On the Fidelity of the Ax+B Equipment Model for Clustered Photolithography Scanners in Fab-level Simulation

James R. Morrison

Industrial and Systems Engineering

WSC’11 – December 11 – 14, 2011 – Phoenix, AZ, USA
Presentation Overview

• Motivation

• Ax+B model → Generalized Ax+B model

• Simulation to test model → Flow Line as a baseline for photo tool

• Should we care?
Presentation Overview

Motivation

- Ax+B model → Generalized Ax+B model
- Simulation to test model → Flow Line as a baseline for photo tool
- Should we care?
Motivation: Semiconductor Wafer Manufacturing

- Global revenue in 2010: US $ 304,000,000,000\textsuperscript{[1]}
- High construction cost for fabricator: US$ 5,000,000,000\textsuperscript{[2]}
- Clustered photolithography tools: US$ 20,000,000+
  - The most expensive tool in a fabricator
  - Typically the bottleneck of the fabricator
  - Key yield and cycle time contributor

\textsuperscript{[1]} HIS iSuppli April 2011
\textsuperscript{[2]} Elpida Memory, Inc., available at http://www.elpida.com
\textsuperscript{[3]} http://www.rocelec.com/manufacturing/wafer_fabrication/
Motivation: 450 mm wafer era

• The anticipated transition to 450 mm diameter wafers
  • Smaller lot sizes and increased setup frequency

• To assess the feasibility and cost of future fabricators, fab-level simulation is one key tool employed

• **Question**: Are the current equipment models sufficient?
Motivation: Empirical Equipment Models

- Equipment models for fab-level simulation

<table>
<thead>
<tr>
<th>Model</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ax+B Model</td>
<td>Used in commercial fab-wide simulation software</td>
</tr>
<tr>
<td>Wood (2007)</td>
<td>Used in commercial fab-wide simulation software</td>
</tr>
<tr>
<td>SEMATECH</td>
<td>First to propose Ax+B model</td>
</tr>
<tr>
<td>EPT Model</td>
<td>Jacobs et. al. (2006) Measure effective process times for sequential processing</td>
</tr>
<tr>
<td>SDF Model</td>
<td>Unbehaun (2007) Focus on concurrent processing, predicting tool behavior with slow down factors</td>
</tr>
</tbody>
</table>

- However,
  - These models are parameterized with empirical data
  - When conditions change, does it still work?
Presentation Overview

- Motivation

Ax+B model $\rightarrow$ Generalized Ax+B model

- Simulation to test model $\rightarrow$ Flow Line as a baseline for photo tool

- Should we care?
Ax + B Models

- **Notation**
  - $S_i$: start time of lot $i$
  - $C_i$: completion time of lot $i$
  - $W_i$: wafers in lot $i$

- **Performance measures**
  - Cycle time of lot $i$
    \[ T_i^{CT} := C_i - S_i \]
  - Throughput time of lot $i$
    \[ T_i^{TT} := \min\{ T_i^{CT}, C_i - C_{i-1} \} \]
Ax + B Models

- Address non-linearity: Setups and first wafer delay
- Time between wafer completions $A_i$
- First wafer delay $B_i$
- Throughput time estimation

$$T_i^{TT} = A_i \cdot (W_i - 1) + B_i$$
Ax + B Models

- Popularly known as affine models
- First proposed by SEMATECH
- Used in many commercial fab-wide simulation software

However,
- No analytic method to obtain the B values
- Empirically obtained B values are best in exact situation for which the data was collected

- Difficult to assessing implications of changes and design configurations
  (e.g. assess of changes in photolithography capacity, determining the capacity loss caused by small lot sizes, determining the number of tools required for 450 mm wafer fab)
Generalized Ax+B model

- Affine models use
  - One A and B value for the tool, or
  - One A and B value per lot class (per tool)

- Cluster tools and clustered photolithography tools may have
  - Sequence dependent setups between lot changes
  - Internal buffers for continued throughput during setups
  - Increased frequency of setups in 450 mm wafer fabs

- Suggest a generalized Ax+B model with sequence dependent parameters

<table>
<thead>
<tr>
<th>$B_{k(i),k(i-1)}$</th>
<th>$k(i-1) = 1$</th>
<th>$k(i-1) = 2$</th>
<th>$k(i-1) = 3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$k(i) = 1$</td>
<td>65</td>
<td>298.69</td>
<td>296.40</td>
</tr>
<tr>
<td>$k(i) = 2$</td>
<td>245.87</td>
<td>60</td>
<td>267.54</td>
</tr>
<tr>
<td>$k(i) = 3$</td>
<td>244.71</td>
<td>267.60</td>
<td>55</td>
</tr>
</tbody>
</table>
Presentation Overview

• Motivation

• Ax+B model  Generalized Ax+B model

Simulation to test model  Flow Line as a baseline for photo tool

• Should we care?
Flow Line Models

• Proposed to address some of the difficulties associated with using affine models for cluster photolithography tools

• Address nonlinear effects in throughput as the lot sizes and setup frequency change

• Use this model as our baseline for clustered photolithography tools

• More detailed than affine models, but require additional computation
Flow Line Models

- Deterministic service time for $P_i$
- One module can hold at most one wafer
- Buffers can be modeled by process modules with zero process time
- EEEs (Elementary Evolution Equations) characterize wafer movement
Flow Line Models

