Department of Industrial Systems Engineering and Management | IE3100M Systems Design Project (AY2017/2018)

Development and Optimization of Gating Factors Based on O²DES Framework

Project Advisors: Dr Liu Yang, Dr Li Haobin (NUS) | Mr. Heng Teng Nai, Mr. Christian Indra Kusuma (Micron)
Team Members (Group 7): Mark Chan Wei Shuen | Jia Yan | Phang Fu Chun | Toh Yong Sheng | Yan Xiaoxuan

OVERVIEW

Problem Definition

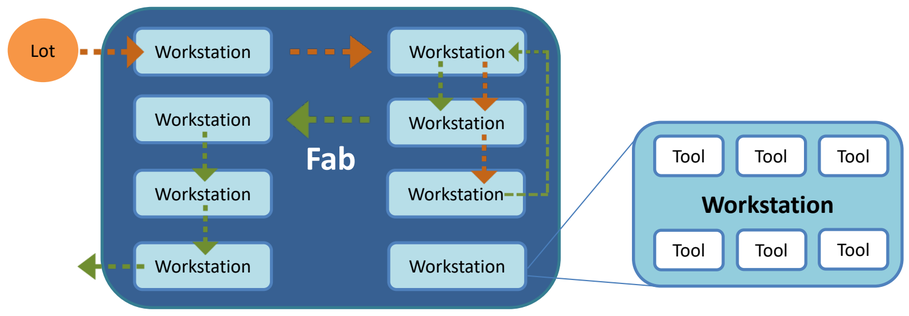
Queue Time Lots (QTL) are wafer lots whose quality is affected by how long they spend queueing. **Maximum Queue Times (Max QTs)** are imposed on such lots along specified process steps throughout fabrication. A **breach** occurs if a QTL's queue time exceeds its Max QT, and the QTL is considered defective. **Gating Factors (GFs),** a ratio between 0 to 1, are imposed and used to quantify the **Work-in-Process (WIP)** accommodable by workstations, without causing any breaches.

Current Practice

Engineers determine the GF based on past experiences and observation of current queuing situation at a workstation. This poses risks of:

- Overestimating GF, which increases the possibility of breaches, and
- Underestimating GF, which results in idle capacity.

Lot Process Flow



During fabrication, wafers are transported as lots and undergo multiple steps for completion. Lots are transported to workstations that contain multiple tools performing the required process steps. Re-entry of lots to previous workstations is also possible due to the complex nature of semiconductor manufacturing.

PROJECT METHODOLOGY

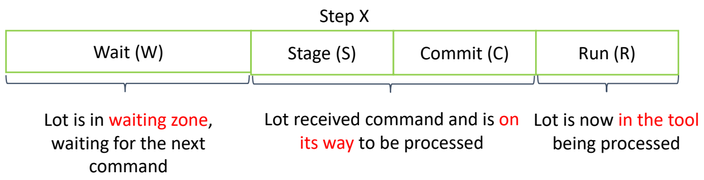
Problem Formulation → Simulation Model Design → Simulator Implementation → Simulator Calibration → Stress Test of Simulator → Data Retrieval and Analysis → Input Modelling and Output Analysis → Gating Factor Optimization

Project Objective: To develop a means that allows the auto-generation of optimal Gating Factors, given real-time Fab dynamics, that minimize the number of queue time breaches without extensively compromising production rate.

KEY CONCEPTS AND ALGORITHMS IMPLEMENTED

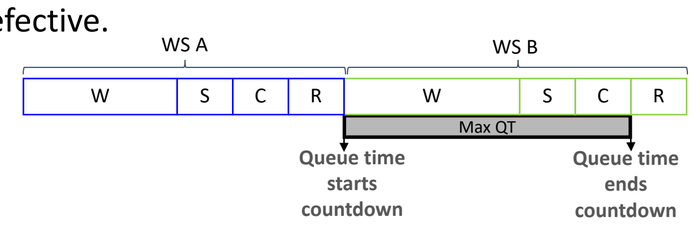
Step Status

Each process step consists of 4 different statuses – Wait (W), Stage (S), Commit (C) and Run (R), indicating the current location and command instructed on a particular lot.



Maximum Queue Time

The maximum amount of time a QTL can spend queuing, between the end of R in WS A and before the start of R in WS B, before it is considered defective.



Gating Factor

Breaching may occur downstream in the production line if there is no regulation on the flow of QTLs. Specific GFs are imposed on every WS to prevent breaches from occurring in that WS.

Gating Decision

It is the decision of whether to release (or hold/gate) a QTL into a downstream WS, made through comparing the downstream WS's current WIP with its accommodable WIP determined by its GF.

Workload Distribution

After determining that a certain amount of lots can be released into the downstream WS, Mixed Integer Programming (MIP) is employed to decide the exact allocation of lots to tools, to minimize the maximum of the assigned workloads to tools within the WS.

