SEMICONDUCTOR CLUSTER TOOL AVAILABILITY

by
Troy A. Bailey

B.S., Electrical Engineering
University of Washington, 1992

Submitted to
the Department of Electrical Engineering and Computer Science and
the Alfred P. Sloan School of Management
in Partial Fulfillment of the Requirements for the Degrees of

Master of Science in Electrical Engineering
and
Master of Science in Management

at the
Massachusetts Institute of Technology
June 1994

© 1994, Massachusetts Institute of Technology

Signature of Author____________________________________
Department of Electrical Engineering and Computer Science
Sloan School of Management
May 6, 1994

Certified by____________________________________ 5/4/94
James E. Chung
Assistant Professor of Electrical Engineering

Certified by____________________________________
Charles H. Fine
Associate Professor of Management

Accepted by____________________________________
Frederic R. Morgenthaler
Chairman, Committee on Graduate Students

SEMICONDUCTOR CLUSTER TOOL AVAILABILITY

by
Troy A. Bailey

Submitted to
the Department of Electrical Engineering and Computer Science and
the Alfred P. Sloan School of Management
on May 6, 1994
in Partial Fulfillment of the Requirements for the Degrees of
Master of Science in Electrical Engineering and
Master of Science in Management

Abstract

Success in today's semiconductor manufacturing industry depends, in part, upon efficient use of expensive capital equipment. The manufacturing tools that a semiconductor company selects and the manner in which those tools are utilized greatly influences the company's financial success.

Recent trends in semiconductor manufacturing tools have favored integrating multiple processes into a single "cluster" tool. Integrating processes can provide technological and economic benefits, however there are costs associated with the increased complexity of cluster tools. Tool availability is identified as one such cost.

Four case studies are used to investigate availability issues associated with cluster tools. Based on the issues identified in the case studies, a set of guidelines for the utilization of cluster tools is presented. Computer models are proposed as a method of quantifying cluster tool availability. Factors relevant to accurate availability models are discussed and general recommendations regarding cluster tool availability modeling are made.

The process of acquiring a new semiconductor production tool and integrating it into an existing production line is also presented. The experiments and procedures used to develop a new process are studied in detail.

Thesis Advisors:

James E. Chung, Assistant Professor of Electrical Engineering
Charles H. Fine, Associate Professor of Management

Acknowledgments

I would like to acknowledge the Leaders for Manufacturing Program for its support of this work.

The research for this thesis was performed at Digital Equipment Corporation's Semiconductor Operations division in Hudson, MA. I would like to express my gratitude to all the people at Semiconductor Operations who welcomed and assisted me during my internship. Particularly, I would like to thank my company supervisor, Brian Lee, for his support and guidance.

The work presented in Chapter 5 is the result of a team effort. I would like to acknowledge the other members of the team: Frank Krupa, Kosta Culafi, Orion Farina, Joe Mensah, and the process and equipment engineers from the tungsten deposition and etch tool vendors.

I also owe a debt of gratitude to my thesis advisors, Jim Chung and Charlie Fine, for their encouragement and direction.

Finally, I would like to thank three people who also helped make this work possible:

> My father, who taught me to be persistent;
> My mother, who taught me to be resourceful;
> and my wife, who taught me to believe in myself.

Table of Contents

1. Introduction 11
1.1 Overview of Semiconductor Manufacturing 11
1.2 Semiconductor manufacturing tools 13
1.2.1 Tool Productivity 14
1.2.2 New Tool Integration 16
1.3 Thesis Organization 17
2. Cluster Tools 19
2.1 Description 19
2.2 Advantages 19
2.3 Disadvantages 21
2.4 Trends 22
3. Case Studies of Cluster Tools 25
3.1 Integrated Tungsten Deposition and Etchback Tool 25
3.1.1 Tool Description 25
3.1.2 Process 27
3.1.3 Availability Issue Illustrated by This Tool 27
3.2 Oxide Deposition Cluster Tool 30
3.2.1 Tool Description 31
3.2.2 Process 31
3.2.3 Availability Issues Illustrated by This Tool 32
3.3 Plasma Etch Cluster tool 34
3.3.1 Tool Description 34
3.3.2 Process 35
3.3.3 Availability Issue Illustrated by This Tool 36
3.4 Highly Integrated Metal Deposition Tool 38
3.4.1 Tool Description 38
3.4.2 Process 40
3.4.3 Availability Issue Illustrated by This Tool 41
3.5 Summary 44
3.5.1 Guidelines for Integrating Processes on Cluster Tools 46
4. Modeling Cluster Tool Availability 49
4.1 Applications of Semiconductor Tool Models 49
4.1.1 Modeling Individual Tools 49
4.1.2 Modeling Entire Fabs 50
4.2 Characteristics Relevant to Accurate Availability Models of Cluster Tools 50
4.2.1 Individual Process Module Reliability Characteristics 50
4.2.2 Interferences Between Process Modules 51
4.2.3 Multiple Processes for a Single Tool 52
4.2.4 Failure Diagnosis Delays 52
4.2.5 Maintenance Policies 52
4.3 Recommendations for Accurate Availability Models of Cluster Tools 53

