Challenges in Gate Level Modeling for Delay and SI at 65nm and Below

Institute for Electronic Design Automation-TUM

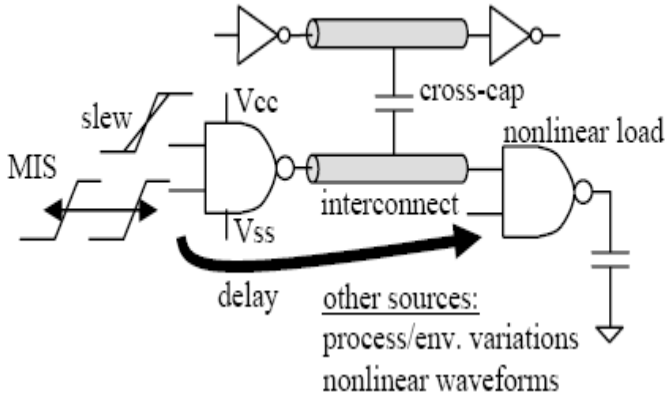
Jelena Frtunik
jelena.frtunik@mytum.de

Outline

- Motivation
- Introduction
- Cell delay model
- Library characterization
- Challenges
- Conclusion

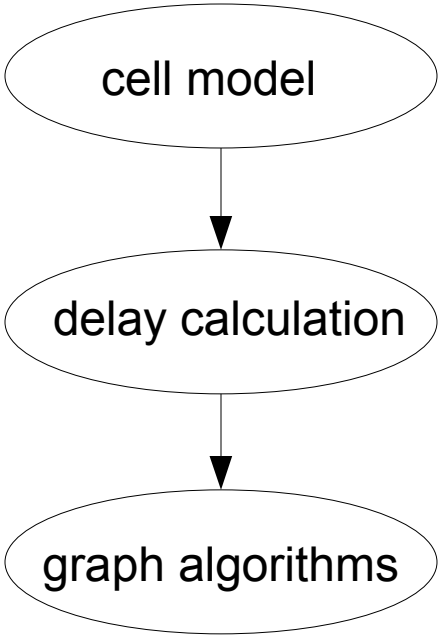
Motivation

- At 90nm and below, physical effects are new challenges for delay calculation.
- Top-level interconnects are becoming more resistive with narrower metal widths, resulting with impedance tha much greater than the drive resistanc the driving cell.
- Miller effect must be accounted in timing analysis.
- At 65nm and below the timing res must always account for the effect o coupling between signal nets, or crosstalk.



Introduction

- Metric for measuring the performance of a chip and releasing it into manufacturing is the result of static timing analysis (STA).
- STA flow



- STA becomes compute intensive with addition of coupling effect and variability computation.
- GOAL: High quality (efficiency and accuracy) STA
- Much of the CPU time during STA is spent on DC => simulation has to be very efficient.

Introduction

- DC can be though as a circuit simulation using cell models.
- DC simulates a small sub-circuit (stage) consisted of driving and receiving gates and (coupled) interconnect parasitics.
- Transistor VS Cell level model
  - Transistor level DC- circuit simulation on the the stage's circuit using a spice-like engine
    - Computationally very expensive
  - Cell level DC - abstraction of the stage's circuit built upon efficient cell models of drivers and receivers and reduced-order models for interconnect.
- DC computes parameters of signal transition relevant for timing analysis (delay and slew) .