OR Formulation of Workload Distribution

Known Parameters:

t_i = *Expected total processing time for lot i, which is a constant*

N_j = *Number of working chambers in tool j*

T_j = *Total processing time for previously assigned lots in W, S, C, and R status in tool j*

Decision Variables:

$x_{ij} = 1$, *if lot i is allocated to tool j, or* 0, *if otherwise.*

$k_{ij} = 1$, *if lot i is qualified to be processed on tool j, or* 0, *if otherwise.*

Objective:

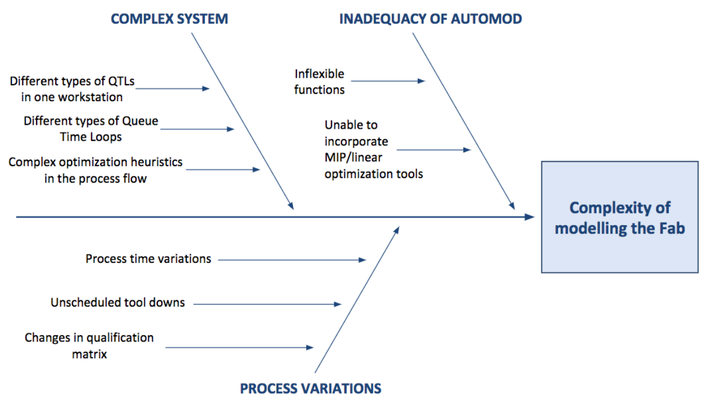
$Min.\ z = max\ \{\sum_{i=1}^{N_j} x_{ij} t_i + T_j, for\ all\ j\}$, *where z is the expected processing time of all lots, in the tool with the highest assigned workload in the current workstation.*

$s.t.\ \sum_{j=1}^{M} x_{ij} = 1$, *for each i where* $1 \le i \le N$ *(each lot i can only be handled by 1 tool)*

$x_{ij} \le k_{ij}$ *(To check if lot i is qualified to be processed on tool j)*

$i, j \ge 1$ *and are integers*

MODELLING COMPLEXITY

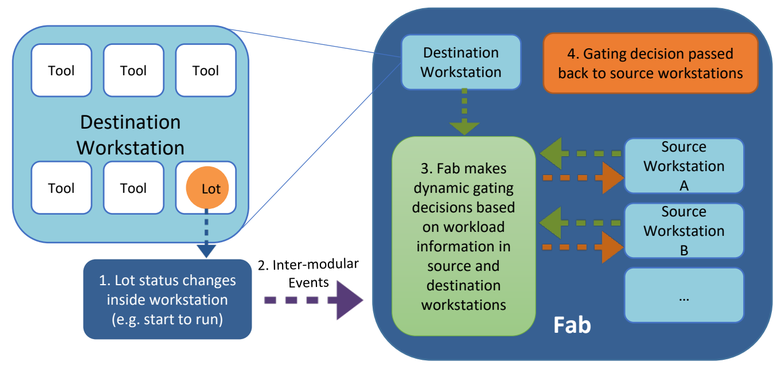


Problem Characteristics

- Sufficiently large scope of simulation required for accurate results
- Large volume of data required for accurate simulation
- Complex algorithms embedded in process flow
- Other optimization steps simultaneously conducted

RESULTS AND DELIVERABLES

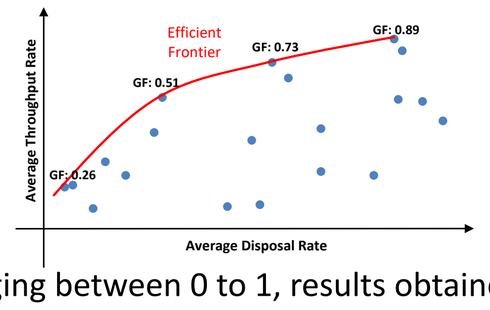
O²DES Micron Fab Simulator



The simulator possesses the following advantages:

- Scalable - simulations can be conducted at full Fab level
- Convenient XML input and output methods
- MIP optimization embedded within algorithms

Efficient Frontier



- For GFs ranging between 0 to 1, results obtained are plotted to determine an efficient frontier
- A set of optimal GFs is obtained for stakeholders' decision-making based on current operational objectives

KEY SKILLS

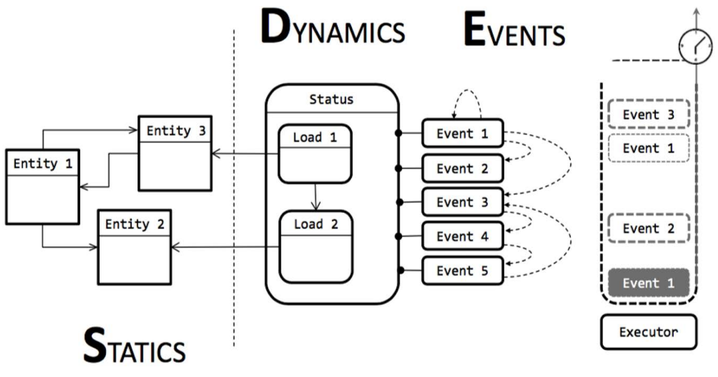
Simulation

Performed multiple simulation runs and applied efficient frontier concept to determine a set of optimal GFs.

Operations Research

Replaced current heuristics method for workload distribution with Mixed Integer Programming to optimally assign workloads to available tools in a WS.

Object-Oriented Programming



Applied OOP under the O²DES Framework to model real-time operation and interaction between a group of workstations.