4.3.1 Incorporate Availability Modeling Into the Existing Tool Models 54
4.3.2 Model Process Modules and Platform as Separate Entities with Interactions 54
4.3.3 Delays for Monitoring and Failure Diagnosis 55
4.3.4 Include All Relevant Maintenance Issues 56
4.3.5 Recommendations for Additional Research 56
5. New Tool Acquisition and Integration 59
5.1 Acquisition 60
5.1.1 Selection 60
5.2 Integration 62
5.2.1 Typical Startup Schedule 62
5.2.2 Process Development 63
5.2.3 Installation & Mechanical Startup 70
5.2.4 Process Startup: 70
5.2.5 Process Refinement 76
5.2.6 Process Verification 86
5.2.7 Release to Production 88
5.3 Discussion 88
5.3.1 A Case for Production Optimized Tools 88
5.3.2 Shifting Priorities 89
5.3.3 The Need to Cooperate 91
5.3.4 Developing a Process In Three Different Locations 92
5.3.5 The Difficulty of Transferring a Process 93
5.3.6 Planning in the Face of Uncertainty 94
5.4 Summary 96
Appendix A: Background, Tungsten Plugs 99
A.1. Tungsten Plugs 99
A.1.1. Purpose 99
A.1.2. Process 99
A.1.3. Critical Plug Characteristics 102
A.1.4. Critical Etch Properties 106
A.2. Vendor 3's Tungsten Etch Process 108
A.2.1. Plasma etching 108
A.2.2. Etch Tool Configuration 109
A.2.3. Etch Chamber Configuration 110
A.2.4. Temperature Effects and Helium Backside Cooling 112
A.2.5. Etch Sequence 113
Appendix B: TiN Etch Experiment 117
Appendix C: Tungsten Etch Experiment 121
Appendix D: Repeatability Experiment 127
References 131

List of Figures

Figure 1.1. Integrated Circuit Fabrication Process 13
Figure 2.1. Basic Types of Cluster Tool Configurations 20
Figure 3.1. Configuration of Tungsten Deposition and Etchback Cluster Tool 26
Figure 3.2. Configuration of Oxide Deposition Cluster Tool 31
Figure 3.3. Configuration of the Plasma Etch Cluster Tool 35
Figure 3.4. Highly Integrated Metal Deposition Cluster Tool 39
Figure 5.1 Tungsten Etch Residues on TiN 65
Figure 5.2 Residue and Roughened TiN Surface After Tungsten Etch 66
Figure 5.3. Tungsten Plug Surface After Extended Etching of Residues 66
Figure 5.4. Cross Section of the Wafer Edge. 68
Figure 5.6. SEM Image of Tungsten Plugs from Overetch Experiment 85
Figure A.1. Tungsten Plug Process Cross Sections. 101
Figure A.2. Thinned Metal Over Recessed Plug. 102
Figure A.3. Cross Sections of Tungsten Plugs. 104
Figure A.4 Cross Section of Tungsten Plug Showing Trenching of TiN 105
Figure A.5. Cross Section of Tungsten Plug Showing Coring. 106
Figure A.6. Block Diagram of the Tungsten Etcher (plan view). 110
Figure A.7. Tungsten Etch Chamber Cross Section. 111
Figure A.8. Typical Endpoint Signal Trace for Tungsten Etchback Process 114
Figure C.1. Contour Plot of Tungsten Etch Rate and Uniformity 125

List of Tables

Table 3.1. Process Sequences for Plasma Etch Cluster Tool 36
Table 3.2. Usable States of the Clustered Etch Tool 37
Table 5.1 TiN Etch Experiment Trends 79
Table 5.2 Tungsten Etch Experiment Trends 81
Table B.1. Coded Experimental Design Matrix for TiN Etch Experiment 117
Table B.2. Actual Factor Levels Corresponding to Experimental Design Codes 117
Table B.3. Results of TiN Etch Experiment 118
Table B.4. Factor Analysis of TiN Etch Experiment Responses. 119
Table C.1. Coded Experimental Design Matrix for Tungsten Etch Experiment 121
Table C.2. Actual Factor Levels Corresponding to Experimental Design Codes 122
Table C.3. Results of Tungsten Etch Experiment 123
Table C.4. Factor Analysis of Tungsten Etch Experiment Responses. 124
Table D.1. Etch Rate and Uniformity Test Results 128
Table D.2. Etch Particle Test Results 128
Table D.3. Machine Particle Test Results 128
Table D.4. Repeatability Experiment Etch Times (seconds) 129

1. INTRODUCTION

Semiconductor manufacturing is a capital intensive industry. A modern, high-volume fabrication facility (fab) can cost in excess of 2 billion dollars[1], with approximately seventy-five percent of that going toward manufacturing tools[2]. A manufacturer's tool selection and utilization strategies are critical to it's success.

1.1 OVERVIEW OF SEMICONDUCTOR MANUFACTURING

Semiconductor circuits are manufactured on circular slices of pure, single-crystal silicon. The slices, called wafers, are typically 100 mm to 200 mm in diameter and approximately 0.5 mm to 1 mm thick. The circuit is manufactured on the surface of one side of the wafer through a complicated process involving hundreds of steps.[8]

Semiconductor integrated circuits are composed of three dimensional structures of conducting, insulating and semiconducting materials. These structures are created by forming a layer of material on the wafer, patterning that film through a photographic process, and removing a portion of the film in an etch process. Subsequently another layer is formed and the process is repeated until, layer by layer, the desired structures are constructed.

By controlling the shape, sequence and characteristics of the layers, various electrical devices may be produced. These devices, when connected electrically by patterned conductive layers, comprise an integrated circuit. The patterns that

define modern integrated circuits include minimum features of 0.5 micron (one millionth of a meter) or less. The most complicated circuits contain millions of individual devices.

Integrated circuits are laid out in rectangular areas with dimensions ranging from 1 mm on a side to as large as a few centimeters. The circuit pattern is repeated across the wafer as few as 50 times to as many as 1000 times depending on the relative sizes of the circuit and the wafer. When processing is complete, the wafer is cut, or diced, separating the individual circuits which are then referred to as die.

Figure 1.1 shows the fabrication process described above. The wafers iterate through the add layer-pattern-etch process 20 times or more for complex circuits. These processes include oxidation, diffusion, ion implantation, evaporation, chemical and physical vapor deposition, photolithography, and wet and dry etch processes. The tools which are used to perform the process steps are discussed in the next section.

