

半导体芯片制造概述

1.1 概 述

1.1.1 半导体产业概况

半导体产业是信息技术产业群的核心和基础。建立在半导体技术进步基础上的全球信息化、网络化和知识经济的飞速发展,使半导体产业的战略地位越来越重要。其技术水平和发展规模已经成为衡量一个国家经济发展和科技进步的重要标志,对国家的综合实力产生巨大的影响和推动作用。有关数据表明:发达国家的国内生产总值(GDP)每增长 100 元,就需要 10 元左右电子工业产值和 1~2 元的半导体产值的支持^[5]。根据美国半导体协会(SIA)预测,到 2012 年,集成电路全行业销售额将达到 5 000 亿~6 000 亿美元,它将支持 6 万亿到 8 万亿美元电子装备和 30 万亿美元电子信息服务业。21 世纪的未来经济是信息经济,目前发达国家信息产业产值已占国民经济总产值的 40%~60%,国民经济总产值增长部分的 65%与半导体产业有关。因此,发展半导体产业是推动国民经济信息化的重要保证,是信息产业发展的重中之重。

截至 2006 年,半导体产业已经历了连续 5 年的增长,从 2007 年的预测数字来看,未来两年还将保持增长势头,如果这一预测准确,那么半导体产业将经历自 1980 年后最为持续的增长阶段。这对整个半导体产业来说是不同寻常的。在 2007 年以前的 7 年中,全球晶圆制造厂(fab)的产能已提高了 2 倍。虽然日本目前还是全球最大的半导体制造国,但从发展来看,随着轻晶圆(fab-lite)策略的实施和集成器件生产商(Integrated Device Manufacturer, IDM)与代工厂

(Foundry)更多的合作,中国大陆、中国台湾和东南亚已成为全球半导体产业的主要增长地区,中国大陆地区 Fab 的产能比例已从 2000 年全球的 1%增长到 2007 年的 8%(图 1-1)。

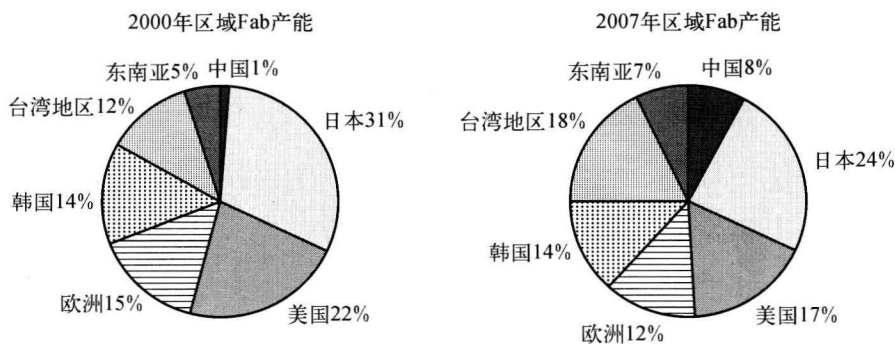


图 1-1 Fab 产能最近 7 年的分布变化

资料来源: SMA World Fab Watch, January 2001 and April 2007.

随着中国的半导体产业的飞速发展,北京、天津、上海、广东、成都、西安等地已相继成立了若干半导体晶圆制造、芯片封装制造基地。以上海为龙头的长江三角洲地区,已经形成了以华虹 NEC、Intel、中芯国际、台积电、宏力半导体等国际半导体生产集团为首的芯片制造加工密集区。2007 年中国大陆正式投产的晶圆制造企业有 9 家,如表 1-1 所示。这 9 家公司的年产能折合 8 in(1 in=2.54 cm)晶圆达到 180 万片,而国内 IC 设计公司对晶圆的年需求不到 50 万片,因此这些晶圆厂必须寻求海外订单。

表 1-1 中国大陆正式投产的晶圆厂

企业名称	晶圆 in 及工厂编号	投产年份	工艺/ μm	规划月产能/万片
中芯国际	8-1	2001	0.13	4.5
	8-2	2002	0.13	4.5
	8-3	2003	0.13	1.5
	12-4	2004	0.11	2.0
	8-7	2001	0.25	3.65
宏力半导体	8-1,1 期	2003	0.25	2.7
	8-1,2 期	2005	0.15	2.0

