

## 제공 밸런스와 장비할당을 통합한 반도체 FAB 생산 스케줄링

정재우<sup>1</sup> 장재진<sup>1</sup> 서정대<sup>2</sup> 구평희<sup>3</sup> 정유인<sup>4</sup>

<sup>1</sup>Department of Industrial and Manufacturing Engineering, University of Wisconsin-Milwaukee, 3200 North Cramer St, Milwaukee, WI 53201, U.S.A

<sup>2</sup>경원대학교 산업정보시스템공학과, 461-701 경기도 성남시 수정구 복정동 산65

<sup>3</sup>부경대학교 시스템경영공학과, 608-739 부산시 남구 용당동 산100

<sup>4</sup>삼성전자 시스템기술팀, 445-701 경기도 화성시 태안읍 반월리 산16

### Abstract

웨이퍼 가공 라인에서 좋은 스케줄링은 생산량을 높일 뿐 아니라, 제공 수준을 낮게 유지함으로써 가공 중에 발생하는 불량률의 발견을 신속하게 하고, 라인 내 particle 발생의 원인이 되는 작업자 수를 감소시켜 수율 향상에도 도움을 준다. 이를 위하여 본 연구는 적절한 납기 만족도와 제공 수준을 유지하며 생산량을 최대화하는 스케줄링 방법을 개발한다. 본 연구에서 사용된 MIP 모델은 pre-bounding 방법을 이용해서 풀이 시간을 단축하였고 제공 밸런스와 장비할당에 대한 통합 모델링으로 상호간 원활한 절충이 가능하게 했다. 또한 상위 시스템인 플래닝과의 적절한 연계 방법에 대해서도 연구하였다.

### 1. 서론

그 동안 자동화된 반도체 스케줄러에 대한 많은 연구가 진행되었지만 이들의 품질은 사람에 의한 수동 스케줄링에 미치지 못해 중요한 의사결정에 대한 작업자의 개입이 불가피했다. 스케줄러의 개발에 있어서 어려운 점들은 크게 보아서, 첫째, FAB 이 다양한 특성을 가진 장비들로 구성되어 있고 Reentrant 라는 독특한 구조를 가져 제어가 복잡하다는 점이 있고, 둘째, 스케줄러에 목표로 주어지는 플래닝의 결과가 최적화 되지 못해서 생산 목표의 불확실성이 존재한다는 점이다.

본 연구에서는 high-volume 반도체 FAB 을 위해서 플래닝에 의해 주어진 제품별 목표 생산을 기준으로 throughput 을 최대화하는 스케줄링 방법을 개발한다. 고가의 장비로 구성된 high-volume 반도체 FAB 은 국내 반도체 매출의 90% 이상을 차지하고 있고 전세계 반도체 시장의 약 75% 를 차지하고 있다. 이 FAB 에서는 throughput 최대화는 가장 중요한 성능 평가의 기준이 되고 있다.

본 연구는 3가지의 새로운 개념을 도입하여 FAB 스케줄링의 성능을 향상시킨다. 첫째, 전공정 후공정으로 나누어져 있는 복잡한 반도체 생산 시스템의 속성상 중장기 생산계획을 수립하는 플래닝이 최적화된 목표를 FAB 스케줄링에 제공 할

수 없다. 플래닝에 의해 주어진 생산 목표는 생산성 향상을 위해 스케줄링에서 적절한 조정이 필수적으로 이루어져야 한다. 본 연구에서는 플래닝에 의해 목표로 주어진 결과를 스케줄러에 반영하는 과정을 개선한다.

둘째, 그 동안 많은 연구들에서 FAB 전체의 스케줄링은 주로 제공 밸런싱과 장비할당의 계층적 단계로 다루어왔다. 이러한 계층적 접근 방법은 모델의 단순화와 풀이 과정의 용이함을 제공하지만 상위 단계의 의사결정이 하위 단계의 의사결정을 제한하는 문제점으로 인해 전체 스케줄링 성능의 저하를 초래한다. 본 연구에서는 이 두 단계의 의사결정을 통합해 상호 원활한 절충이 가능하게 했다.

마지막으로, 그 동안 실제 high-volume 반도체 FAB 에서 MIP 를 이용한 스케줄링은 시간의 제약으로 인해 불가능하게 여겨졌다. 본 연구에서는 MIP 모델의 풀이 시간 감소를 위해 pre-bounding 제약을 추가했고 Full-scale FAB (웨이퍼 투입기준 35K) 에서 제안된 모델을 실험을 통해서 검증했다.

### 2. 반도체 생산관리

반도체 제조 공정은 크게 전공정 (FAB 또는 Front-end) 과 후공정 (모듈 또는 Back-end) 으로 나뉘어지고 대략 85% 이상의 자원이 전공정에서 사용된다. 300 개 이상의 단위 공정으로 구성된 전공정은 절연체와 비절연체의 패턴을 가공함으로써 웨이퍼 전자회로를 만들어간다. 알려진 바와 같이 FAB 의 특성은 랜덤 수율, 낮은 장비 유연성, 다양한 장비 특성, Reentrant 구조, 혼류 생산 (Uzsoy, 1992; Fowler, 2003) 등이 있다.

반도체 생산관리 시스템은 플래닝, 스케줄링, 디스패칭의 계층적 의사결정 구조로 되어 있다. 각 모듈은 상위 모듈에서 목표를 받는데 하위 시스템 일수록 자세한 현장 정보를 요구하는 대신 계획 대상 범위 (time horizon) 는 짧다.

플래닝은 전공정과 후공정을 대상으로 제품별, 월별 혹은 주별 목표 생산량을 결정한다. 계산 범위는 수개월에서 일년에 이른다. 이렇듯 폭 넓은 범위를

대상으로 한 의사결정에는 여러 가지 어려움이 수반된다.