- EEEs (Elementary Evolution Equations)
  - $X_{w,i}$: entry time of wafer $w$ to process $i$
  - $\tau_i$: service time for process $P_i$

$$X_{w,i} = \max \{ X_{w,i-1} + \tau_{i-1}, X_{w-1,i+1} \}$$

Wafer is ready to enter

Process is ready (empty)
Flow Line Models

- Detailed example

\[ X_1(i, 1) = \max\{a(i), X_{p+1}(i-1, W_{i-1})\} + \tau_{\text{setup}} \text{ if } c(i) \neq c(i-1), \]
\[ X_1(i, 1) = \max\{a(i), X_2(i, W_{i-1}-R_1+1)\} \text{ if } c(i) = c(i-1), \]
\[ X_m(i, w) = \max\{X_{m-1}(i, w) + \tau_{c(i)}^{m-1}, X_{m+1}(i, w-R_m)\} \text{ for } i \neq 1 \text{ and } m \neq 1, \]
\[ X_{s+1}(i, 1) = \max\{X_s(i, 1) + \tau_{c(i)}^s, X_{s+2}(i, W_{i-1}-R_{s+1}+1)\}, \]
\[ X_M(i, w) = \max\{X_{M-1}(i, w) + \tau_{c(i)}^{M-1}, X_M(i, w-R_M) + \tau_{c(i)}^M\}, \]
\[ C_M(i,w) = X_M(i, w) + \tau_{c(i)}^M \]
for \( w = 1, \ldots, W.\)
Flow Line Models

- **Accuracy of Flow Line Models**

  ![Photolithography Tools](image)

  ![Flow Line Models](image)

- **Control of cluster tool robot:**
  - Robot essential: Petri net, MIP models, etc.

- **Throughput of clustered photolithography tools?**
  - Robot overhead can be incorporated into process times
  - Bottleneck behavior dictates throughput
Flow Line Models

• **Throughput of cluster tools:**
  - Good robot policy provides bottleneck throughput
  - Typical robot overhead is small and easy to include
  - Practical study: 1% throughput error (Morrison 2011)
  - Cycle time estimation for cluster tools (Park and Morrison 2011)

![Graph](image-url)
Numerical Experiment

- 18,000 lots and 10 replications
- 3,000 lots for warm-up period
- JIT (Just in time) arrivals
- Pre-scan track setups between lot changes
- Setup duration: Uniform [240, 420]
- Reticle alignment: Uniform [120, 300]
Simulation Outline

**Estimation of Affine model**

Flow Line Model

1. **Affine Model for T = 6**
   - Get A and B values

2. **Estimate Throughput for T = 3**

**Flow Line (Baseline)**

Flow Line Model

1. **Calculate Throughput for T = 3**

Compare it!!
Simulation Results

• Generate **Affine Model (AX+B model)**
  • Average train size of $T = 6$ lots between setups
  • (through flow line model) Generate data for throughput time $T_i^{TT}$
  • Generalized A and B parameters
• Just as is done in the industry except generalized A and B

< A values >

<table>
<thead>
<tr>
<th>Class of Current Lot</th>
<th>Class of Previous Lot 1</th>
<th>Class of Previous Lot 2</th>
<th>Class of Previous Lot 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>65</td>
<td>65</td>
<td>65</td>
</tr>
<tr>
<td>2</td>
<td>60</td>
<td>60</td>
<td>60</td>
</tr>
<tr>
<td>3</td>
<td>55</td>
<td>55</td>
<td>55</td>
</tr>
</tbody>
</table>

< B values >

<table>
<thead>
<tr>
<th>Class of Current Lot</th>
<th>Class of Previous Lot 1</th>
<th>Class of Previous Lot 2</th>
<th>Class of Previous Lot 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>65</td>
<td>298.69</td>
<td>296.40</td>
</tr>
<tr>
<td>2</td>
<td>245.87</td>
<td>60</td>
<td>267.54</td>
</tr>
<tr>
<td>3</td>
<td>244.71</td>
<td>267.60</td>
<td>55</td>
</tr>
</tbody>
</table>
Simulation Results

- Using the basic A/B parameters obtained
  - Estimate $T = 3$ behavior
  - Vary the wafers per lot from 24 down to 1
Simulation Results

- Using the generalized A/B parameters obtained
  - Estimate T = 3 behavior
  - Vary the wafers per lot from 24 down to 1
Simulation Results

- Average throughput % error

![Graph showing simulation results]

- Generalized Ax+B Model
- Ax+B Model
Presentation Overview

- Motivation
- $Ax+B$ model $\rightarrow$ Generalized $Ax+B$ model
- Simulation to test model $\rightarrow$ Flow Line as a baseline for photo tool

Should we care?
Simulation Results

• **What does the result mean for industry?**

  • Clustered photolithography tools: US$ 20,000,000+
    - The most expensive tool in a fabricator
    - Typically the bottleneck of the fabricator
    - Key yield and cycle time contributor

  • Fewer wafers per lot for 450mm wafer
    - (Pettinato and Pillai 2005) and (Kondo 2006)
    - 3, 5, 10 wafers per lot are recommended for 450 mm wafer fabs

  • **Affine models trained on existing data may not give good predictions of the throughput for photolithography tools in 450 mm wafer fabs**
Concluding Remarks

• Motivation

• Ax+B model  Generalized Ax+B model

• Simulation to test model  Flow line as a baseline for photo tool

• Should we care? Probably!
  • New improved generalized Ax+B model does not well predict thpt
  • Lot sizes of 3, 5, 10 wafers are recommended for 450 mm fabs
  • What models should we use?