(续表)

企业名称	晶圆 in 及工厂编号	投产年份	工艺/ μm	规划月产能/万片
和舰科技	8-1	2003	0.15	3.2
	8-2	2005	0.13	2.0
先进半导体	5-1	1994	1	2.5
	6-2	1997	0.5	1.5
	8-3	2003	0.25	4.5
华虹 NEC	8-1	1999	0.25	3.5
	8-2	2004	0.18	5.5
贝岭微电子	4-1	1997	0.8	1.6
华润上华	6-1	1998	0.35	6
	8-2	2005	0.35	3
台积电	8-10	2004	0.25	3.5
柏玛	8-1	2004	0.35	3.0

2009年上海半导体产业量已占到全国的二分之一以上。浦东新区作为“上海国家产业基地”的核心,基本形成了以晶圆制造业为主体、产业链各环节逐渐完备的微电子制造产业群,同时成为我国大陆地区微电子产业投资规模集中、技术水平先进、产业体系完整、生产规模庞大以及最具发展潜力的微电子制造基地。

1.1.2 半导体晶圆制造系统特征

半导体芯片制造的整个过程可以分为晶圆制造、晶圆针测、装配与封装以及最终测试四个阶段。其中,晶圆制造与晶圆针测被合并称为芯片制造的“前端制造”(front-end operations),而装配封装和最终测试被总称为“后端制造”(back-end operations)。表 1-2 详细地阐述了半导体芯片制造过程。

表 1-2 半导体芯片制造过程

阶段/流程	描 述
晶圆制造 (fabrication)	本工序的主要工作是在晶圆上制作电路及电子元件(如晶体管、电容、逻辑开关等),其处理程序通常与产品种类和所使用的技术有关,但一般基本步骤是先将晶圆适当清洗,再在其表面进行氧化及化学气相沉积,然后进行涂膜、曝光、显影、蚀刻、离子植入、金属溅镀等反复步骤,最终在晶圆上完成数层电路及元件加工与制作

(续表)

阶段/流程	描 述
晶圆针测(probe)	经过上道工序后,晶圆上就形成了一个一个小格,即晶粒(die),一般情况下,为便于测试,提高效率,同一片晶圆上制作同一品种、规格的产品;但也可根据需要制作几种不同品种、规格的产品。在用针测(probe)仪对每个晶粒检测其电气特性,并将不合格的晶粒标上记号后,将晶圆切开,分割成一颗颗单独的晶粒,再按其电气特性分类,分别装入不同的托盘中,不合格的晶粒则舍弃
装配与封装 (assembly and packaging)	将单个的晶粒固定在塑胶或陶瓷制的芯片基座上,并把晶粒上蚀刻出的一些引接线端与基座底部伸出的插脚连接,以作为与外界电路板连接之用,最后盖上塑胶盖板,用胶水封死。其目的是用以保护晶粒避免受到机械刮伤或高温破坏。至此才算制成了一块集成电路芯片(即我们在电脑里可以看到的那些黑色或褐色,两边或四边带有许多插脚或引线的矩形小块)
最终测试 (final test)	测试可分为一般测试和特殊测试,前者是将封装后的芯片置于各种环境下测试其电气特性,如消耗功率、运行速度、耐压度等。经测试后的芯片,依其电气特性划分为不同等级。而特殊测试则是根据客户特殊需求的技术参数,从相近参数规格、品种中拿出部分芯片,做有针对性的专门测试,看是否能满足客户的特殊需求,以决定是否须为客户设计专用芯片。经一般测试合格的产品贴上规格、型号及出厂日期等标识的标签并加以包装后即可出厂。而未通过测试的芯片则视其达到的参数情况定为降级品或废品

上述四个过程中,晶圆制造是最为复杂和重要的环节,晶圆制造系统(Semiconductor Wafer Fabrication System, SWFS)是最为复杂的制造系统^[3,64]。SWFS 主要包括以下特征:

1) 工艺流程复杂

晶圆制造是要在以半导体材料硅或砷化镓为衬底的单晶圆薄片上制造出几十甚至几百个特定的集成电路芯片^[1,2]。晶圆的加工周期比较长,一般多为一个月到两个月。不同产品的工艺流程也不尽相同,简单的工艺流程可能有几十道工序,复杂的则可能有多达数百道工序。典型的半导体产品的加工流程往往包括 200~600 个工序,使用设备的种类有 60~80 种,设备台数多达上百台。