Introduction

- CHALENGE in delay cell modeling is: efficient construction of a parameterized timing model of a design, representing the design characteristics as functions of environmental and process variations.
    - Two main approaches to constructing timing model:
        - Analysis in multiple corners of the parameter space (MC)
        - Compute and propagate distributions of arrival times (SSTA)
- Statistical STA:
    - considers manufacturing variations
    - eliminates the drawbacks of corner-based analysis

Cell delay model

- Cell models must accurately represent the behavior of the circuit
- PAST: only one cell model
- NOW : different models for the driving and receiving gates because they affect delay and slew in different ways
- DRIVER MODELS:

1) Constant delay model
2) Load depended model
3) SLDSM
4) ECSM
5) CSM

→ VRM models

→ CSM model

- TRANSITION from VRM to CSM

Driver Models

1) Constant-delay model- defines delay of each gate independent of input slew and output load.
2) Load(fanout) dependent delay model(single-parameter model) - delay depending only on the capacitive load driven by the gate.
3) Input-slew/output-load delay-slew model (SLDSM)(two-parameter model)-output transition parameters (delay and slew) defined as functions of output load and input slew.
4) Extended SLDSM, or ECSM -same parameters as in SLDSM. In addition to the informations about delay and slew the model contains more detailed profile of the voltage response (transition).
5) Current-source model (CSM)- defines drawn current as a nonlinear function of several parameters input voltage and time, or input and output voltages).

Driver models

- PAST: VRM Models (define characteristics of the voltage response at the gate output as function of input slew and output load.)
  - Fast way of determining gate's delay and output slew using simple look-up tables and interpolation.
- NOW: CSM Models (nonlinear voltage-controlled current source, which approximates the current drawn by gate for certain value of input voltage,time, output voltage)
  - To determine delay and slew DC must perform circuit simulation.
- Major factors for transition from VRM to CSM:

More complex (coupled) interconnects
Crosstalk
Strong impact of waveform details

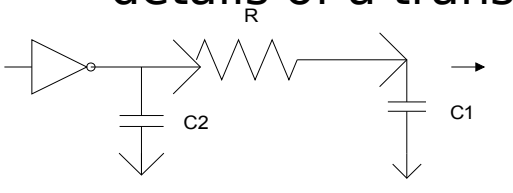
IR-drop
Env/process variations
Back-Miller effect.

Voltage-Response model

- Delay and output slew in Non Linear Delay Model (NLDM) defined as:

$$D = F_d(T_{in}, C_{load}) \qquad T_{out} = F_s(T_{in}, C_{load})$$

- Works good in the case:
  - interconnect parasitics can be approximated by a single linear capacitor
  - crosstalk is small
  - details of a transition other than slew can be neglected



$$C_{eff} = C_1 + C_2\left[1 - \frac{RC_1}{t_D - t_x/2} + \frac{(RC_1)^2}{t_x(t_D - \frac{t_x}{2})} e^{\frac{-(t_D - t_x)}{RC_1}} \left(1 - r^{\frac{-t_x}{RC_1}}\right)\right]$$

  - The Ceff model is accurate up to the point at which the gate begins to behave like a resistance.

Voltage-Response model

- Functions Fd and Fs are represented through 2-D tables.
- Calculation of Delay and output slew (Tout) is done via 2-D interpolation on the delay and slew tables.
  - ADVANTAGE:
    - Delay and slew values on the grid points are exact.
      - The success of evaluation depends on number of grid points and their values, and the interpolation method.
  - DISADVANTAGE:
    - Problem when Tin or Cload are outside of the corresponding characterization range.
      - Solution: Finding robust extrapolation technique

Current-Source Model (CSM)

- Model where each output pin is described using a nonlinear voltage-controlled current-source.
- General form:

$$I_{drv} = F_i(V_o(t), t, p_1, \ldots, p_N)$$

- GOAL of DC : determine the output voltage response. The current is determined with numerical solution of state-space system describing voltage responses on the nodes of a stage

$$E\frac{dx}{dt} = Ax + Bu(v,t)$$

$$v = Cx$$

Channel-Connected Component

- Current drawn by a cell having a SINGLE Channel-Connected Component (CCC) can be modeled as:

$$I_{drv} = F_i(V_i, V_o) + G\left(\frac{dV_i}{dt}, \frac{dV_o}{dt}, V_i, V_o\right)$$