최적화된 플래닝을 어렵게 하는 요소들로는 FAB에서의 복잡한 생산구조, 전공정과 후공정 사이의 제품 분기 및 제품 등급, 대체의 발생으로 인한 불확실성, 시장 상황에 따른 제품 수요의 불확실성 등이 있다. 이와 같은 이유로 인해, 두 가지의 상반된 목표 (납기 만족과 throughput 최대화) 사이의 가중치에 대한 결정은 때때로 어려운 과제가 된다.

최적화된 플래닝의 어려움과 FAB 운영의 어려움으로 인해, FAB에는 pull 방식 보다는 push 방식의 생산 모델이 적용되어 왔다. (Lee 2001). Push 방식의 생산 모델에서 플래닝은 스케줄링에게 납기를 내려주는 것이 아니라 주별, 목표 생산 (wafer) 매수를 내려준다.

FAB의 생산 스케줄링은 장비별 공정별 목표 생산 매수를 결정하는데 계획 시간 단위 (time bucket) 는 한 shift 에서 몇시간 정도이다. 웨이퍼 투입 스케줄링도 FAB의 throughput 과 재공 수준에 영향을 주는 중요한 결정이다. 본 연구에서 제안되는 스케줄링은 throughput 최대화와 주어진 생산 목표의 만족이라는 상반되는 목표들에 대해서 최적해를 찾고자 한다.

FAB 이 reentrant 구조이기 때문에 좋은 재공 밸런싱은 장기적으로 좋은 throughput 을 보장한다. 재공 밸런싱은 FAB 공정의 시작에서 끝까지 균일한 재공 상태를 유지하고자 하는 것인데 궁극적으로 장비의 비가동 시간을 줄여 가동률을 향상 시키는데 목적이 있다.

장비 부하 밸런싱은 병목공정의 장비 할당을 위한 매우 중요한 개념이다. 재공 밸런싱이 장기적인 throughput 을 위해 중요한 개념이라면 부하 밸런싱은 현재 shift 의 throughput 을 위한 개념이다. 부하 밸런싱은 동일 공정의 장비들에 대기하고 있는 재공들을 균일하게 하여 장비의 비가동 시간을 최소화하고자 하는 개념이다. <그림 1>에서, 장비 안에 원으로 표시된 색깔의 공정을 각 장비에서 처리 할 수 있는 웨이퍼의 종류라고 한다면, 3대의 장비 중에서 M2 는 회색 대신 검은색 웨이퍼를 선택해야 한다. 그렇지 않으면, M1 이 곧 처리할 물량을 확보하지 못하게 될 것이고 검은색으로 표시된 종류의 웨이퍼들은 대기 시간이 길어질 것이다.

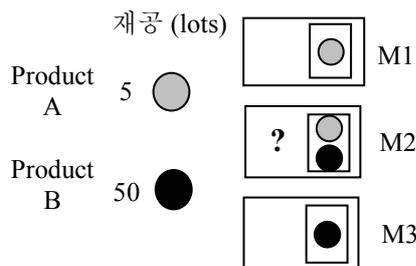


그림 1. FAB 에 있어서 장비 부하 발란싱

비록 FAB 스케줄링이 복잡한 생산 환경 하에서 복수의 목적을 최적화해야 하는 어려운 문제이지만, 스케줄링의 초점을 병목 공정인 포토 공정에 둠으로써 문제를 용이하게 접근할 수 있다. 모든 layer 들은 포토 공정을 거쳐야 하고 절대적인 병목 공정이기 때문에 포토의 운영은 전체 라인의 효율을 결정한다. 그러므로 다른 공정들이 비교적 단순한 절차로 제어되는 반면, 이 공정을 위해서는 가장 정교한 제어가 필요하다.

### 3. 기존 연구

80년대부터 90년대 초반까지 반도체 스케줄링에 대한 연구는 주로 lot 투입, 병목 스케줄링, 묶어서 하기 (batching), 납기 스케줄링 등 단위 기능에 초점이 맞추어져 있었다. 90년대 중반 이후에 Uzsoy 등 (1992, 1994)에 의해 반도체 생산관리의 어려움과 특징들이 정확히 진단 되기 시작하면서 이러한 단위 기능들을 통합해서 다루는 연구가 보편화 되었다. 이 연구에서 자주 사용되는 방법이 계층적 접근 방법 (Hierarchical Approach) 이다.

Lot 투입 방법은 주로 공기와 생산량을 이용하여 병목 공정의 굵는 현상을 최소화하는데 초점이 맞추어져 왔다. (Glassey 와 Resende, 1988; Wein, 1988; Lou 와 Kager, 1989). 병목 공정 스케줄링을 위해서 많은 연구들이 수리적인 방법과 휴리스틱을 이용한 방법으로 연구를 진행했는데, 주로 마스크 사용, 준비 교체 시간 등 공정의 미세한 모델링에 대해 연구했다. (Lu 등, 1994; Sohl 와 Kumar, 1995; Liao 등, 1996; Akcali 와 Uzsoy, 2000; Kim 등, 2002).

묶어서 하기 (batching) 는 확산, 산화, hardbake 등의 공정에 주로 사용되는데 Glassey 와 Weng (1991), Fowler 등 (1992), Weng 과 Leachman (1993), Robinson 등 (1995), Saito 와 Arima (2002) 등에 의해 연구되었고 look-behind 와 look-ahead 접근 방법이 자주 이용되었다. 납기만족에 대한 연구는 Lu 와 Kumar (1991), Kim 등 (1998), Sloan (2003), Kim 등 (2003) 에 의해 진행되었는데 파운드리 업체나 ASIC FAB 이 주로 대상이다.