2) 重入型加工流程

重入型加工是晶圆制造系统区别于 Job-shop 和 Flow-shop 系统的主要特征。根据工艺要求,晶圆制造的整个过程由一定数量的重复工艺阶段(stages)构成,而每一阶段又包括若干道工序(steps)。由于 SWFS 中功能相同或相近的

机器都被安排到同一个机台组,故工序操作类型相同但位于不同加工阶段的晶圆将要访问相同的机台(群组)。这种产品在制造过程中多次访问同一机台(群组)的生产流程称为重入型制造流程^[3,4]。例如,工艺为 $0.25\ \mu\text{m}$ 的8 in晶圆的加工过程大概需要120个加工阶段,关键区域如蚀刻区的重入次数可多达20~30次。

3) 系统造价昂贵

晶圆制造系统中的机台是非常昂贵的。其中黄光区的一个主要机台,其价格就能达到几千万元人民币。以一座月产能为3.5万片的8 in(200 mm, 1 in=25.4 mm)晶圆厂为例,该系统包含了300~400多台加工机器以及若干条大型物料传送系统和大量辅助生产设备,总造价约为100多亿元人民币^[130]。而建设一座自动化的12 in(300 mm)晶圆厂,则需要高达150多亿元人民币的巨额投资。

4) 机台特性复杂

半导体生产线上不同的加工机台具有不同的加工方式,一般将其分为单批加工机台和多批加工机台。另外这些机台特性复杂,操作方式和运行模式各不相同。而且,各种精密设备还容易出现无法预测的故障并导致停机,且修复时间难以确定。另外,处于同一阶段的不同产品以及处在不同阶段的同一产品在同一台机器加工,其工艺参数也不尽相同,这又形成了复杂的依赖加工顺序的机台设置问题,更增加了调度的复杂性。

5) 多产品混线生产

与过去主要生产少数几种固定产品的面向库存生产(Make-to-Stock, MTS)的模式不同,晶圆代工是一种直接面对来自全球范围内的客户定制产品(Application-Specific Integrated Circuit, ASIC)市场,面向订单组织生产(Make-to-Order, MTO)是典型的大批量、超多品种混合的生产模式。一般月产能3万片的晶圆厂,其同时在线生产的产品甚至多达100多种。

1.1.3 半导体晶圆制造系统调度现状

1.1.3.1 调度理论的发展

在生产调度中,调度问题可以被描述为:在某一时间期限内分配一组机器来执行生产订单任务^[6]。更具体的,生产调度是针对一组可分解的生产任务,在尽可能满足约束条件(如交货期、工艺路线和资源情况等)的前提下,通过下达生产指令,安排生产任务的具体操作使用哪些资源、其加工时间及加工顺序,以获

得生产任务执行时间或成本的最优化。

从 20 世纪 50 年代起,应用数学、运筹学、工程技术等领域的科学家开始重视调度问题研究并解决了一系列有代表意义的调度和优化问题。但是,人们普遍把 Conway, Maxwell 和 Miller 三人有关调度的研究工作^[7]作为调度理论研究的正式开始,他们三人也被称为调度理论的奠基人。调度理论的发展历程如表 1-3 所示。

表 1-3 调度理论的发展历程

发展阶段	调度理论研究概述
20 世纪 50 年代	主要是提出了针对一些特殊和规模较小问题的解析优化方法,一般仅适用于单机和简单的流水车间(Flow-shop)问题,研究范围较窄,但这些早期的工作为经典调度理论后来的发展奠定了基础
20 世纪 60 年代	开始用一些普通的优化方法解决调度问题,主要是一些数学规划方法。60 年代末期,一个完整的经典调度理论体系开始初步形成
20 世纪 70 年代	问题复杂性方面的理论工作开始了,人们发现大多数调度问题都是 NP-complete 问题或 NP-hard 问题,一般很难找到解决这些问题的快速算法,因此解决调度问题的启发式算法开始成为求解调度问题的重要算法。到了 70 年代末期,经典调度理论开始趋于成熟
20 世纪 80 年代	调度理论与实际的结合已成为调度研究的首要问题。这个富有挑战性的课题吸引了来自机械制造、自动化技术、计算机科学、企业管理、系统工程和应用数学等多个领域的学者,许多学科领域的技术和方法被应用到调度问题研究中,调度成为一种跨学科的研究领域
20 世纪 90 年代至今	对调度问题的研究进入了高潮,各种研究手段得到了充分的发挥,同时还不断有新的研究工具被应用到调度研究当中。比较有代表性的技术有:系统仿真、遗传算法、人工神经网络和模糊数学等,同时学者们继续研究其他各种先进的启发式算法