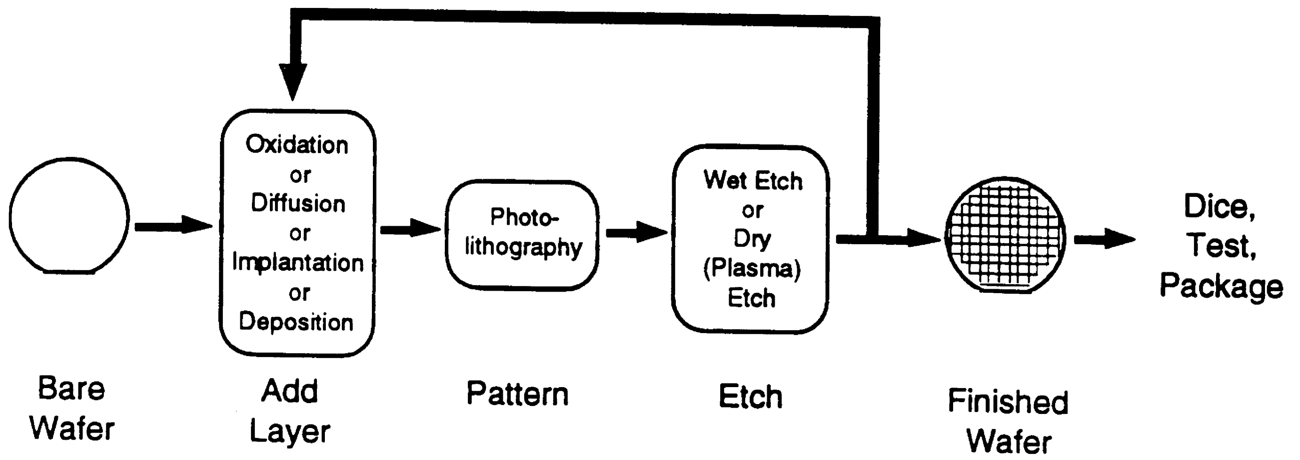


Figure 1.1. Integrated Circuit Fabrication Process

Complicating the entire fabrication process is the fact that it must be conducted in an ultra clean environment. The successful fabrication of integrated circuits requires that the wafer surface be free from any form of contamination; a single particle the size of the smallest feature on the chip can destroy the whole chip. Modern fabs employ elaborate environmental control systems and cleanliness procedures to ensure that there is less than one particle, larger than one micron, per cubic foot of air. Even with such extreme measures, wafer contamination continues to be a major factor limiting integrated circuit production yield.

1.2 SEMICONDUCTOR MANUFACTURING TOOLS

Semiconductor manufacturing tools are among the most complicated and expensive manufacturing equipment used today. Integrated circuit production lines include tools utilizing the latest in advances in chemical processing, vacuum technology, plasma physics, microscopic imaging, high-energy physics,

materials science, and robotics. Increases in the performance of each successive generation of integrated circuits depends largely on advances in manufacturing technology. To compete in the high-performance integrated circuit markets, manufacturers and tool vendors continually push the tools and technology to its limits.

A modern fab includes hundreds of tools ranging in price from hundreds of thousands to several million dollars. Ironically, these tools may be capable of manufacturing state-of-the-art circuits for only a few years before advances in technology and competition dictate that they be replaced. In order to obtain an adequate return on their investment in manufacturing tools, semiconductor companies must select and operate their tools in such a manner as to maximize productivity.

1.2.1 Tool Productivity

To evaluate or compare manufacturing tools, it is common to establish a measure of productivity. In accordance with the goals of manufacturing businesses, productivity measures typically include quality of production and rate of production. Aside from these similarities, productivity measurements vary greatly from industry to industry and from tool to tool.

Many Japanese manufacturer's use a standard tool productivity measurement that, due to it's simplicity, can be applied to many tools. This standard measure expresses manufacturing tool productivity as:

Productivity = Usable Output÷{Theoretical Output Rate x Total Time}. [3]

This equation is powerful in its simplicity as it removes much of the detail of more specialized measures and leaves only an essential assessment of how well the tool satisfies the manufacturing business goal: to produce acceptable quality products in a timely manner.

The expression above is a ratio of what was actually produced to the ideal output. The difference between these numbers comes from the facts that not all output is of acceptable quality, tools are not always operated at their maximum output rate, tools are not always in an operable condition, and when they are operable, they are some times idle due to inefficient scheduling or material flow.

An equivalent expression which explicitly refers to the non-idealities mentioned above is

Productivity = Output Quality x Output Rate x Utilization x Availability

where

Output Quality = acceptable product/total product processed,
Output Rate = total product processed/time spent processing product,
Utilization = time spent processing product/ time available for processing, and
Availability = time available for processing/total time.

This expression identifies availability as one of the factors that determine how well the tool performs its intended purpose.

1.2.1.1. Availability

All complicated manufacturing tools require periodic maintenance and are susceptible to breakdowns. As a result, they are not available for production one hundred percent of the time. In addition to unscheduled repairs and periodic maintenance, manufacturers typically perform preventive maintenance and process monitoring activities that also reduce the amount of time a tool is available for production. Depending on tool reliability and process complexity, non-productive periods or "downtime" can account for a significant portion of total time. Availability as defined in the equation above and in the following discussions includes the concepts of tool reliability and maintainability.

1.2.2 New Tool Integration

Another aspect of semiconductor tool use which can affect a manufacturer's return on tool investment is the procedure that company uses to integrate a new tool into the production line. In many manufacturing industries, starting up a new tool is a matter of bolting it down and plugging it in. The complexity of semiconductor manufacturing tools and processes requires considerably more effort and the time required to bring a new tool on line can represent a significant portion of the tool's useful life.

1.3 THESIS ORGANIZATION

This thesis examines semiconductor tool application strategies by investigating the issues relating to availability for a particular type of equipment referred to as "cluster tools" and by observing the procedure used to startup a new tool. The research work was conducted at Digital Equipment Corporation's semiconductor fabrication facility in Hudson, Massachusetts.