통합된 스케줄링을 연구 대상으로 많은 연구들이 재공 밸런싱과 병목공정 할당이라는 계층적 방법을 이용하고 있다. TB (Two Boundary) 방법은 재공 밸런싱 방법으로 Hung (1996), Kim (2001), Dabbis (1999), Dabbis (2001), Dabbis (2003), Lee (2002), Pai (2004), Shen (2003) 등을 비롯해서 많은 연구에서 사용되었다. TB 는 Lou 와 Kager (1989) 에 의해 처음 FAB 스케줄링에 도입되었다. 이후, Houmin et al (1996) 은 TB 방법의 효율에 대한 추가로 검증했다

Vargas-Vilamil 등 (2003) 은 복잡한 FAB 스케줄링을 위해서 3 개의 계층적 단계로 이루어진 접근 방법을 도입했다. 첫번째 단계는, 여러 개의 제품을 하나의 제품으로 묶는 단계이고, 두번째 단계는 최적화 단계로써, 장비 가동률과 생산량

목표를 결정하는 단계이다. 마지막으로 디스패칭 로직을 이용해서 lot 을 직접 제어한다. 이 연구는 복잡한 FAB 의 스케줄링을 다루는 방법에 대한 좋은 아이디어를 제공하기는 하지만, 여러 제품을 하나의 제품으로 모델링 함으로써 제품 준비교체 시간이나 묶어서 하기 등의 미세한 모델링을 어렵게 한다.

위의 대부분의 연구에서는 플래닝과 스케줄링의 연계 방법을 중요하게 다루지 않았다. 플래닝에서 주어진 목표는 미리 최적화 된 것으로 간주하고 일 단위 또는 Shift 단위로 그 목표를 만족시키려고 하고 있다. 하지만 실제 FAB 에서 플래닝은, 전후 공정구조, 복잡한 FAB 제어 구조, 불안정한 시장 상황 등의 어려움들을 안고 있어서 최적화된 플래닝은 가능하지 않으며 플래닝에 의해서 주어진 목표는 스케줄링에서 적절한 조정이 이루어져야 한다. 그렇지 않으면, 플래닝에서 주어진 목표가 스케줄링을 지나치게 제한함으로써 전체적 생산성의 저하를 초래한다.

#### 4. 새로운 스케줄링 방법

##### 4.1. 문제 설명

- $i$ : Device 인덱스 ( $1 \leq i \leq I_j$ )
- $j$ : Layer 인덱스 ( $1 \leq j \leq J_i$ )
- $k$ : 현재 Layer 인덱스, ( $1 \leq k \leq J_i$ )
- $t$ : Shift 인덱스
- $m$ : 포토 장비 인덱스 ( $1 \leq m \leq M$ )
- $w$ : 주 (week) 에 대한 인덱스
- $C_{ij}$ : ( $i, j$ )의 공기, 대기시간 포함
- $C_{ij}^*$ : ( $i, j$ )의 누적 공기
- $K_{ij}$ : ( $i, j$ )의 가용 마스크 수
- $X_{ij}$ : ( $i, j$ )의 목표 생산 매수
- $X_{ijm}$ : Device  $i$ , layer  $j$ , 장비  $m$  의 할당 매수
- $X_{ijm}^c$ : Binary 인덱스, Shift 시작 할 때 장비  $m$  이( $i, j$ ) 를 생산하고 있었으면  $X_{ijm}^c = 1$ , 그렇지 않으면  $X_{ijm}^c = 0$  이 됨
- $P_{ij}$ : ( $i, j$ )의 공정 시간
- $Q_{ij}$ : ( $i, j$ )의 Shift 초 재공
- $Q_{ij}^*$ : ( $i, j$ )의 Shift 말 재공
- $D_{ij}$ : ( $i, j$ )의 평균 목표 생산 매수 (주별 목표 생산 매수에서 구해짐)
- $D_{ij}^p$ : RPH 방법에서 구해진 ( $i, j$ ) 의 목표 생산 매수
- $L_{ij}$ : ( $i, j$ )의 lateness 허용 한계.
- $O_{ij}$ : Device 의 마지막 공정 지수,  
 $j=J$  일 때  $O_{ij} = 2$  그렇지 않으면  $O_{ij} = 1$
- $Y_{ijm}$ : ( $i, j$ ) 의 장비  $m$  에서 가용 여부.
- $Y'_{ijm}$ : Binary 변수, ( $i, j$ ) 가  $m$  에 할당되었으면 1 그렇지 않으면 0 이 됨.

$Y'_{ij}$ : ( $i, j$ ) 의 할당된 장비 수 (정수 변수임),

$$Y'_{ij} = \sum_{m=1}^M Y'_{ijm}$$

$CP_m$ : 장비  $m$  의 한 Shift 동안 생산 능력

본 연구에서는 재공 밸런싱과 병목 공정 할당 기능을 통합하고 적절한 제품별 생산 목표량과 재공 수준 하에서 throughput 을 최대화 할 수 있는 스케줄링 방법을 연구한다. 이 모델에서 사용하는 입력 자료로는: device 별 ID, layer ID, 공기, 현재 재공, 포토 장비 마스크 수, 장비 공정 시간, 현재 생산중인 device-layer 등이 있다. 하나의 layer 는 증착, 식각, 검사, 확산, 포토 등의 공정을 통합한 것으로 매 layer 의 마지막에 하나의 포토 공정이 있다. 이 모델에서는 포토 공정 이외에서는 장비에 대한 상세한 모델은 하지 않는다. 이 모델에서는 lot 투입, lot 별 스케줄링 등은 고려하지 않지만, 이것들은 일단 포토 공정의 스케줄이 결정되면 forward 스케줄링 방식으로 쉽게 결정할 수 있다. 이 모델의 최종 결과로는 한 shift 동안의 포토 각 장비에 대한 제품 layer 별 목표 생산 매수이다.

