pulse generator

The output from the sense gate will be high only if both the signals at B and E are low. Signals B and E can be both high for a little bit per period due to delay. Therefore, signal F can be seen as a pulse generator, creating a pulse each period.

The sense gate spans 5 gates of the ring. When only ring-115 is running, there is a pulse generated per period due to the delay introduced by 5 gates. When both ring-113 and ring-115 are running freely, `OUTPUT9` is only high when both sense gates for ring-113 and ring-115 are high for the same period of time (the sense gates are inverted before connected to a NOR gate). Since the number of inverters for the two rings is different, the oscillation frequencies are different. The frequency of `OUTPUT9` goes high should be less than or equal to the ring-115 oscillation frequency. It corresponds to the overlapping of the sensing gates high signals from the two rings.

Having the supply voltage set at $0.55V$, it is increased until the circuit just no long works. This is at $1V$.

The supply voltage is set to be $0.64V$ to start, and then increased gradually. As the supply voltage increases, the width of the pulse decreases. The pulses disappear as the supply voltage reaches $0.9V$.

OUTPUT8 works as expected. Its frequency is twice the OUTPUT9 frequency. Over the range $0.60V$ to $0.80V$, OUTPUT8 works as expected. Outside this range the rise or fall of the input edge may be too slow or too fast to clock the divide-by-2 stage.

We can see pulses coming from OUTPUT10, having a period of $3ms$. Therefore, the gate delay is $\frac{3\times10^{-3}}{2\times115} = 13\times10^{-6}s$. The pulse width is $0.26ms$, and there are 5 gates being spanned over, so the gate delay is: $\frac{0.26\times10^{-3}}{5} = 52\mu s$. The delay is 4 times as much as the actual value. This could be due to loading.

Conclusion

Gate delay of an inverter can be measured by direct measuring method, or by measuring the period of a ring oscillator and then dividing it by twice the number of gates, or by measuring the width or period of the overlapping pulse. The most efficient and accurate way is from measuring the period of a ring. A frequency counter is used in this experiment. It is seen that there is a tradeoff between power consumption and delay time when varying the supply voltage. The effect of having the width of p-channel devices bigger is also analyzed.

Tables and Plots

Index	Interconnect type	Description	Area/μm^2	Capacitance/fF
1	polysilicon	for input A	342	$342\times4.03\times10^{-2} = 13.8$
2	polysilicon	for input B	354	$354\times4.03\times10^{-2} = 14.3$
3	metal	for V_{DD}	804	$804\times2.3\times10^{-2} = 18.5$
4	metal	for V_{SS}	689	$689\times2.3\times10^{-2} = 15.8$
5	metal	for output Y	823	$823\times4.03\times10^{-2} = 18.9$
6	polysilicon	for output Y	283.5	$283.5\times4.03\times10^{-2} = 11.4$
7	metal 2	contact pins	147	$147\times1.73\times10^{-2} = 2.5$

Table 1: Interconnct capacitances

Port identifier	Area references	Capacitance/fF	Total capacitance/fF
A	342	13.8	14.6
	$\frac{147}{3} = 49$	$\frac{2.5}{3} = 0.83$	
B	354	14.3	15.1
	$\frac{147}{3} = 49$	$\frac{2.5}{3} = 0.83$	
Y	823	18.9	31.1
	283.5	11.4	
	$\frac{147}{3} = 49$	$\frac{2.5}{3} = 0.83$	

Table 2: Capacitances associated with ports

Label	Source		Drain		Gate	
	area μm^2	perimeter μm	area μm^2	perimeter μm	area μm^2	perimeter μm
N_A	105	44	87.5	39	21	20
N_B	94.5	41	87.5	39	21	20
P_A	388.5	95	462.5	99	111	80
P_A	462.5	99	444	98	111	80

Table 3: Areas and perimeters of transistors

Input conditions	Rise/Fall	Rise/Fall(ns)	Delay(ns)
A falls	Rise	$62.36 - 61.64 = 0.72$	$61.85 - 61.5 = 0.35$
B falls		$62.44 - 61.72 = 0.72$	$61.99 - 61.5 = 0.49$
Both fall		$62.51 - 61.87 = 0.64$	$62.06 - 61.5 = 0.56$
A rises	Fall	$11.09 - 10.75 = 0.34$	$10.88 - 10.5 = 0.38$
B rises		$11.33 - 10.72 = 0.61$	$11.04 - 10.5 = 0.54$
Both rise		$10.9 - 10.68 = 0.22$	$10.77 - 10.51 = 0.26$