Chapter 2 is an introduction to cluster tools. It gives a brief description of cluster tool configurations as well as presents some of the advantages and disadvantages of cluster tools vis-a-vis non-integrated tools. Chapter 3 contains the results of case studies of several cluster tools used by Digital with the purpose of identifying factors which may lead to reduced availability for this type of tool. Chapter 4 considers modeling as a quantitative method that might be used to compare the availability of various tool configurations. In Chapter 5 an account is given of the procedure used to acquire a new semiconductor manufacturing tool and integrate it into a manufacturing line.

2. CLUSTER TOOLS

Technical and economic advantages have been gained by combining, or clustering, two or more semiconductor manufacturing processes into a single tool. Typically this is done by connecting multiple processing modules to a central platform with a wafer handler that moves wafers between the processes.

2.1 DESCRIPTION

Clustering can be parallel, serial, or a combination of both as shown in Figure 2.1. In parallel clustering, several processing modules that perform the same process are connected to one handler and each wafer is processed through a single unit. In serial clustering, processing modules of different types are combined into one tool and wafers are processed through all of the different modules . Combining parallel and serial clustering provides redundancy and allows for balancing processes with different cycle times.

2.2 ADVANTAGES

Technical advantages are gained from clustering when processes can be run in quick succession or when the wafers can be held in a controlled environment between process steps. One area that has shown particular advantages has been metalization. When exposed to atmosphere, metal films are susceptible to oxidation. By holding the newly deposited film under vacuum until a passivating process is run, oxidation can be avoided. Cluster tools also reduce the number of

times that wafers are handled by operators further reducing the possibility of contamination and physical damage.

Another kind of technical advantage of clustering is realized in semiconductor development. Process developers often select tools before they are certain of the specific processes they will be running. The flexibility of cluster tools, which can be reconfigured by removing and adding processing modules , is attractive to developers.

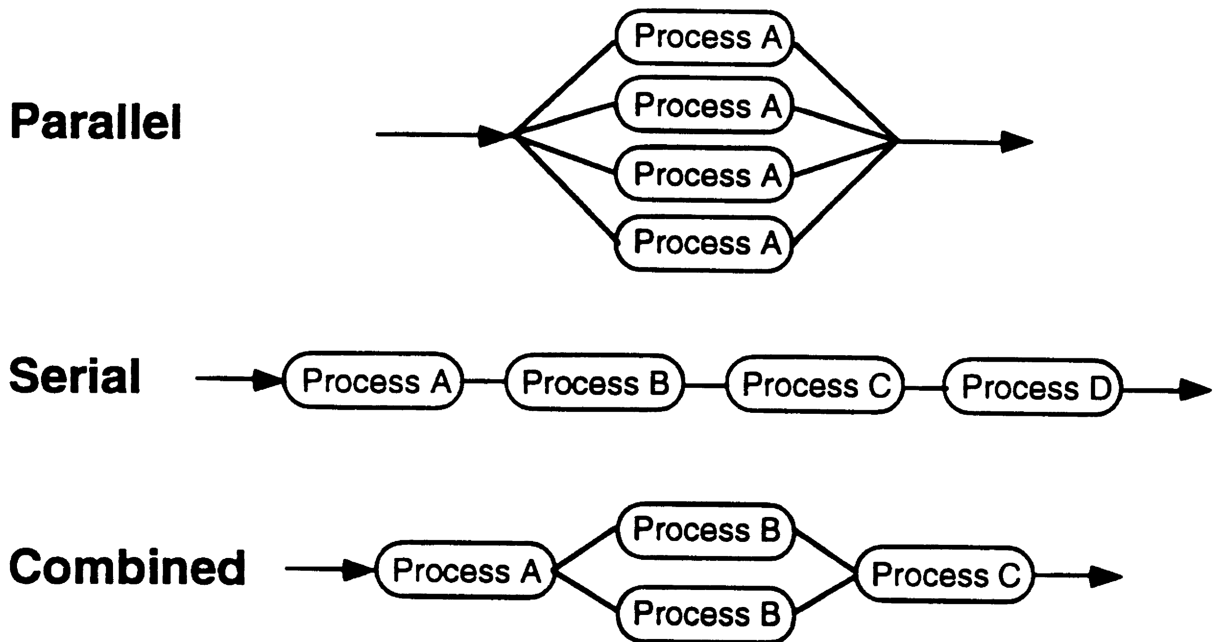


Figure 2.1. Basic Types of Cluster Tool Configurations

Economic advantages are gained through cycle time and cost reduction. With serial clustering, wafers are moved directly from one process to the next, without being removed from the tool; wafer handling and cycle time are thus reduced. Parallel clustering can also reduce cycle time by processing more than

one wafer at a time. Clustering process around a shared platform reduces the floor space required for each process. Because the fab environment is expensive to maintain, a smaller footprint also translates to an economic advantage.

2.3 DISADVANTAGES

Cluster tools are inherently more complex than non-cluster tools[4]. As more processing modules are added to a tool, more gas lines, vacuum pumps, exhaust lines, monitors and controllers must also be added. This complexity and the need to conserve on space often results in a tool that is more densely constructed and difficult to maintain.

When different types of processes are clustered together, developing and controlling the overall process is more difficult. Interactions between processes, although often quite subtle, can be critical. Such interactions may exist for non-clustered tools as well, but the difficulty of testing and understanding these interactions is compounded when wafers are subjected to several processes between tests.

The disadvantages mentioned above result in lower availability for cluster tools. Many engineers familiar with cluster tools are quick to point out low availability as a drawback of clustered designs. However, part of this opinion may be perceptual. Cluster tools, by definition, are several tools connected together. The aggregate availability of any collection of tools will be lower than the individual tool availabilities. Some of the people that work with cluster tools tend to consider them as a single unit, leading to a lower perceived availability. The case

studies in Chapter 3 try to identify cluster tool availability issues that are not simply a result of considering multiple tools as a single unit.

2.4 TRENDS

Since the late 1980's when cluster tools were introduced into full production fabs, the number of cluster tools in use has grown steadily. However the development and proliferation of cluster tools has lagged behind early industry forecasts.