1.1.3.2 SWFS 调度现状

根据上述对 SWFS 的特征描述可以得知,晶圆制造系统的调度问题属于大规模复杂动态调度问题。在半导体晶圆生产线上,流动着不同工艺流程的几十种甚至上百种产品,这些工艺流程具有多重入特征,产品制造周期长达数十天,

而且执行这些流程的机台具有类型繁多、特性复杂及高度动态不确定等特征,使得生产线上的在制品对机台设备使用权的竞争非常激烈。

由于晶圆制造过程的复杂性、多目标性、动态随机性和不确定性等因素,导致建立既能精确描述制造系统特性,又易于求解的调度模型非常困难,即使经过简化的模型也是 NP-hard 的,求解非常困难。受到实际需求的刺激,人们对 SWFS 的调度与控制这一富有挑战性的问题进行了大量的研究,但是尚未形成一套行之有效的方法和理论,以期能在合理的时间里得到系统的最优解。

在 SWFS 的调度领域,调度方法的效率(运算代价或时间)和解的精确性(对于最优解的逼近程度)似乎是一对很难调和的矛盾。目前被广泛使用的方法主要有三大类:数学规划方法、调度规则和智能算法。这些算法具有不同的特点,它们在各种条件下的性能也各不相同,对它们的比较分析非常重要。表1-4为几类调度方法的比较分析。

表 1-4 SWFS 调度方法比较分析

调度方法	分析评价(有效性、适应性和计算复杂性)
数学规划法	数学规划方法在理论上能够求出问题的最优解,它们的有效性当然是最高的,适应范围也比较广,几乎能适应任何一种调度问题,但是计算复杂性很高,这使得它们的计算规模容易受问题规模的限制,即使用来验证其他算法时,也是在小规模问题上验证,因此难以得到实际的应用
调度规则	调度规则的特点是算法简单,计算复杂度低,精确性不高,但是基本调度规则经过适当的组合和变形往往可以得到非常好的调度效果,有时甚至能得到问题的最优解,从总体上看,调度规则的最大优点是计算量小,单个的调度规则有效性较低,适应性也较窄,经适当的组合,后两项指标都会有较大提高,正因为如此,人们对这种方法的热情经久不衰,到目前为止,调度规则仍然被普遍认为是最实用的调度算法
智能算法 (人工智能、 计算智能、 群体智能)	一定程度上依靠枚举,一般来说收敛到最优解很慢,而且很难表示解的最优性。某些文献中涉及的调度算法,虽然能够得到很好的调度,但是仅适用于规模相当有限的系统,对设备数量和任务数量有很大的限制。例如:① 人工神经网络的效率受训练影响很大,并且在问题规模较大时,存在计算速度慢与结构参数难以确定的弱点;② 模拟退火法能以一定的概率接受差的能量值,因而有可能跳出局部极小,但它的收敛速度较慢,难以用于实时动态调度环境;③ 遗传算法在调度领域中已经得到了比较广泛的应用,几乎每一种调度问题都有人用遗传算法求解,因此可以认为遗传算法的适用性是比较好的;遗传算法的有效性一般与其相关参数的选择有关,随着问题规模的增大,算法有效性呈下降趋势;遗传算法的计算复杂性一般是通过实验获得的,理论上的分析不多见

1.2 半导体晶圆制造系统简介

1.2.1 半导体晶圆制造过程^[8,9]