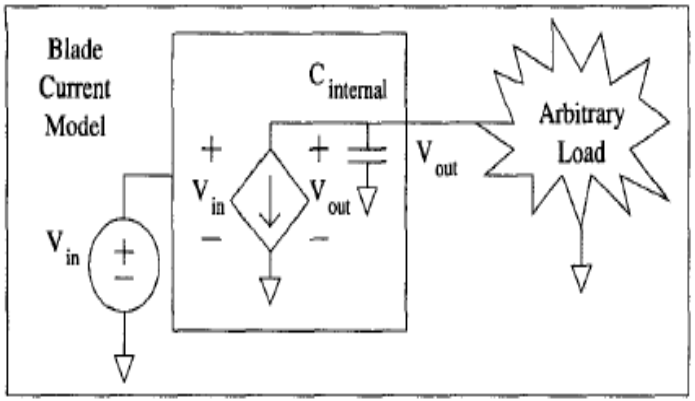
- Nonlinear STATIC current source depending on instantaneous values of Vi and Vo. (characterized using a dc-simulation).
- DYNAMIC current that is from nonlinear capacitors of the cell's transistors.

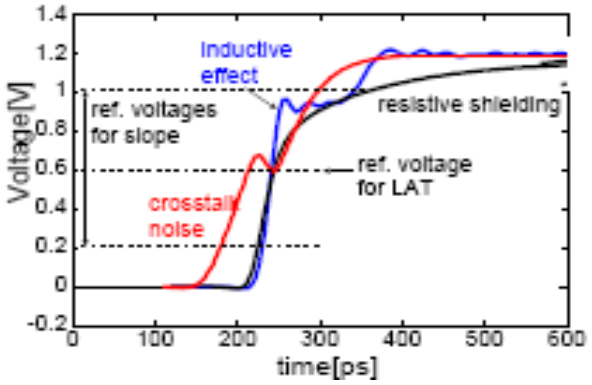
$$G = c_m(V_i, V_o)\frac{dV_i}{dt} + c_g(V_i, V_o)\frac{dV_o}{dt}$$

- Coefficients of two terms correspond to Miller and output capacitance respectively, and are nonlinear functions of Vi,Vo.

Blade model (CSM)

- Blade is the first high-speed cell model that can:
  - Produce and consume arbitrary voltage waveforms, including noisy waveforms, at near-SPICE accuracy.
  - Operate on arbitrary loads including lumped-C, PI-models etc.
- Razor - interconnect model.
  - Accurately calculates the resulting waveform that is used to drive the next Blade model, all at speeds 5-6 orders of magnitude faster than SPICE.





Channel-Connected Component

- Current drawn by a cell having a MULTI Channel-Connected Component (CCC) can be accurately modeled as:

$$I_{drv} = F_i(w, V_o) + G\left(\frac{dw}{dt}, \frac{dV_o}{dt}, w, V_o\right) \qquad w(t) = J(V_i)$$

- The output current is not a function of Vi, but of some other transition, w(t). w(t) can be shifted in time, distorted and/or inverted compared to Vi.
  - For simple multi CCC w(t) is a transition at the input of the last CCC.
- Operator J can be:
  - based on a lookup table representing the slew of $w(t)$ as a function of slew of vi.