Early responses to the introduction of cluster tools included expectations that fully modular, independently controlled process modules would be linked together via an open architecture of hardware and software standards. The individual process modules would be independently controlled and communicate through a central cluster controller on the platform. The interchangeable nature of cluster tool modules would allow manufacturers to select and integrate process modules from different vendors and thereby obtain the "best of breed"[5] processing tools.

Since the early 1990's a group of tool vendors has been working to develop an open architecture like the one described above. For a number of reasons the modular equipment standards committee (MESC) standard architecture has taken longer than expected to develop, with agreement upon the final software interface standards only recently achieved[6]. Although the MESC standard is beginning to make inroads, a variety of proven platforms and process modules does not yet exist. The majority of cluster tools in use today are based on closed, proprietary architectures.

Cluster tools were initially expected to reduce the downtime associated with repairs. With fully modular processes, a down module could be wheeled away and replaced with a functional one[7]. This method of repairs has not become a reality. Because of the difficulty of disconnecting all the process module systems, particularly hazardous process gases, and the need to qualify the new module, this method would likely provide no time savings over repairing the down module in situ.

For the near future, cluster tools will continue to grow in capability and market share. The potential benefits offered by an open architecture will likely mean that the MESC standard will gain more and more acceptance. Although some enthusiasts have envisioned integrating hundreds of process steps, it appears that the complexity of cluster tools will increase only slowly as process development and increasing tool reliability make highly clustered tools practical. Rather than completely replacing conventional tools, cluster tools will continue to be used strategically in the areas where they provide a clear advantage.

3. Case Studies of Cluster Tools

The case studies presented in this chapter were selected to illustrate issues unique to (or exaggerated for) cluster tools and which have a detrimental effect on availability for production. The case studies include a description of the tool, the process performed with the tool and the availability issue(s) illustrated by the tool. Many of the issues identified apply to more than one of the tools discussed here, but the strongest or clearest example has been chosen in each case. The tools presented in the following cases are tools that Digital has used in production or is in the process of purchasing for future production.

3.1 Integrated Tungsten Deposition and Etchback Tool

The tool described in this case study illustrates how two fairly simple processes can interfere with each other when they are integrated into a cluster tool. These interferences are a product of integration and translate into reduced availability for the cluster tool.

3.1.1 Tool Description

The tungsten deposition and etchback cluster tool consists of four modules and two cassette stations surrounding a central load lock and wafer handler (See Figure 3.1 below.) Two of the modules, B and C are identically configured tungsten deposition chambers. Modules A and D are identically configured plasma etch chambers. This configuration is a combination of serial and parallel

processes; chambers B and A are used as a serial deposition and etchback process which runs in parallel with an identical process running on chambers C and D.

The load lock serves as an interface between the wafers in the two cassettes and the process chambers. While the deposition and etchback processes are performed at low pressures (measured in millitorr), the wafers in the cassette are at atmospheric pressure. The load lock cycles between atmospheric pressure-when the robot moves a wafer in or out of the cassette-and low pressure-when it moves a wafer in or out of the modules. An internal buffer holds a portion (up to eight) of the wafers under vacuum while they are processed. The internal buffer reduces the number of times the load lock must be vented to atmosphere and pumped down thereby reducing the time required to process a lot.

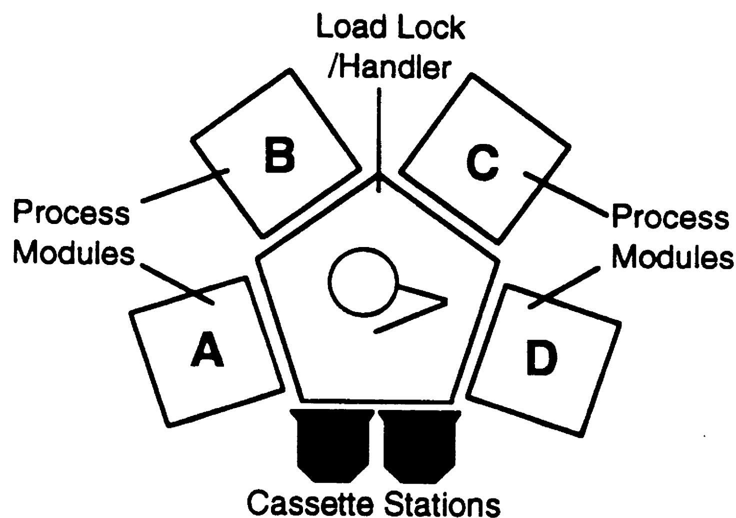


Figure 3.1. Configuration of Tungsten Deposition and Etchback Cluster Tool

3.1.2 Process

A single deposition and etchback sequence (or series of process steps) is performed on this tool. The following is a detailed description of how the wafers move through the tool during processing. After a cassette holding up to 25 wafers is loaded onto one of the cassette stations, the load lock vents to atmospheric pressure and the robot loads eight wafers into the load lock. The load lock then pumps down and the robot moves the wafer into one of the deposition chambers.

After the wafer has been deposited with tungsten the robot moves it back into the load lock. The robot then moves the wafer into one of the etch chambers. The robot continues to move wafers into the deposition and etch chambers as they become available.

This process continues until all of the wafers in the load lock are deposited and etched back. At that point, the load lock vents to atmosphere, the robot replaces the wafers in the cassette and another batch of (up to eight) wafers is moved into the load lock.

3.1.3 Issue Illustrated by This Tool

Processes that are integrated into a single tool can interfere with each other both physically and logistically. These interferences reduce the amount of time the process modules are available for processing.

3.1.3.1. Physical Interference: Particle Contamination

One source of physical interference between the two processes on this tool is particles. During the tungsten deposition process tungsten is deposited on the chamber surfaces as well as on the wafer . If this deposition is allowed to accumulate it will flake off and may settle on the wafer destroying the product. Plasma etching cycles can be used after deposition to remove the tungsten from the chamber, but etch by-products may also be a source of particulates.