晶圆制造是半导体芯片制造的关键环节,需要花费数十天的时间在晶圆上完成数百步复杂的化学和物理操作。由于大多数制造流程都发生在晶圆表面的几微米以内,整个工艺过程非常精密。CMOS(Complementary Metal-Oxide-Semiconductor Transistor,互补型金属氧化物半导体)技术在集成电路工艺中最具有代表性,下面将以 $0.18\ \mu\text{m}$ 的 CMOS 集成电路制造工艺为例,简单介绍晶圆制造的多重入型工艺流程。

如图 1-2 所示,半导体晶圆在首先经过清洗后,送到热炉管,在含氧的环境中,以加热氧化的方式在晶圆表面形成一层二氧化硅(SiO_2),紧接着在刚刚长成的二氧化硅上以化学气相沉淀的方式形成一个氮化硅(Si_3N_4)层,然后整个晶圆将进行曝光显影制程,即先在晶圆表面涂上一层光阻,再将光罩上的图案曝光光刻到光阻上面。接着利用蚀刻技术,将部分未被光阻保护的氮化硅层除去,留下的就是所需要的集成电路图。随后,以磷为离子源对整个晶圆进行离子植入,并

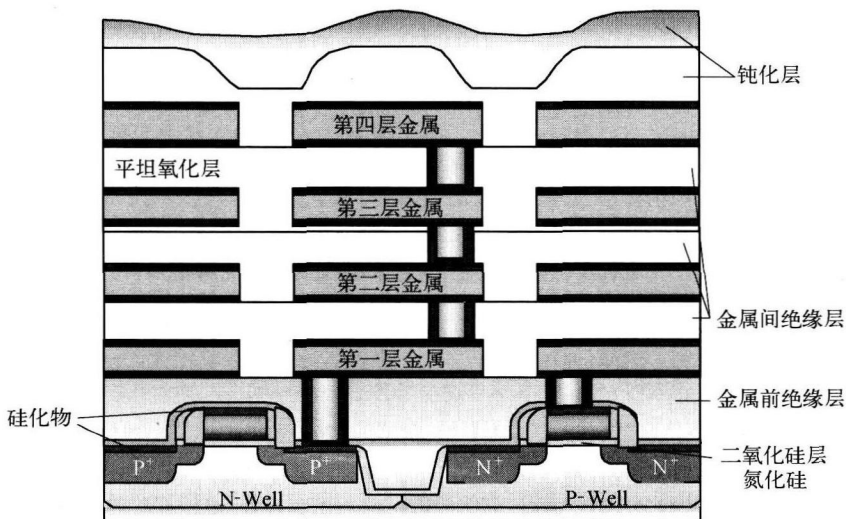


图 1-2 CMOS 工艺流程图

将光阻剂去除。最后接着进行金属化制程,制作金属导线,以便将各个晶体管与组件加以连接,增加芯片密度能够在晶圆表面放置更多的组件,这实际上就减少了表面连线的可用空间,这个问题的解决方法就是利用2~4层独立金属层的多层金属结构,到2012年,芯片上的金属层可望达到9层。到此,一个层次(layer)的集成电路制造完毕。此时必须进行一些电性或物理特性测量,以检验加工结果是否在技术规范内,即晶圆级电学在线测试。如此反复以上步骤,通过重入型加工过程完成集成电路的剩余第二层、第三层、直至第 n 层的电路加工。

上述基本工艺过程是在几个主要加工区域,即扩散区、光刻区、刻蚀区、淀积区、金属淀积区、离子注入区和化学机械研磨区内完成的,所有区域均处在晶圆厂的超净环境中。具体介绍如下:

(1) 扩散区(diffusion): 扩散区是进行高温工艺及薄膜淀积的区域。其主要设备是高温扩散炉和湿法清洗设备。高温扩散炉可以在近 $1\ 200^{\circ}\text{C}$ 的高温下工作,并能完成多种工艺流程,包括氧化、扩散、淀积、退火及合金。湿法清洗设备是扩散区中的辅助工具。晶圆在放入高温炉之前必须进行彻底清洗,以除去晶圆表面的玷污以及自然氧化层。

(2) 光刻区(litho): 独特的黄色荧光灯照明使得光刻区与晶圆厂中的其他各个区域明显不同,故又称其为黄光区。光刻的目的是将电路图形转移到覆盖于晶圆表面的光刻胶上。光刻胶是一种光敏的化学物质,它通过深紫外线曝光来印制掩模板的图像。光刻胶只对特定波长的光线敏感,例如深紫外线和白光,但对黄光不敏感。

