Parameter Variations
Variations

- Wafer to Wafer
- Inter-die
- Intra-die

Environmental
- \(V_{dd}\)
- Temperature

Process
- Device
- Interconnect
Process Variations

Source: IBM, ISSCC’00
Sources of Process Variations

- **Film thickness variations**: Tox is critical but is relatively well controlled. Vertical variations caused by Chemical-Mechanical Planarization (CMP); Inter-layer distances (dielectric thickness).

- **Horizontal**: Poly line-width & $L_{\text{eff}}$ variation comes from:
  - Mask, exposure and etch variations (photolithography)
Statistical Description: Lumped Statistics

\[ L = L_0 + \Delta L \]

\[ \Delta L = N(\mu = 0, \sigma^2_L) \]

\( L_0 = \mu_L \) is the sample mean of \( L \), and \( \sigma^2_L \) is estimated as the sample variance. The combined set of underlying deterministic and random contributions are lumped into a combined “random” statistical description.

It is important to remember the basis for the distribution being described, and to apply the distribution only to similar samples. For devices on one wafer, the distribution (mean and variance) for \( L \) can be different from devices within a single die.
Separation of inter-die and intra-die variations

Sources of parametric variation are separated into: interdie and intradie variation.

Since the variation within the die - due to layout pattern dependencies – may be larger than variations across wafer, thus matching of devices is different on one chip vs. on one wafer.

\[ P = P_0 + \Delta P_{\text{intradie}} + \Delta P_{\text{interdie}} + \Delta P_e \]

- \( P_0 \) = nominal design value
- \( \Delta P_{\text{intradie}} \) = intradie variation (within a given chip)
- \( \Delta P_{\text{interdie}} \) = interdie variation (from one chip to another)
- \( \Delta P_e \) = remaining “random” or unexplained variation

\( P \) indicates a structural or electrical parameters such as \( W, t_{\text{ox}}, \) device parameters, \( V_{\text{th}}, \) channel mobility, coupling capacitances, line resistances.
**Intradie Variation**

It is the deviation occurring spatially within any one die. In contrast to interdie variation, intradie variation contributes to the loss of matched behavior between structures on the same chip. Two important sources for intradie variations are: (1) wafer level trends and (2) die pattern dependencies.

(1) **Wafer scale variation** can result in small trends that are reflected across the spatial range of the chip. For example, some deposition processes suffer from a gentle “bowl” or concentric ring pattern in thickness from the center of the wafer outwards. These can cause systematic trends across the die.

(2) **Die-pattern dependencies** can create variations that have become increasingly problematic in IC fabrication. Two interconnect lines that are designed identically in different regions of the die may result in lines of different width, due to photolithographic interactions, plasma etch micro-loading, ..etc. Distortion in lens and other elements of a lithographic system are also known to create systematic variations across the die.

Within any randomly selected die, wafer level variation can be approximated as a random bias function of the coordinates of the die within the wafer. Due to the small die area w.r.t. wafer, it is reasonable to assume that the wafer level variation can be modeled within the die as a linear function of position:

\[ P_{\text{intradie}}(x, y) = W(\omega_0, x, y) = \omega_0 + \omega_x x + \omega_y y \]

where \( \omega \) denotes the coefficients \( \omega_0, \omega_x \) and \( \omega_y \) are random variables describing the plane (example: slanted plane)
Supply Voltage Variation

- Activity changes
- Current delivery RI and L(di/dt) drops, wire planning
- Within-die variation

**V\textsubscript{dd} Profile**

**IBM Chip**
- 0.13\(\mu\)m CMOS Tech.
- 160K Macros
- 8mm X 8mm
- \(V\text{\textsubscript{dd}}=1.2V\)

**Power** = 48W, 20% leakage

**Variations in \(V\text{\textsubscript{dd}}\)** = 3% to 15%

**Hot spots** = High power density regions

Temperature Variation

• Higher temperature results in slower transistors, higher interconnect resistance and exponentially higher subthreshold leakage

• Activity & ambient change
• Within-die variation
• Floorplan
• Power distributions

IBM Chip

0.13\textmu m CMOS Tech.

CPU Core of \textmu Processor

2.5mm X 4.7mm

\(V_{dd}=1.0V\)

Power = 5.6W, 20\% leakage

Variations in temperature = 0.8\textdegree C to 30.3\textdegree C

Hot spots = High power density regions

Motivation

Increasing uncertainty in Timing

- Technology-generated uncertainty
  - Process variations are increasing with every new technology generation
  - Coupling noise impact on timing
  - Power supply variations and its impact on timing
  - Inaccuracies in the delay calculator

Statistical timing analysis regarded as a key area in the industry and in technology roadmap (ITRS)
**Frequency & Subthreshold Leakage**

- **Normalized Frequency**
- **Normalized Leakage (Isb)**

**Source:** Intel
$V_t$ Distribution

- 0.18 micron
- ~1000 samples
- ~30mV

<table>
<thead>
<tr>
<th>$\Delta V_t$ (mv)</th>
<th># of Chips</th>
</tr>
</thead>
<tbody>
<tr>
<td>-39.71</td>
<td></td>
</tr>
<tr>
<td>-25.27</td>
<td></td>
</tr>
<tr>
<td>-10.83</td>
<td></td>
</tr>
<tr>
<td>3.61</td>
<td></td>
</tr>
<tr>
<td>18.05</td>
<td></td>
</tr>
<tr>
<td>32.49</td>
<td></td>
</tr>
</tbody>
</table>