##### 4.2. 스케줄링 접근 방법

제안된 새로운 스케줄링 방법은 두 개의 계산 단계로 되어 있다. 첫번째로, 플래닝에 의해 주어진 주별 목표 생산 매수를 바탕으로 layer 별 목표 생산 매수와 웨이퍼들의 평균 lateness 가 계산되는데, RPH (Rolling and Pulling Horizon) 방법을 이용한다. lateness 는 각 layer 의 재공들이 공기와 비교했을 때 목표 생산 매수를 맞추기 위해서 얼마나 앞서 있거나 혹은 뒤쳐져 있는지를 계산한다. RPH 방법은 TB 방법을 개선한 것으로 pulling 기간에 해당되는 평균 목표 생산 매수가 더해진다는 점이 기존의 TB

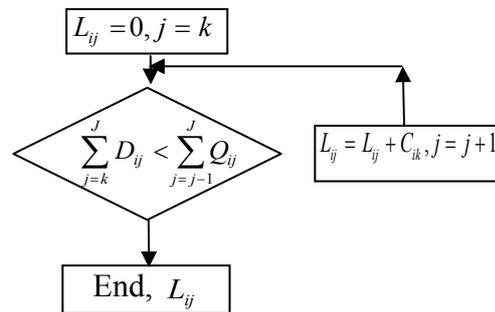


그림 2. lateness 계산 흐름도

방법과는 다른 점 이다. 두번째, 새로운 MIP 모델에서 가능한 재공을 장비에 할당하는데 RPH 에 계산된 layer 별 목표와 lateness, 포토 장비의 부하가 주로 고려된다.

새로운 모델은 기존 계층적 접근 방법과는 다르게 재공 밸런싱과 장비 할당을 두번째 단계에서 한꺼번에 처리한다. 즉, 첫번째 단계에서 계산된 layer 별 목표

생산 매수는 실제 포토 작업 능력보다 훨씬 커서 두번째 장비 할당 단계에서 큰 제약이 되지 않고 단지 상한선 역할만 한다.

본 모델에서는 재공 밸런싱과 병목 장비 할당의 통합을 위해서 새로운 개념의 재공 밸런싱 개념을 도입한다. 기존 연구들은 재공 밸런싱을 위해서 layer 별 목표 생산 매수를 주로 이용했다. 본 연구에서는 목표 생산 매수 대신 layer 별 재공의 유연성을 고려한 장비별 부하를 계산하여 재공 밸런싱의 지표로 이용한다. Layer 별 목표 생산 매수를 이용하지 않고 장비별 부하를 재공 밸런싱 지표로 이용하는 것이 장비의 가동률을 높이려는 재공 밸런싱의 본래 목적에 더 잘 부합할 것이다.

새롭게 제안된 MIP 모델은 위의 재공 밸런싱 방법과 장비 할당 방법을 동시에 고려한다. 장비 할당 방법은 주로 throughput 최대화를 추구한다. 그래서 MIP 모델은 재공 밸런싱과 throughput 최대화라는 두개의 결정 변수로 구성되어 있다.

이 MIP 모델에서 IP 변수의 숫자는 device 수, device 별 layer 수와 장비수의 곱과 같다. FAB의 상황에 따라서 IP 변수 숫자가 증가해서 때때로 풀이 시간이 문제가 될 수 있는데 이것을 해결하기 위해서 본 연구에서는 pre-bounding 방법을 이용한다. pre-bounding 방법은 스케줄링 풀이 시간을 줄이기 위해 결정 변수의 범위를 MIP 풀이 이전에 제한하는 방법이다. 본 연구에서와 같이, 현장의 제약, 개략적인 해를 구하는 heuristic 계산 방법을 이용하는 것도 가능하고 또한 현장 관리자의 직감 등을 통해 pre-bounding 제약을 만들 수도 있을 것이다.

#### 4.3. RPH (Rolling and Pulling Horizon) 방법: layer의 목표 생산 매수 결정

RPH 방법은 다음의 두 단계로 구성된다. (1) 주간 목표 생산 매수의 layer 별 분배 (2) layer 별 목표 생산 매수의 결정. <표 1>은 RPH를 이용한 layer 별 생산 매수의 결정 과정을 보여주고 아래의 식과 같이 표현된다.

단계 1: 주간 목표의 layer 별 분배

$$D_{ij} = C_{ij} \times D_w,$$

만약  $7(w-1) < C_{ij}^* < 7w$

$$D_{ij} = \left(\frac{C_{ij1}}{C_{ij}} \times D_w\right) + \left(\frac{C_{ij2}}{C_{ij}} \times D_{w+1}\right),$$

만약,  $7(w-1) < C_{ij}^* < 7w$   
 $C_{ij1}$ 에서  $l$ 은  $C_{ij}$ 에서  $l$ 의 비율.