(3) 刻蚀区(etching): 蚀刻是在晶圆上没有光刻胶保护的地方留下永久的图形,最常见的工具是等离子体蚀刻机、等离子去胶机和湿法清洗设备。等离子体蚀刻机是采用射频能量在真空腔中离化气体分子的一种工具。蚀刻工艺主要有两大类:湿法刻蚀和干法刻蚀。两种方法的主要目标是将光刻掩模板上的图案精确地转移到晶圆表面或将光刻胶精确地移除。

(4) 离子注入区(ion implantation): 气体携带要掺杂的杂质,例如砷(As)、磷(P)、硼(B),将在注入机中离化(采用高电压和磁场来控制并加速离子)。高能杂质离子能够穿透涂胶晶圆的表面。离子注入完成后,要进行去胶和彻底清洗晶圆。

(5) 化学气相淀积区(Chemical Vapor Deposition, CVD): 主要负责生产各个步骤中的介质层与金属层的淀积。化学气相淀积中所采用设备的温度低于扩散区中设备的工作温度。该工艺的大致步骤如下:首先,将晶圆装载到反应室

内,装载过程通常是在惰性气体的环境下进行的。然后,晶圆被加热到预定温度,将反应气体引入到淀积薄膜反应室进行反应。最后,将参与反应的化学气体排除。

(6) 金属淀积区(metal deposition): 金属淀积区设备主要负责芯片表面金属连线。把各组件连接到一起的材料、工艺,联机过程一般称为金属化工工艺(metallization process)。由于淀积多层金属及合金的需要,以及对金属淀积的阶梯覆盖度的更高要求,使得溅射技术成为大型集成电路制造的标准淀积技术。

(7) 化学机械研磨区(Chemical Mechanical Polishing, CMP): 化学机械研磨的目的是使晶圆表面平坦化,这是通过将晶圆表面突出部分减薄到下巴部分的高度实现的。晶圆表面凹凸不平给后续加工带来了困难,而 CMP 使这种晶圆表面的不平整度降到最小。

1.2.2 生产绩效指标

生产绩效指标可将生产系统之运作状态以量化的形式表达,因此,管理人员能借此判断系统的实际绩效与规划目标的差异程度,并协助其确认现场潜在的问题,以便适当地调整生产策略。然而,由于晶圆制造的高复杂性导致在此产业中无法以单一的指标来衡量其生产绩效,需要同时考核数个指标(如产出率、在制品量、生产周期等)才能合理地表达系统成效。

过去的研究文献在探讨晶圆制造的生产绩效指标时,数位学者都各以不同的观点提出其分类方式。如表 1-5 所示。

表 1-5 生产绩效指标分类

学 者	指 标 类 别	指 标 项 目	
Atherton ^[10]	决定 Fab 的生产运作模式	设备集合 设施 制程优先权 调度	可用度 产品组合 控制法则
	反应既定生产运作模式的结果	晶粒良率 生产周期 设备利用率	产出量 调度稳定度
	兼具前两种模式的特性	动态产能 系统瓶颈	存货状态

(续表)

学 者	指 标 类 别	指 标 项 目	
Leachman 等 ^[11]	品质	生产线平均良率 缺陷密度	晶粒平均良率 Fab 与晶粒类别整合良率
	生产力	每日步进机种的晶圆作业数 每日直接人工的生产力 每日全体人工的生产力 主要设备的生产率	
	生产速度与产出目标	每光罩层平均生产周期 平均交期达成率 新制程导入至量产的效率	
林悦慈 ^[12]	产品	每个层次(layer)的生产周期 总生产周期 实际/理论的生产周期比率 总在制品量利用率 延迟工件个数平均符合大小	
	设备	可用时间比率 加工时间比率 当机及预防保养比率	
	其他的资源	生产成本 作业员的效率 品质因素	

周煜智^[13]将生产控制指标区分为产品与设备两大类,并依生产绩效再细分为交期率(delivery)、在制品(WIP)、产出量(throughput)、生产周期(cycle time)和设备效率(equipment efficiency)等五种类别,如表 1-6、表1-7 所示。