- High Freq
- High Isb
- Medium Isb
- Low Freq
- Low Isb

Source: Intel
Leakage with $V_{dd}$ and Temp. Variations

Leakage considering environmental variations
• Accurate leakage model of actual $V_{dd}$ and temperature profile

Large leakage variations regions correspond to the hot spots in the $V_{dd}$ and Temp. profiles

Temperature Dependence

- Gate Leakage is unaffected by temperature
- Subthreshold Leakage has a super linear relationship to Temperature
  - At room temperatures, gate leakage is dominant for future processes
  - At higher temperatures (where chips normally operate), subthreshold leakage dominates
  - This situation may change in future with technology scaling ($I_{\text{gate}}$ scales faster)
- For a 50nm device with 1.5nm $T_{\text{ox}}$

# Leakage with Process Variations

## Monte-Carlo Analysis

### NMOS
- $T_{ox}$ and Length have greatest effect
- The mean of length variations is not equal to nominal
- Larger mean and variance when all parameters are varying

### PMOS
- Much larger dependence on length

<table>
<thead>
<tr>
<th>Parameter varied</th>
<th>Mean Leakage (pA)</th>
<th>Standard Deviation</th>
<th>SD/ Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>None</td>
<td>42.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nch</td>
<td>42.5</td>
<td>1.8</td>
<td>4.2%</td>
</tr>
<tr>
<td>$T_{ox}$</td>
<td>42.9</td>
<td>9.0</td>
<td>21.0%</td>
</tr>
<tr>
<td>$L_{drawn}$</td>
<td>44.1</td>
<td>9.6</td>
<td>21.8%</td>
</tr>
<tr>
<td>All above</td>
<td>45.9</td>
<td>15.7</td>
<td>34.2%</td>
</tr>
</tbody>
</table>


<table>
<thead>
<tr>
<th>Parameter varied</th>
<th>Mean Leakage (pA)</th>
<th>Standard Deviation</th>
<th>SD/ Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>None</td>
<td>26.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nch</td>
<td>26.5</td>
<td>1.0</td>
<td>3.8%</td>
</tr>
<tr>
<td>$T_{ox}$</td>
<td>27.0</td>
<td>6.2</td>
<td>23.0%</td>
</tr>
<tr>
<td>$L_{drawn}$</td>
<td>32.0</td>
<td>22.0</td>
<td>68.8%</td>
</tr>
<tr>
<td>All above</td>
<td>33.6</td>
<td>27.8</td>
<td>82.7%</td>
</tr>
</tbody>
</table>

Impact on Circuits and Micro-architectures
Delay Paths

Path delay variability due to variations in $V_{dd}$, $V_t$, and Temp impacts individual circuit performance and power.

Objective: full chip performance, power, and yield. Multivariable optimization of individual circuit-constituent elements.

Optimize each circuit for full chip objectives by guard-banding.
Delay Paths: Problem of Correlations

gate delay pdfs  Arrival time pdf

Arrival time pdf

Arrival time pdf

I1

A

B

C

D
Performance of high-speed synchronous digital systems is reduced significantly by clock skew, even though the H-tree clock distribution network is used.

When circuits run at giga-Hz, clock skew becomes a significant part of the clock period.

Random and wafer level variation impact

Interconnect sensitivity analysis

Statistical Interconnect Impact

$L_{\text{eff}}$ Variations:
- Impact of Poly-Silicon mask, lithography and etch at the chip level
- Each buffer (total of 65) has a unique value of $L_{\text{eff}}$
Interconnect Sensitivity Analysis

- Take derivative of delay with respect to variable of interest
  - $V_t$
  - Channel length
  - Gate oxide thickness
  - ILD thickness
  - Wire thickness
  - IR drop
  - Temperature gradients
Control of Parameter Variations
Body Biasing Techniques

Forward Body Bias: $V_t$ Modulation

![Graph showing normalized operating frequency vs. forward body bias (mV)]

- Normalized operating frequency
- Forward body bias (mV)

**1.2V 110°C**

![Graph showing $F_{\text{max}}$ (MHz) vs. $V_{dd}$ (V)]

- $F_{\text{max}}$ (MHz)
- $V_{dd}$ (V)

**Body bias chip with 450 mV FBB**

**NBB chip & body bias chip with ZBB**

Intel's 6.6M transistors communications router chip

S. Narendra et al., “1.1V 1GHz Communications Router with On-Chip Body Bias in 150nm CMOS,” ISSCC pp.270-271, 2002
Body Biasing Techniques

Reverse Body Bias: Leakage Reduction

RBB reduces subthreshold leakage
Less effective with: shorter L, lower $V_t$, & scaling

A. Keshavarzi et al., “Effectiveness of Reverse Body Bias for Leakage Control in Scaled Dual $V_t$ CMOS ICs,” ISLPED, pp.207-210, 2001
Body Biasing Techniques
Adaptive Body Bias

Supply Voltage Control

Bin improvement by adaptive $V_{dd}$
20% of dies are pushed from Bin1 to Bin2 + recovered dies that fell below Bin 1

Supply Voltage Control

$V_{dd}$ Variation Reduction

On die decoupling capacitors reduce $V_{dd}$

- Cost area, and gate oxide leakage concerns

Temperature Control

When temperature exceeds the threshold
  1. Lower frequency (activity)
  2. Lower $V_{dd}$

This leads to a power consumption drop followed by a drop in on-die temperature