단계 2: layer 별 목표 생산 매수의 결정

$$D_{ij}^p = \sum_{j=k}^J D_{ij} + (\text{Two days' demand}) - \sum_{j=k+1}^J Q_{ij}$$

#### 4.4. RPMIP (Rolling and Pulling horizon with MIP) 모델링: 포토 장비의 한 shift 스케줄 계산 과정

(1) 장비별 가용 재공 (MAW) 계산

$WPM_{ij}$ : 재공 per machine

$$WPM_{ij} = \frac{Q_{ij}^*}{\sum_{m=1}^M Y_{ijm}}$$

$MAW_m$ : Machine Available WIP

$$MAW_m = \sum_{i=1}^I \sum_{j=1}^J (WPM_{ij} \times Y_{ijm})$$

$AAW$ : Average Available WIP

$$AAW = \frac{\sum_{m=1}^M MAW_m}{M}$$

여기서,  $MAW_m - AAW$ 가 재공 밸런싱 변수로 이용된다.

(2) lateness 계산

<그림 2>는 lateness 계산 과정을 보여준다. 만약,  $(i, j)$ 의 누적 분배된 주간 목표가 누적 현재공 보다 작으면  $C_{ik}$ 만큼 더해지고  $k$ 를 증가시켜 반복 계산한다. 누적 분배된 재공이 현재공 보다 많으면 계산을 중지한다.

(3) 목적 함수

새로운 모델의 목적함수는 두 가지의 서로 다른 결정 변수를 갖고 있는데 하나는 현 shift의 생산 매수를 최대화하기 위한 것이고 두번째는 나중 shift의 throughput을 위해서 재공 밸런싱을 유지하기 위한 것이다. 이 두 결정 변수 사이의 가중치(아래식의 M과 N)는 실험을 통해 결정될 수 있다.

$$\text{Min } Z = M \sum_{i=1}^I \sum_{j=1}^J (D_{ij}^p - X_{ij} \times O_{ij}) - N \sum_{m=1}^M (MAW_m - AAW)$$

(4) 제약 조건

● 가용 재공 제약:

$$X_{ij} \leq Q_{ij}$$

표 1. Layer 별 목표 생산 매수 계산 과정

Layer ID	Layer 별 목표	Layer 별 누적 목표	Pulling 기간의 목표	현재공	누적 현재공	Layer 별 목표 생산 매수
1	800	6600	500	700	5600	1500
2	800	5800	1000	1400	4200	2600
3	1500	5000	1600	800	3400	3200
4	1500	3500	2200	900	2500	3200
5	1000	2000	2800	1500	1000	3800
6	1000	1000	1100	1000	0	2100

- 준비 교체시간 고려를 위한 변수:  
 만약  $X_{ijm}$  이 0 보다 크면,  $Y'_{ijm}$  이 되고 1, 그렇지 않으면, 0 이 된다.  $Y'_m$  는 정수로써 장비  $m$  의 준비 교체 회수가 된다.

$$X_{ijm} \leq MY'_{ijm}$$

$$X_{ijm} \geq Y'_{ijm}$$

$$Y'_m = \sum_{i=1}^I \sum_{j=1}^J Y'_{ijm}$$

- 장비의 생산 능력 제약:

$$\sum_{i=1}^I \sum_{j=1}^J (X_{ijm} \times P_{ijm}) + (Y'_m - 1) * SetupTime \leq 8hr$$

- Shift 말 재공을 계산을 위한 수식:

$$Q_{ij}^* = Q_{ij} + X_{ij-1} - X_{ij} \quad (j \neq 0)$$

$$Q_{i0}^* = R_i$$

- 목표 생산 매수 제약:

$$X_{ij} \leq D_{ij}^p$$

- 가용 마스크 수 제약:

$$Y'_{ij} \leq K_{ij}$$

- lateness 제약:

$$Y'_{ij} \geq 1, \text{ if } L_{ij} \geq 2$$

- 가용 장비 제약:

$$Y'_{ijm} = 0, \text{ when } Y_{ijm} = 0$$

- 현재 진행 중인  $(i, j)$  (device, layer) 고려:

$$Y_{ijm} = 1, X_{ij}^c = 1, Q_{ij} > 0 \text{ 일 때}$$

- 비 음수 제약

$$X_{ijm} \geq 0 \text{ for all } i, j, m$$

$$Q_{ij}^* \geq 0 \text{ for all } i, j$$

(5) pre-bounding 제약 추가

- 한 장비의 할당 layer 수 제한

같은 장비에서 생산 될 수 있는 device 와 layer 의 개수는 생산될 수 있는 lot 의 숫자 보다 적어야 한다. 통상의 실제 FAB 에서는 한 장비에 생산되는

layer 수가 10 개를 넘지 않는데 풀이 시간의 단축이 중요해 지면 MB (Minimum Batch) 의 크기를 늘려 할당 되는 종류를 더 줄일 수 있다.

$$Y'_m \leq (H / P_m^*) / MB$$

- 초과 재공을 가진 layer 강제 할당

만약 어떤  $(i, j)$  의 재공이 한 장비의 생산능력의 3 배가 넘는다면 제약을 통해서 적어도 한대의 장비가  $(i, j)$  에 할당되게 함.

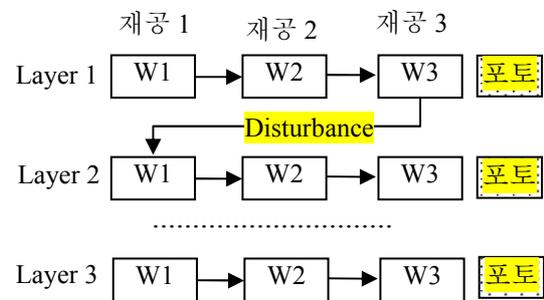


그림 3. 성능 평가를 위한 재공 모델링

$$Y'_{ij} \geq 1 \text{ where } Q_{ij} / (H / P_{ij}^*) \geq 2$$

5. 성능 실험

5.1. 실험 모델

새롭게 제안된 모델의 성능을 평가하기 위해서 웨이퍼 투입 기준 35 K FAB 모델을 바탕으로 테스트 프로그램이 개발 되었다. 이 모델에는 50 대의 포토 장비, 4 개의 서로 다른 생산 매수를 가진 device 와 device 당 25 개의 layer 수가 있다.