Another source of particles is the wafer itself. Tungsten does not adhere well to the underlying oxide film. Prior to tungsten deposition, a titanium nitride "glue layer" is deposited on the oxide to improve tungsten adhesion. If there are any areas of the wafer which do not receive titanium nitride deposition (such as the backside of the wafer) that are exposed to the tungsten deposition, the tungsten will peel off and become a source of particles.

Finally, as mentioned before, the process of etching tungsten can produce by-products which condense as particles within the etch chamber. Thus the etch chamber can also be a source of particles.

Controlling particles on this tool often requires maintaining process parameters within a narrow operating window. If these parameters drift, particles may form in which case the tool cannot be used until the particles have been eliminated. The remedy for particles is often a "wet clean", which means venting the process chamber to atmosphere, opening it and wiping all the surfaces down with a cleaning solution (de ionized water and isopropyl alcohol are often used). The

chamber is then resealed and pumped back down to process pressure. Exposing the chamber to atmosphere and the cleaning solution allows the ceramic parts within the chamber to adsorb moisture. Upon pump-down this moisture out-gases, extending the time required to pump down to process pressure. The entire wet clean procedure can take up to 24 hours on this tool.

Particle concerns are more or less common to tungsten processing tools in cluster and non-cluster configurations. The complication with cluster configurations comes from the fact that the modules share a common load lock. When the wafer is transferred between the load lock and process modules it passes through a slit valve which separates the chambers. The load lock pressure is approximately one order of magnitude higher than the process chamber pressure. The pressure differential between the two chambers can cause a rush of gases to flow between the chambers when the valve is opened. Because particles can be redistributed by turbulent gas flows, particles generated in one module may move through the load lock and contaminate the other modules. Thus, multiple modules may require cleaning due to a particle problem with a single module. In non-clustered tools, the only mechanism for transporting particles between processes is the wafers themselves. For processes which are prone to producing particles, cleaning steps such as rinsing, may be performed between non-clustered processes to reduce the opportunity for inter-process particle contamination.

3.1.3.2. *Logistic Interference: Blocking*

Another type of interference that is exaggerated for cluster tools is blocking. When one process module fails, related modules in a multi-step process may be input or output constrained such that they are unusable; although they are capable of normal operation, flow of product through the operable modules is blocked.

Most semiconductor manufacturing lines include more than one tool capable of processing any given step. This is done either to achieve a higher capacity or to provide backup in case of failure. For a line using non-cluster tools, if a process module fails, it may be possible to re-route product flow to other similar modules to maximize through put. On the other hand, cluster tools running serial processes have production procedures and process recipes which are designed for running several process steps in a single integrated run. These conditions can make it difficult, or impossible, to take advantage of functioning modules when other modules on the same tool have failed.

3.2 OXIDE DEPOSITION CLUSTER TOOL

The oxide deposition tool considered in this case was one of the first modular cluster tools that Digital purchased. In fact it was likely among the first cluster tools of its type used in a production fab. The tool is configured with four processing modules arranged around a central loadlock and handler as shown in Figure 3.2. The original process developed on this tool was a complex serial process including all modules with some of them being used more than once in

the process. This tool is a clear example of the affect of process complexity on cluster tool availability.

3.2.1 Tool Description

This tool is used to deposit inter-metal dielectrics via chemical vapor deposition (CVD) and to etch the dielectrics via reactive ion etch (RIE). Three of the four process modules are used for oxide deposition, the fourth is used for etching. Module A is configured to deposit Type 1 oxide; module B is configured to deposit Type 2 oxide; module C is configured to deposit Type 2 and Type 3 oxide. Module D is capable of etching all three types of oxide.

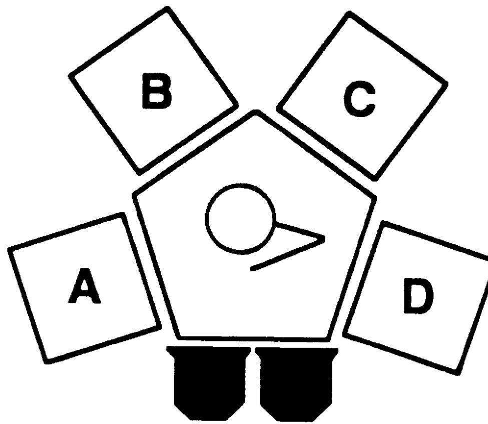


Figure 3.2. Configuration of Oxide Deposition Cluster Tool

3.2.2 Process

The manufacturing steps preceding oxide deposition create a very uneven surface on the wafer, making subsequent processing difficult. Thus the purpose of the oxide deposition process is not simply to deposit an insulating layer but

also to smooth or planarize the surface of the wafer. Planarization involves depositing and partially etching a series of different oxides each of which have different deposition and etch characteristics.

The process initially considered for this tool consisted of depositing Type 2 oxide, etching, depositing Type 3 oxide, etching, depositing Type 2 oxide, depositing Type 1 oxide, etching, depositing Type 1 oxide, etching, and, finally, depositing Type 2 oxide. To do this the wafers traveled from the load lock to modules B, D, C, D, C, B, A, D, A, D, and C. A full cassette of 25 wafers took almost 12 hours to process.

3.2.3 Availability Issues Illustrated by This Tool

This tool provides an extreme example of the complex processes that are possible on a cluster tool. The time required to perform process control and failure diagnosis on such a process can greatly reduce availability.

3.2.3.1. Process Control

To insure that this process ran successfully it was necessary that all steps ran within a narrow range of critical process parameters. Monitor wafers were routinely processed through individual steps to gauge process performance. Because there were significant interactions between some steps, it was necessary to run additional monitor wafers through selected combinations of steps as well. One or more measurements were performed for each wafer and process parameters were calculated from the measurements. The parameters were then

recorded on Statistical Process Control (SPC) charts which were designed to highlight any out-of-range parameters.