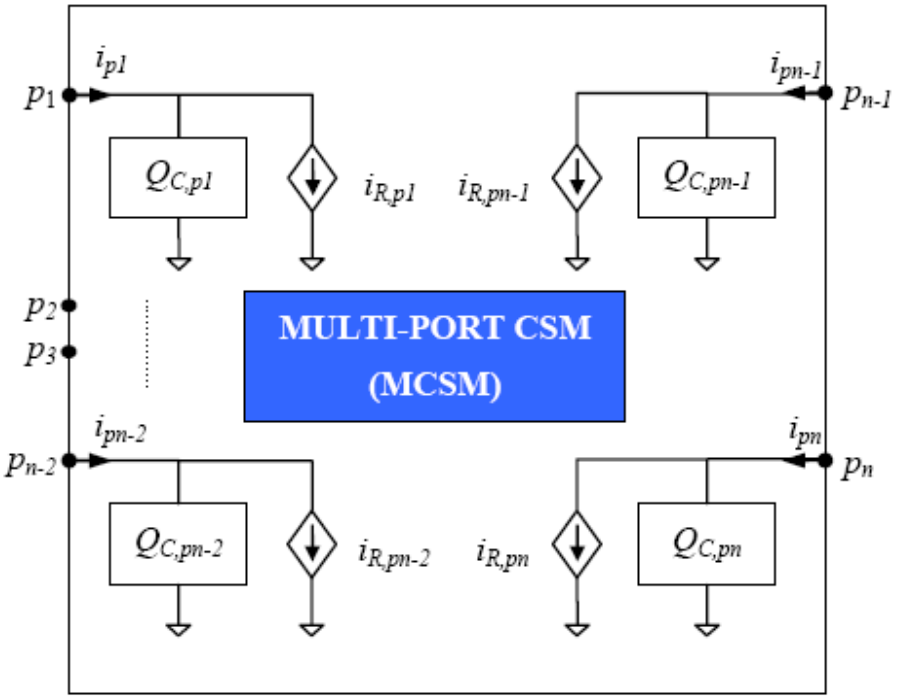
Channel-Connected Component

- All models are based on a partition of the cell's circuit into CCCs and characterizing each CCC with a dc-current table.
  - Advantage: more compact representation of the current
  - Disadvantage: need to understand the internal topology of the cell
- FUTURE: New technique of constructing a multi-CCC CSM based on a piece-wise polynomial macro-model
- Alternative approach for describing CSM is constructing current model by treating cells as black-boxes. One example is Synopsys CCS (Composite Current Source) model.
  - The drawn current is described as an explicit function of input slew, output load and time

Receiver Model

- RM of gate is consisted from two parts:
  - Load Model (LM) -impact from the receiving gate on the driving stage
    - current injected into the gate or time-dependent capacitive load
  - Delay metric
- The impact of the receiver gate can be described either as current injected into the gate, or time-dependent capacitive load because: $I = C\frac{dV}{dt}$
- LM model
  - Past: single capacitance value;two values (for rise and fall) ; capacitance range(for min and max delays);
  - Now: capacitance as a function of input slew and/or output load; voltage-dependent capacitance; CSM-based load model.

Multi-port current source model (MCSM)



$$i_{R,p_i} = f_i(V_{p_1},....,V_{p_n})$$

$$Q_{C,p_i} = g_i(V_{p_1},....,V_{p_n})$$

- MCSM for a two input nand gate has three ports:
  - One for each input/ output

- Miller capacitance at the input nodes is captured by the Qc because it depends on the voltages at the other ports in the cell including the output port.

- This approach can allow modeling simultaneous switching of multiple inputs.

Receiver Model

- !!! Detailed modeling of pin capacitance is helpful in the cases where load capacitance is dominated by the nonlinear pin capacitance of the receivers, which causes distortions of the waveform.

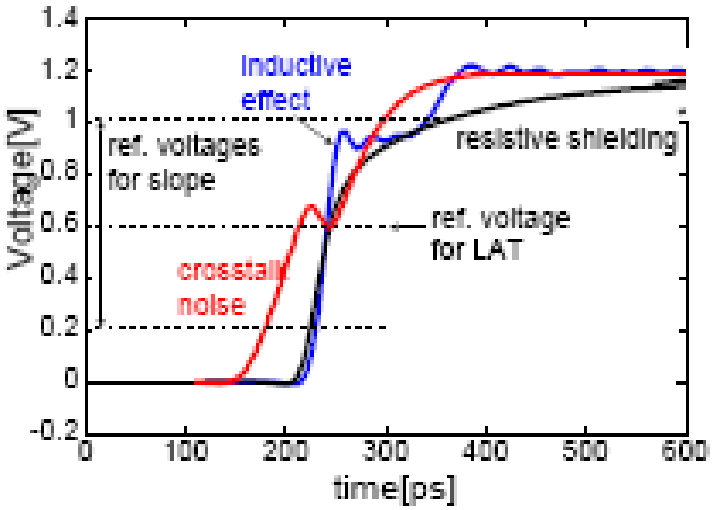
- The MCSM model accurately models the nonlinear pin capacitance and so-called back-Miller effect associated with coupling between input and output pins of the first CCC of the receiving gate.
  - DISADVANTAGES:
    - Not practical to use on large scale design because it requires modeling of all receiver nodes as ports.
    - Computation of response becomes costly because it requires simulation of a state-space model with multiple current sources.