Table 4: Rise, Fall and Delay times for transient inputs for original setup

Input conditions	Rise/Fall	Rise/Fall(ns)	Delay(ns)
A falls	Rise	$62.63 - 61.67 = 0.96$	$62.01 - 61.5 = 0.51$
B falls		$62.65 - 61.77 = 0.88$	$62.1 - 61.5 = 0.6$
Both fall		$62.74 - 61.87 = 0.87$	$62.17 - 61.5 = 0.67$
A rises	Fall	$11.26 - 10.8 = 0.43$	$10.97 - 10.5 = 0.47$
B rises		$11.55 - 10.74 = 0.81$	$11.11 - 10.5 = 0.61$
Both rise		$11.03 - 10.71 = 0.32$	$10.85 - 10.5 = 0.35$

Table 5: Rise, Fall and Delay times for transient inputs for external load being $67.3pF$

Input conditions	Rise/Fall	Rise/Fall(ns)	Delay(ns)
A falls	Rise	$63.56 - 62.01 = 1.55$	$62.45 - 61.5 = 0.95$
B falls		$63.8 - 62.18 = 1.62$	$62.65 - 61.5 = 1.15$
Both fall		$63.83 - 62.3 = 1.53$	$62.76 - 61.5 = 1.26$
A rises	Fall	$11.91 - 11.05 = 0.86$	$11.35 - 10.5 = 0.85$
B rises		$12.47 - 11.08 = 1.37$	$11.71 - 10.51 = 1.2$
Both rise		$11.41 - 10.94 = 0.47$	$11.1 - 10.5 = 0.6$

Table 6: Rise, Fall and Delay times for transient inputs for supply voltage at $3V$

Input conditions	Rise/Fall	Rise/Fall(ns)	Delay(ns)
A falls	Rise	$62.42 - 61.63 = 0.79$	$61.89 - 61.5 = 0.39$
B falls		$662.47 - 61.73 = 0.74$	$62.01 - 61.5 = 0.51$
Both fall		$62.56 - 61.87 = 0.69$	$62.08 - 61.5 = 0.58$
A rises	Fall	$11.39 - 10.88 = 0.51$	$11.07 - 10.5 = 0.57$
B rises		$11.94 - 10.89 = 1.05$	$11.37 - 10.5 = 0.87$
Both rise		$11.13 - 10.8 = 0.33$	$10.92 - 10.5 = 0.42$

Table 7: Rise, Fall and Delay times for transient inputs for supply voltage at $3V$

Parameter	Description	Value(ns)
T	Oscillation period	$133 - 104 = 29$
t_1	Delay from `ENB` to `OUT1`	$104 - 100 = 4$
t_2	Delay from `OUT1` to `OUT2`	$113 - 104 = 9$

Table 8: Timing specifications of NOR gate

Voltage(V)	Frequency(kHz)	Voltage(V)	Frequency(kHz)
0.67	0.827	0.7	1.369
0.81	6.8	1.07	96.39
1.50	412	1.85	740.9
2.08	961	2.27	1149.8
2.56	1442.6	2.77	1643
2.97	1837	3.03	1902
3.06	1921	3.21	2010

Table 9: ring-113 oscillation frequency varies with supply voltage

	$V_{dd} = 3V$		$V_{dd} = 5V$	
	$21 - 13$	$28 - 45$	$21 - 13$	$28 - 45$
Frequency(kHz)	1910	2090	3720	3682
Gate delay(ns)	2.32	2.12	1.19	1.20
Percentage difference	8.6%		0.83%	

Table 10: Performance comparison of $21 - 13$ and $28 - 45$

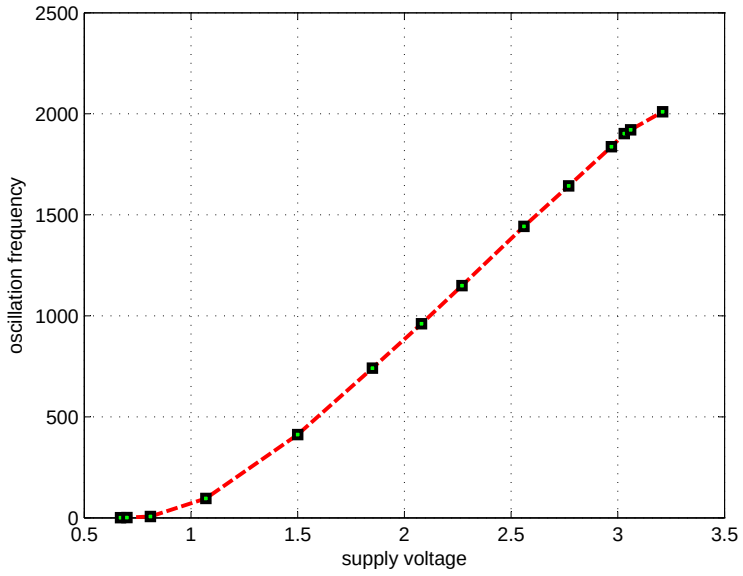

Figure 1: ring-113 oscillation frequency versus supply voltage