The procedure of monitoring and charting process parameters is standard for many manufacturing tools. For cluster tools such as this one, however, the procedure can be quite time consuming. For the process described above, a total of 12 monitor wafers were regularly processed and over 40 calculations were required. At one point in time, there were 76 SPC charts for this process.

3.2.3.2. *Failure Diagnosis*

Semiconductor tool failures are not always signaled by an alarm or a flashing light. Often, the signal that a tool has failed to operate correctly is a wafer, or batch of wafers, that fail a subsequent quality check, such as surface particles or film thickness. When this type of failure occurs on a cluster tool, its is necessary to determine which process module has failed. Diagnosis often begins with processing one or more of the process control monitor tests described above. If such monitors do not indicate which module is not operating correctly, diagnosis continues with additional tests designed to isolate the failure. If the nature of the failure is intermittent, it may be necessary to run several series of monitors to recreate the failure.

The greater the number of processing modules integrated into a tool, the more complicated and lengthy the failure diagnosis procedure. In the case of cluster tools like the oxide deposition tool, the diagnosis, or "trouble shooting", phase of a failure repair can take literally days.

The process engineers working with this tool were able to develop this process, however, with tool availabilities of less than 30%, they decided that the process was too complicated. In the end, this process was broken into several shorter processes which were divided between this tool and two others. The elegance of running several steps in a single process on a single tool is attractive, however keeping such a process under control and operating acceptably can present a great challenge.

3.3 PLASMA ETCH CLUSTER TOOL

This case study examines a plasma etch cluster tool that is used to etch various integrated circuit levels including oxide, nitride, resist, and polycrystalline silicon. In Chapter 1, cluster tools were described as being parallel, serial or a combination of both. In some ways the plasma etch cluster tool falls outside those definitions. The processing modules are used for different processes and as such are not parallel in the sense that they don't provide high capacity of a single process. Neither are the processes generally used in a serial manner with wafers visiting each processing unit. The tool is used to perform five different processing sequences which are applied at various points in the manufacturing process. This cluster tool is essentially four different tools which share a common handler and loadlock.

3.3.1 Tool Description

The tool consists of four processing modules and two cassette stations clustered radially around a central load lock and a single wafer handler as shown in Figure 3.3. All four modules are used for etching. Modules A and B are configured

identically, however they are dedicated to run two different processes because one of the processes has a considerable "seasoning" effect which makes the module on which it runs unsuitable for running the other process. Likewise, modules C and D are similar configurations, but are dedicated to different processes to insure consistency (C is used for two processes and D for one).

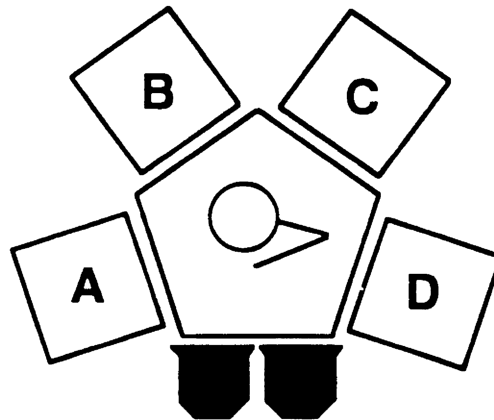


Figure 3.3. Configuration of the Plasma Etch Cluster Tool

3.3.2 Process

The five process sequences run on this tool are described in Table 3.1 below. As the table shows, processing module C is used in all sequences except number 2. Sequence number 2 requires only module B. The tool is capable of processing wafers from both cassette stations which allows process sequence 2 to be run on one cassette of wafers while any of the other process sequences are run on another. None of the other process sequences are run simultaneously as each requires module C and little would be gained over running the sequences consecutively.

Table 3.1. Process Sequences for Plasma Etch Cluster Tool

Number	Process Sequence
1	C_1, A
2	B
3	C_1
4	C_2
5	C_2, D, C_2, D

3.3.3 Availability Issue Illustrated by This Tool

The question of availability of cluster tools is often complicated by the fact that a tool can have several different types of availability. Table 3.2 lists the usable states for the plasma etch cluster tool. Some of these states provide such limited utility that, under most production policies, the entire tool would be turned over to repair. However, several of the states allow four or five of the process sequences to be run and it is feasible that the tool could be operated in those states, particularly if duplicate tools existed to provide the process sequences unavailable on this tool. Production decisions and policies are more complicated for tools that have multiple availabilities.

Table 3.2. Usable States of the Clustered Etch Tool

Usable States				**Sequences Available for Given State**				
Processing Module				Sequence Number				
A	B	C	D	1	2	3	4	5
0	0	1	0			√	√	
0	0	1	1			√	√	√
X	1	0	X		√			
0	1	1	0		√	√	√	
0	1	1	1		√	√	√	√
1	0	1	0	√		√	√	
1	0	1	1	√		√	√	√
1	1	1	0	√	√	√	√	
1	1	1	1	√	√	√	√	√

1= Processing Module is Up
0 = Processing Module is Down
X= Either Condition

Accurate capacity planning for a single tool or an entire fab requires an understanding of tool availability. Defining availability on this tool is not simple and thus accurately assessing availability presents a challenge. Because the various process modules have different reliability statistics, tool availability for each of the process sequences will be different. An accurate capacity planning system used with this tool would need to account for the availability of each process sequence.

3.4 HIGHLY INTEGRATED METAL DEPOSITION TOOL

The tool described in this section is used for depositing metal films. Physical vapor deposition (PVD) is a method of depositing metal films by condensing metal atoms out of a vapor phase onto the surface of the wafer.

The metalization process is particularly sensitive to contamination and oxidation, as such it is a prime candidate for integration. Depending on the specific metalization process, it may be desireable to deposit several layers of different metals (e.g.: aluminum, cobalt, titanium, tungsten etc) while keeping a wafer in a high vacuum, low particle environment. By integrating several metalization and cleaning modules into a cluster tool with wafer transport under vacuum, it is possible to achieve such a process.