Receiver Model

- DM-Delay metric interprets the voltage response in terms of delay and slew.
  - Delay and …ge waveforms by finding the … thresholds,Vref for delay a…
  - FUTURE: E…

    Calcul…ased on Least Squar… waveform close to the …

    The sc…eforms affected by res…ire inductance without library extensions.



Library characterization

- Library characterizer may perform millions of transistor-level simulations and take days to produce a high accuracy standard cell library.
- A typical library characterization flow consists of 4 major steps:
  - sensitization of input vectors
  - setting up and performing circuit simulations
  - collecting and processing results
  - generating libraries

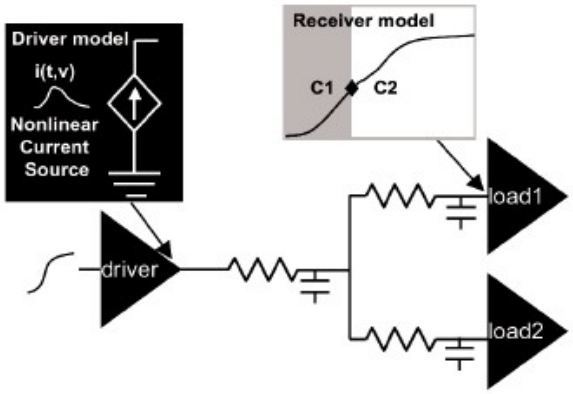
- DRIVER Model
  - Depends on the type of driver model (VRM or CSM)
- RECEIVER Model

Characterization algorithm for VRM model

- VRM model -steps:
  1) For given an input slew and an output load, a transistor level simulation is performed by applying a voltage source at the input and a lumped capacitive load at the output.
  2) The slew and the delay in the case of NLDM are measured from the simulated voltage waveform.
  3) The process is repeated for a list of predefined input slews and output loads leading to several 2-D tables.
- DISADVANTAGES:
  - Not so clear how to choose values on input slew and output load in order to minimize the interpolation errors in DC.
  - Measurement of slew and delay on the output transition involves interpolation between time points which introduces a source of error.

Composite Current Source

- Composite Current Source (CCS from Synopsis) modeling technology is the first in the industry to deliver a complete open-source current based modeling solution for timing (composed of driver and receiver model) noise and power.



· Synopsys's CCS is similar in terms of characterization flow to the VRMs (like ECSM.)

  • Instead of measuring and storing voltage WFs, CCS measures and stores current WFs consisting of time-current pairs along the transition. This leads to larger disk size of the library.

Characterization algorithm for CSM Model

- CSM model
  - Usage of DC simulations which are cheaper than transient simulations required for VRMs
  - Less disk space because they are based on a voltage-controlled current sources which are independent the parameters like input slew and output load.
  - An additional runtime cost of waveform processing is necessary because the measured current is likely to be much noisier than voltage waveforms.

Characterization algorithm for Receiver Model

- Cell characterization of RM can be described as:
  - Simplest method is based on adding the average gate and diffusion capacitances of the devices connected to the cell's input pin.
  - Measuring current through a saturated ramp voltage source attached to the cell's input pin. With averaging over time (integration) we determinate the pin capacitance value. More expensive but more accurate.