테스트에서는 새롭게 제안된 RPMIP 모델과 계층적인 접근 방법을 이용한 두개의 모델을 비교하였다. 두개의 모델은 각각 TBH (TB and Heuristic) 와 TBMIP (TB and MIP) 모델인데 두

모델 모두 재공 밸런싱 방법으로 TB 방법을 이용한다. 병목 공정 할당에 있어서 TBH는 휴리스틱 방법을 이용하고 TBMIP 방법은 MIP를 이용한다. 휴리스틱 방법은 가장 큰 재공을 선택해서 가장 가동률이 낮은 장비에 할당하는 방식을 사용했다.

연속적인 21개의 shift에 대해 실험을 진행하였는데, 한 shift에서의 스케줄링 결과가 확률 분포(유니폼 분포)에 따라 생산되어 다음 shift의 입력 자료로 이용되었다. 연속된 shift간의 반복 실험은 주로 재공 밸런싱의 성능을 평가하기 위함이다. 좋은 재공 밸런싱은 shift가 지나도 throughput이 높게 유지되게 한다.

<그림 3>은 테스트 프로그램에서 사용한 재공 흐름에 대한 모델링을 설명한다. 한개의 layer는 3개의 소그룹의 공정들로 구성되어 있고 그림에서 재공 1과 재공 2는 한 shift 후에 재공 2와 재공 3으로 각각 넘어간다. 재공 3에 있는 재공이 실제 포토에서 가용한 재공인데 실제 생산 예측 매수만큼 다음 shift에서 다음 layer의 재공 1로 넘어간다. device 별로 첫 layer의 재공 1은 매 shift 투입 매수와 같은데 본 테스트에서는 device 별 재공 수준과 생산량을 고려하여 FAB 재공이 일정하게 유지될 수 있게 일정하게 투입했다.

다양한 FAB 모델에 대한 신뢰성을 검증하기 위해서 서로 다른 4개의 장비 유연성 수준과 장비 고장률 수준 하에서 실험을 진행하였다. 유연성 수준은 layer 당 진행 가능한 포토 장비의 숫자에 의해서 결정되었고 고장률 수준은 Quinn과 Bass(1999)에 의해 작성된 세마텍의 자료를 참고하였다. 가동률 수준이 100%이고 유연성이 최대인 수준은 default 수준으로 정의되었다.

실험에 필요한 MIP solver는 CPLEX 8.2, C++ 버전을 이용하였다. Solver에서 제공하는 기능(EpGap)을 이용하여 최적값의 오차를 lot(25매) 단위로 하였다.

5.2. 실험 결과

<표 2>의 실험 결과는 포토 장비의 가동률을 3개의 모델에 대해서 비교한 자료이다. 이

실험 모델에서 모든 포토 장비가 100% 가동되었다고 가정할 때 9,943매/shift의 웨이퍼 생산 가능하다.

이 표는 중간 수준의 장비 유연성과 고장률을 바탕으로 작성되었다. 표에 나타난 바와 같이 RPMIP 방법이 가동률 기준으로 TBH에 비해서 약 6%, TBMIP에 비해서 1.3% 높게 나타나고 있다. 1.3%는 한달 생산 웨이퍼 매수 기준으로 약 450매인데 이는 고가의 장비로 구성된 FAB에서는 연간 수백억원의 추가적인 매출 증가를 기대할 수 있는 수치이다. <표 2>에서 한가지 더 주목할 점은, 병목 장비 할당에 있어서 MIP 방법(TBMIP, RPMIP)을 이용한 방법이 휴리스틱을 이용한 방법(TBH)보다 두드러지게 좋은 성능을 보이고 있다는 것이다.

<그림 4>는 장비의 유연성 수준이 중간 수준으로 고정된 상태에서 고장률 수준에 따른 각 스케줄링 모델별 throughput의 차이를 보여주는 그래프이다. throughput은 한 shift 동안에 포토 장비들에 할당된 웨이퍼 매수의 합이고 21개 shift 평균이다. 3개의 모델 모두에서, 고장률 수준이 높아 질수록 throughput이 감소하는 가운데 RPMIP 모델이 가장 적은 감소율을 보인다. 이는 서로 다른 고장률 수준에서도 RPMIP 모델이 다른 모델에 비해서 좋은 재공 밸런싱과 병목공정 할당의 성능을 보인다는 것을 입증한다.

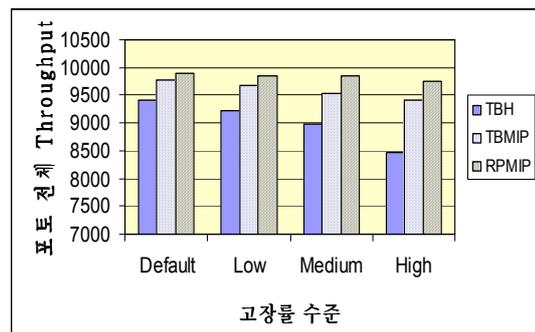


그림 4. 고장률 수준에 따른 throughput

<그림 5>는 서로 다른 장비의 유연성 수준 하에서 각 스케줄링 모델의 throughput 성능을 비교하였다. 이때 고장률 수준은 중간으로 고정되었다. 위의 다른 고장률 수준에서의 결과와 비슷하게 이번 실험 결과에서도 RPMIP 모델이 가장 좋은 성능을 보였다. 특히 장비 유연성이 낮을수록 모델별로 더 큰 성능 차이가 나는 것을 볼 수 있는데 이는 장비의

표 2. 실험 결과-포토 장비의 가동률 (21개 shift 평균)

스케줄링 모델	TBH		TBMIP		RPMIP	
	매수	가동률	매수	가동률	매수	가동률
스케줄링 결과	9014	90.7%	9530	95.8%	9671	97.3%
생산 예측 매수	8743	87.9%	9244	93.0%	9381	94.3%

스케줄링 결과는 포토 장비에 할당된 웨이퍼 매수의 합이고 생산 예측 매수는 스케줄링 결과에 유니폼 분포(0.74, 1.23) 분포를 곱해서 계산된 값이다. 이는 약 90%의 장비 가동률과 같다.