表 1-6 与产品相关的生产控制指标

类 别	指 标 与 描 述	单 位
与交货期相关指标	工单无法如期完工(number of tardy jobs)	#
	工单如期完工的比率(交货率)(on time delivery ratio)	%
	工单的准确性(mean lateness)	Time

(续表)

类别	指标与描述	单位
	工单的延误性(mean tardiness)	Time
	工单之准确度(lateness variance)	Time
	工单延误之比率(percentage of tardy)	%
	显示预测结果变异之大小(mean square error)	Time
	配合成本考量(mean absolute error)	Time
与在制品相关指标	系统存在的 WIP 总数(total WIP)	#
	系统曾出现的最大 WIP 总数(max WIP)	#
	系统存在执行于生产的 WIP 总数比率(active WIP ratio)	%
	系统存在执行于制程测试的 WIP 总数比率(engineering WIP ratio)	%
	系统存在执行于暂停活动的 WIP 总数比率(hold WIP ratio)	%
	系统存在执行于重加工的 WIP 总数比率(rework WIP ratio)	%
	在制品水准控制能力(WIP effective)	%
与产出量相关指标	单位时间内产出量(throughput)	#
与在制品、 产出量相关指标	系统总 WIP 数与产出量的比率(WIP-to-out ratio)	#/Time
	系统 WIP 流动的平均速率(turn over rate)	
与生产加工 周期相关指标	产品总生产周期(total cycle time)	Time
	产品生产周期的标准差(cycle time std. var.)	Time
	产品流经各层级的生产周期(cycle time per layer)	Time
	实际生产周期与理论生产周期之比值(actual-to-theoretical cycle time ratio)	

表 1-7 与设备相关的生产控制指标

类别	指标与描述	单位
与产出量相关指标	单位时间内设备完成的工单数(total movement)	#
	每小时设备产出的晶圆数(wafer per hour)	#
与等待周期 相关指标	机台前等候列的平均等候时间(mean queuing time)	Time
与设备效率相关指标	实际加工时间与理论加工时间的比值(process time ratio)	
	设备处于可利用状态的比例(up time ratio)	%
	设备处于当机状态的比例(down time ratio)	%
	设备执行加工作业之比例(utilization)	%
	设备平均装载率(average load size)	#
	设备处于使用状况的比例(loading rate)	%
	设备真正使用于计划生产的比例(OEE indexes)	%
	瓶颈设备真正使用于计划生产的比例(CUBES indexes)	%

尽管生产绩效指标为数众多,但大多属于辅助性衡量指标。一般而言,接单式生产环境(如晶圆代工厂)特别着重在生产周期与交期达成率,而存货式生产环境(如整合元件制造厂)则偏重在产出水准绩效^[14]。

林悦慈^[12]并指出产出量及良率为晶圆厂主要的重点绩效指标,也是考核绩效的依据。然而影响绩效指标的重要关键因素则为:生产周期、在制品量及机台利用率,故一般生产管理人员会将重点放在此关键因素上。虽然不同的生产目标必须有不同的指标来加以衡量,但整体而言,晶圆制造厂的生产绩效包括总产出、生产周期、在制品量、设备利用率、移动及良率^[15]。

过去研究对生产绩效的见解与应用,通常依照不同的管理需求选择适用的绩效管理指标。对晶圆制造厂而言,为满足生产管理的需求,亦有各种不同方面的衡量指标。晶圆制造厂在对生产系统做排程规划与控制时,最主要关注的焦点仍不外乎为产出量、在制品水准、生产周期、移动等绩效指标。

1.3 半导体晶圆制造系统建模研究

1.3.1 研究状况简介

如前所述,SWFS 所具有的若干复杂特性使得为其建模非常困难。然而,由于 SWFS 仍是以事件为驱动的、内部存在大量同步与异步操作,以及资源争夺与共享的复杂生产系统,故它可被描述为一个复杂的离散事件动态系统(Discrete Event Dynamic System, DEDS)。从 DEDS 理论的角度出发,制造系统可以用状态、事件(生产活动)以及它们之间的关系来描述^[16]。目前,用于制造系统的 DEDS 建模方法大体可分为数学与图形两类:前者包括正规语言与自动机(formal language and automata)、算术代数与布尔函数(arithmetic algebra and boolean functions)、排队网(Queueing Network, QN)等;后者则包括事件图(event graph)、活动周期图(activity cycle diagram)以及 Petri 网(Petri Nets, PNs)方法^[17]。从当前的研究现状看,应用于 SWFS 的建模方法可分为如下四种:

1) 排队网模型(Queueing Networks, QN)

使用最早的模型,曾被广泛用于生产少品种大批量晶圆产品的 SWFS 建模^[25]。学者 Kumar^[26,27]的很多理论分析都建立在排队网模型的理论基础上,使用最多的是 Markov 排队网,即通过对 Markov 过程的分析得到系统的线性约束方程。但他们忽略了晶圆制造过程中的随机性重做和次品报废等事件,并认为加工流程是确定的。从降低制造周期的目标出发,Kumar 提出了一类基于最小松弛的波动光滑策略,通过选择不同的松弛参数,如以缓冲区的剩余制造周期为松弛参数,得到使制造周期平均值为最小的 FSMCT(Fluctuation Smoothing Policy for Mean Cycle Time)策略;以投料时间为松弛参数,得到使制造周期方差最小的 FSVCT(Fluctuation Smoothing Policy for Variance of Cycle Time)策略等,并用仿真实验的方法显示 FS 策略产生的良好性能,FSMCT 比 FIFO(First In First Out)降低制造周期均值 22.4%,FSVCT 比 FIFO 降低制造周期方差 52.0%。然而,随着现代 SWFS 规模和复杂性的增加,特别是在大批量多品种混合生产的模式下,QN 模型已经越来越不能满足实际生产的需要,目前已经很少被采用。

2) Brownian 模型

QN 理论的延伸,它将模型中的控制问题近似为排队网络的调度问题,可适当降低计算的复杂性^[21]。该模型在从设备角度进行派工策略的研究中应用广泛。例如,Harrison^[21]和 Wein^[22]用 Brownian 模型研究了两机多任务的一般闭环排队网。文中定义了反映两机间负载不均衡程度的负载不均衡变量,并将它近似为一维 Brownian 过程。通过对此过程的分析 and 求解,发现了设备空闲和最低优先级任务之间的关系,即当负载不均衡过程运行到边界时,设备才发现功能空闲,两个极点对应此刻处于最低优先级的两类工件。作者由此得到了一种负载平衡的调度规则 WBAL(Work Loading Balancing Rule)。Chevalier 和 Wein^[23]还将该方法推广到多机多任务的一般闭环系统。Brownian 模型仍然采用解线性方程的方法获得调度策略,当问题可以归结为一维 Brownian 过程的解析解时还可以处理,但当得到多维线性方程组的时候,便很难或根本无法得到解析解,所以 Brownian 模型只适用于少量机台的派工研究。

3) 连续流模型(Fluid Networks, FN_s)

Chen 和 Yao 等人指出描述连续系统的连续流网模型,在函数强大数定理保证下,该模型可用来研究排队网络系统的渐近行为,其优点是降低了系统的维数。连续流网模型由流点和站点组成,生产系统中的一个加工任务的一个工步表示为一个流点,一类设备表示为一个站点,通过求解系统资源满足的线性方程,得到使整个计划期内线性库存费用最低的设备能力分配方案。用流网模型解决动态调度问题,所需的输入信息少,仅需投料速率、加工时间和首批订货数。该模型基于系统全局信息,可将任何时刻作为初始点的动态方法,具有良好的鲁棒性,适用于随机环境。但该连续流网模型的缺点仍然是难以适应目前多品种、大批量混合生产的模式^[24]。

4) Petri 网模型(Petri Nets, PN_s)

Petri 网最早由德国的 Carl A. Petri 博士于 1962 年在他的博士论文中提出,用于描述计算机系统事件之间的因果关系。清晰的图形化界面和完善的数学分析工具使 PN_s 成为离散事件动态系统领域建模和调度研究的得力工具。目前,基于 PN_s 的建模方法已经在先进制造系统领域中得到广泛应用,如自动制造系统(Automated Manufacturing System, AMS)及柔性制造系统(Flexible Manufacturing System, FMS)等,并取得了一系列令人瞩目的成果。由于 SWFS 仍然是由离散事件驱动的、并由离散事件按照一定运行规则相互作用