Extension to Process/Environment variations

- Variation-aware STA requires specially constructed libraries providing way of computing delays for different corners, and sensitivity of delay and slew with respect of process and environmental parameters.
- Two types of variation-aware STA :

  1) Multi-corner STA -requires multiple standard-cell timing libraries, generated using different process, voltage and temperature (PVT) parameters in the transistor level simulation

     - Parallelize and run on multiple machines.

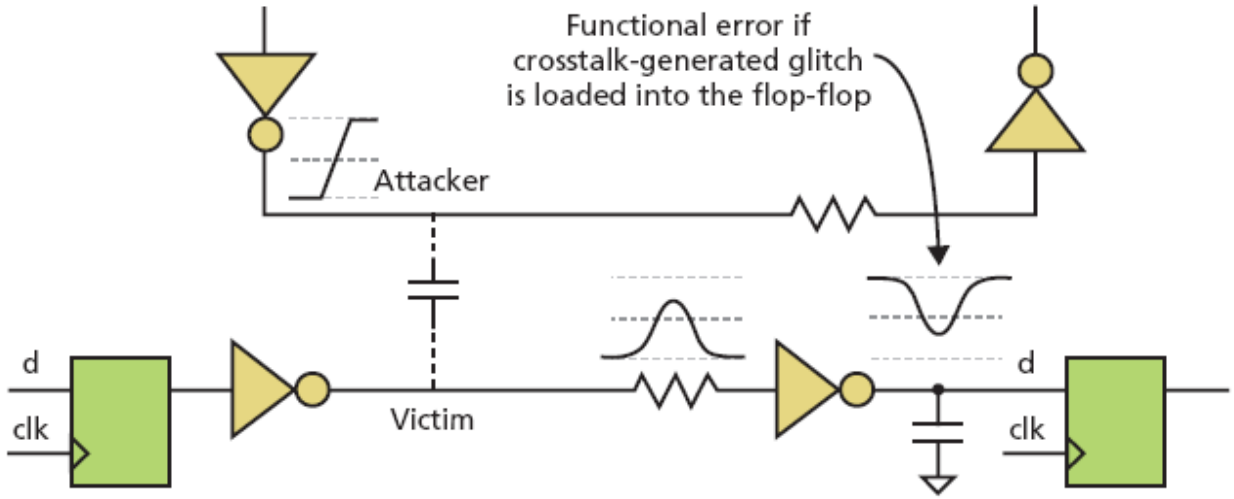
Extension to Process/Environment variations

2) SSTA - the timing, power, and signal integrity characteristics of each element in the device are represented as Probability Distribution Functions (PDFs).

- Similar to MC: A full set of simulations for characterizing the library is performed under a small change of the concerned process parameter. The resulting measurement is then used to compute the sensitivities. This is repeated for each process parameter.
- A set of linear equations for sensitivity is formulated and solved in transistor-level simulator. The set is derived from the original circuit equations by linearizing them around the nominal set of process/env parameters and responses.

SI

- Signal integrity (SI) means ensuring that signals faithfully propagate to their intended destinations within their allocated time frames.
- Beca m a
  highe tly more
  sever
- Low-p different
  switc
  - H more
    vu
  - Lo
    aggressive as attackers.
  - The timing analyzer should fully account this scenario, and report potential violations accordingly.



Challenges

- Compact variation-aware cell model, as well as a compact and efficient (for access) storage of variation-aware library
- Construction of cell model that supports efficient and accurate:
  - Timing
  - Noise
  - Power analyses

- Timing analysis engine have to be electromigration-, thermal-, and lithographic aware in order to have more accurately assess to the signal integrity, power, and performance characteristics of the chip.

Conclusion

- At 65nm and below, we MUST consider crosstalk and variability effects.
- VRM models have to be replaced with CSM because they can handle:
  - Crosstalk,
  - Complex distributed load
  - Nonlinear pin capacitance
  - Back-Miller effect.

THANK YOU FOR THE ATTENTION!

QUESTIONS???