유연성이 낮을수록 부하 밸런싱과 재공 밸런싱이 더욱 중요해진다는 것을 보여준다. 특히, 최근의 나노 반도체와 같이 공정이 까다로울수록 장비의 유연성이

낮아질 수 밖에 없고 최적화된 장비 스케줄링의 중요성은 더욱 강조 될 것이다.

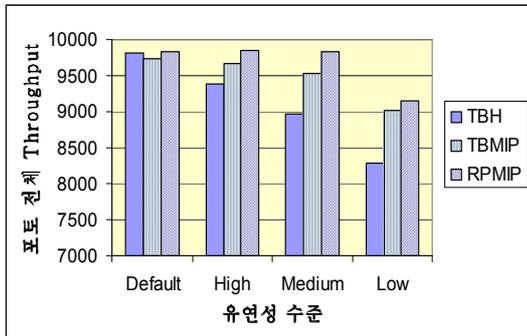


그림 5. 장비 유연성 수준에 따른 throughput

이 실험과 관련해서 한가지 더 주목할 만한 점은, TBMIP 와 RPMIP 모델의 경우 낮은 유연성 수준에서 갑자기 성능이 저하된다는 것이다. 이는 실제 FAB에서 장비의 유연성을 관리하는 임무가 중요하다는 것을 보여준다. 따라서, 유연성에 따른 throughput 영향을 진단하고 적절한 수준을 찾아내는 것도 추후 흥미로운 연구가 될 것이다.

FAB 이 납기를 기준으로 생산하는 Make-To-Order 방식이 아니기 때문에 납기 만족에 대한 정확한 성능 평가에는 다소의 어려움이 있다. 본 연구에서는 납기 만족에 대한 성능을 평가하기 위해서 lateness 를 이용하여 공기보다 늦게 진행되고 있는 웨이퍼의 매수를 합하여 평가 기준으로 하였다.

플래닝에서 주어진 목표 생산량 대비 진도가 2 일 이상 늦은 재공의 합은 TBMIP 에서는 전체 재공 중 0.65% 였고 RPMIP 에서는 2.4% 였다. RPMIP 에서 다소 높게 나타났으나 이는 실제 FAB 관점에서 본다면 여전히 낮은 상태이다. 앞에서도 설명했듯이 플래닝에서 내려 온 목표는 FAB 의 생산성을 위해서 때때로 조정이 요구된다. 따라서 2.4%로의 증가는 이런 조정의 결과로 볼 수 있을 것이다.

Pre-bounding 방법의 성능을 평가하기 위해서 pre-bounding 제약을 추가했을 때와 제거했을 때의 계산 소요시간을 비교했는데 pre-bounding 방법으로 계산 시간과 최적화 수준이 같이 향상 되었음을 알 수 있었다. Pre-bounding 제약을 추가하기 전에는 15 분 계산 후 IP 해와 LP 해의 상한 (Upper Bound) 사이에 평균 1.8% 의 gap 이 있었다. 제약을 추가한 후에는 약 10 분의 계산시간 후에 gap 의 평균이 0.35 % 로 줄었고 많은 실험의 경우에는 15 분 동안에 IP 최적해에 도달 할 수 있었다. 15 분 동안의 스케줄링 시간은 일반적으로 한 shift 의 스케줄링을 위해 받아 들여지고 있는 시간이다.

## 6. 결론 및 추후 연구 방향

위에서 살펴 본 바와 같이 본 연구에서는 3 가지의 새로운 접근 방법을 이용하여 FAB 의 throughput 을

향상시키는 스케줄링 방식을 제안하였다. 첫째, 플래닝에서 내려 온 목표 생산량을 스케줄러에 반영하는 방법에 대해서 새로운 방법을 제안하였다. 둘째, 병목공정 장비의 부하를 고려하는 새로운 개념의 재공 밸런싱 방법을 도입하였고 재공 밸런싱과 병목공정 할당의 기능을 통합한 MIP 모델링 방법을 제안하였다. 마지막으로, MIP 모델링의 계산 시간을 줄이기 위해서 pre-bounding 방법을 도입하였다.

추가적인 FAB 스케줄러의 성능 향상을 위해서 최적화 관점에서의 성능 향상 외에도 FAB 에 대한 미세한 모델링이 필요하다. 이와 같이 미세함을 추가적으로 필요로 하는 것들로써 비정규 웨이퍼 (Nonproductive Wafer, NPW) 의 처리, maintenance 를 고려한 스케줄링, 완전 자동화를 위한 lot 단위 스케줄링 등이 있다. 미세한 상황을 고려한 스케줄링은 작업자 개입의 필요성을 감소시키고 원활한 시스템적 운영을 가능하게 한다는 점에서 지속적으로 추진 되어야 할 과제이다. 이를 위해서는 모델링의 유연성과 최적화 품질을 동시에 개선해야 할 것이다.

## Acknowledgement

이 논문은 2004 년도 한국학술진흥재단의 지원에 의하여 연구되었음. (KRF-2004-042-D00226)

## 참고 문헌

- Alkali, E., Uzsoy, R. (2000) "A sequential solution methodology for capacity allocation and lot scheduling problems for photolithography," Electronics Manufacturing Technology Symposium (Twenty-Sixth IEEE/CPMT International), 374 - 381.
- Fowler, J.W., Hogg, L. G. and Mason, C. (2002) "Workload control in the semiconductor industry," Production planning and control, v 13, n 7, 568-578.
- Fowler, J. W., Hogg, G. L. and Phillips, D. T. (1992) "Control of multiproduct bulk service diffusion/oxidation processes," IIE Transactions, v 24, n4, 84-96.
- Glassey, C. R. and Resende, M. G. C. (1988) "Closed-loop job release control for VLSI circuit manufacturing," IEEE Transactions on semiconductor manufacturing, v1, n1, 36-46.
- Glassey, C. R., Weng, W. W. (1991) "Dynamic batching heuristic for simultaneous processing," IEEE transactions on semiconductor manufacturing, v4, n2, 77-82.
- Hung, Y. F. and Leachman, R. C. (1996) "Production planning methodology for semiconductor manufacturing based on iterative simulation and linear programming calculations," IEEE Transactions on Semiconductor Manufacturing, v 9, n 2, 257-269.
- Kim, S., Yea, S.H. and Kim, B. (2002) "Shift scheduling for steppers in the semiconductor wafer fabrication process," IIE Transactions, v34, 167-177.
- Kim, Y.D., Choi, O.S. and Hwang, H. (2003) "Simplification methods for accelerating simulation-based real-time scheduling in a semiconductor wafer fabrication facility," IEEE Transactions on semiconductor manufacturing, v16, n2, 290-298.
- Kim, Y.D., Kim, J.U., Lim, S.K. and Jun, H.B. (1998) "Due-date based scheduling and control policies in a multi-product semiconductor

- wafer fabrication facility," IEEE Transactions on semiconductor manufacturing, v11, n1, 155-164.
- Lee, Y. H. (2001) "Supply chain model for the semiconductor industry," Journal of systems integration, v 10, n 3, 189-206.
- Lee, Y.H., Park, J. and Kim, S. (2002) "Experimental study on input and bottleneck scheduling for a semiconductor fabrication line," IIE transactions, v34, 179-190.
- Liao, D. Y., Chang, S. C., Pei, K. W. and Chang C. M. (1996) "Daily scheduling for R&D semiconductor fabrication," IEEE transactions on semiconductor manufacturing, v 9, n 4, 550-561.
- Lou, S. X. C. and Kager, P. W. (1989) "A Robust production control policy for VLSI wafer fabrication," IEEE Transactions on semiconductor manufacturing, v2, n4, 159-164.
- Lu, S. C. H., Ramaswamy, D. and Kumar, P. R. (1994) "Efficient scheduling policies to reduce mean and variance of cycle-time in semiconductor manufacturing plants," IEEE Transactions on semiconductor manufacturing, v7, n3, 374-388.
- Lu, S.H. and Kumar P.R. (1991) "Distributed scheduling based on due dates and buffer priorities," IEEE Transactions on automatic control, v36, n12, 1406-1416.
- Pai, P. F., Lee, C. E. and Su, T. H. (2004) "A daily production model for wafer fabrication," International journal of advanced manufacturing technology, v 23, 58-63.
- Quinn, T. and Bass, E., (1999) "300 mm Factory Layout and Material Handling Modeling: Phase I Report," International SEMATECH, Technology Transfer # 99023688B-ENG.
- Robinson, J. K., Fowler, J.W. and Bard, J. F. (1995) "The use of upstream and downstream information in scheduling semiconductor batch operations," International journal of production research, v3, n7, 1849-1869.
- Saito, K., Arima, S. (2002) "A simulation study on periodical priority dispatching of WIP for product-mix fabrication," IEEE/SEMI advanced semiconductor manufacturing conference, 33-37.
- Shen, Y. and Leachman, R.C. (2003) "Stochastic wafer fabrication scheduling," IEEE transactions on semiconductor manufacturing, v16, n1, 2-14.
- Sloan, W. T. (2003) "Shop-floor scheduling of semiconductor wafer fabs: exploring the influence of technology, market, and performance objectives," IEEE Transactions on semiconductor manufacturing, v 16, n 2, 281-289.
- Sohl, D. L. and Kumar, P. R. (1995) "Fluctuation smoothing scheduling policies for multiple process flow fabrication plants," IEEE/CPMT international electronics manufacturing technology symposium, 190-198.
- Uzsoy, R., Lee, C.Y. and Martin-vega, A. L., (1992), (1994) "A review of production planning and scheduling models in the semiconductor industry part1: system characteristics, performance evaluation and production planning," IIE Transactions, v24, n4, 47-60.
- Uzsoy, R., Lee, C.Y. and Martin-vega, A. L., (1994) "A review of production planning and scheduling models in the semiconductor industry part2: Shop-floor control," IIE Transactions, v 26, n 5, 44-55.
- Vargas-Vilamil, F. D., Rivera, D. E. and Kempf, K. G. (2003) "A hierarchical approach to production control of reentrant semiconductor manufacturing lines," IEEE Transactions on control systems technology, v11, n4, 578-87.
- Wein, L. M. (1988) "Scheduling semiconductor wafer fabrication," IEEE Transactions on semiconductor manufacturing, v1, n3, 115-130.
- Weng, W. W., and Leachman, R. C. (1993) "An improved methodology for real-time production decisions at batch-processing work stations," IEEE Transactions on semiconductor manufacturing, v6, n3, 219-225.