The PVD tool considered in this section is one of the most highly integrated cluster tools available today. Most of the PVD tools in use today are capable of performing several processes in order to meet the metalization process requirements mentioned above. This tool, however, may be configured to perform several different process sequences with a high degree of complexity. It may be configured with duplicate process modules for parallel processing, different modules for serial processing, or a combination of both.

3.4.1 Tool Description

The configuration of this tool, shown in Figure 3.4 below, is very flexible. It can be configured with up to four ultra-high vacuum (10^{-9} torr) PVD chambers in

positions A-D. Positions E and F may be configured with high vacuum chemical vapor deposition CVD modules. Module G is a wafer orienter/de-gas chamber and position H may be configured with an optional cool-down module (wafers are heated during the deposition process and cool only slowly when held in a vacuum). I and L are transfer chambers equipped with robotic arms which move wafers between the other modules. Module J is a pre-clean chamber used to prepare the wafer surface for deposition. Finally Module K is a cool down chamber. Chambers I, J, K and L are built in to the main frame of the tool. At the bottom of the diagram two cassette stations in separate load locks are shown.

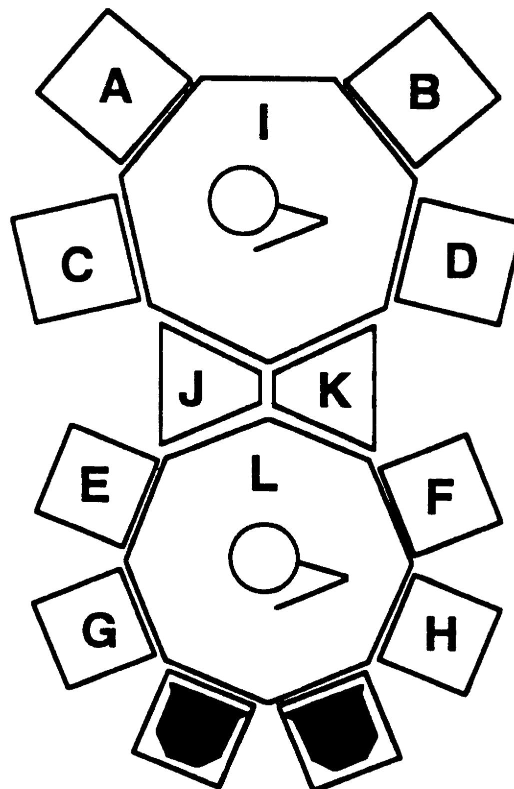


Figure 3.4. Highly Integrated Metal Deposition Cluster Tool

3.4.2 Process

3.4.2.1. Sputter Deposition

This tool performs physical vapor deposition of metals through a process known as sputter deposition. In sputter deposition, high energy ions are directed at a target which is made of the metal to be deposited. When the ions collide with the target, some of the atoms are dislodged (sputtered) from the surface of the target. The sputtered atoms are then transported to the wafer surface where they condense and form a thin film[8]. Because the target is being sputtered away, it must be periodically replaced.

3.4.2.2. Typical Process Sequence

This flexible tool is capable of many different process sequences and, as it is a new tool for Digital, specific process sequences have not been finalized. The following describes a typical processing sequence.

Cassettes of wafers to be deposited with metal are loaded into one of the load locks which is pumps down to 10^{-5} torr. A wafer is transferred into the buffer chamber which then pumps down to 10^{-6} torr. The wafer is then transferred into the wafer orienter which rotates it to a specific alignment to insure consistent processing. If the tool is configured with CVD chambers, the wafer might be transferred to one of them for deposition prior to PVD processsing. The wafer is then transferred to the pre-clean chamber where a high energy plasma is used to etch any native oxide from the surface of the wafer to insure low-resistance interfaces between deposited layers. The pre-clean is conducted at 10^{-7} torr.

The wafer is then moved into the transfer chamber which pumps down to 10^{-8} torr before transferring the wafer to one of the PVD chambers where deposition takes place, at 10^{-9} torr. The wafer may be subsequently transferred to a different PVD chamber to be deposited with a different type of film. After deposition is complete, the wafer is transferred to the cooldown chamber in position K. Once cooled, the wafer is transferred to the buffer chamber. If required, the wafer may be transferred to one of the CVD chambers for additional deposition or to the cooldown module in position H before being returned to the cassette. As the first wafer is transferred through the tool, the second wafer enters the module vacated by the first. The third wafer follows the second and so on until all wafers in the cassette have been processed.

The scheme of moving wafers through successively lower pressures saves processing time. When a vacuum chamber is vented to a higher pressure, time is required to pump it back down to its base pressure. The greater the pressure differential and the lower the base pressure, the more time required to pump down. In normal operation of this tool the modules are only exposed to pressure differentials of a factor of ten, thus pump down times are reduced. The exception to this is the load locks which cycle between atmosphere (760 torr) and 10^{-5} torr, a procedure which requires approximately 8 minutes.

3.4.3 Availability Issue Illustrated by This Tool

The tool described in this section illustrates the importance of maintenance policies for cluster tools. In addition to the target replacements mentioned above, periodic maintenance (PM) procedures, including replacing certain parts within

the chamber, are required after depositing 2000 microns (μm) of metal film. The physical act of replacing the target or PM parts is relatively quick, requiring less than one hour to complete. However the ultra-high vacuum chamber is exposed to atmosphere during replacement. Pumping the chamber back down to base pressure and conditioning the chamber takes approximately 8 hours. Tests are then conducted to verify that the chamber has been returned to an operable state.

The frequency of scheduled maintenance varies with the number of wafers being processed and the amount of film deposited on each wafer. As an approximate example, for an aluminum deposition chamber depositing 1 micron thick films on 20 wafers per hour, the chamber could run for approximately 100 hours before requiring a 2000 μm maintenance. Depending on tool utilization, this would mean opening each aluminum chamber for maintenance approximately once a week. In addition to scheduled maintenance, non-scheduled maintenance would require that the chambers be opened even more frequently. The effect of maintenance on tool availability depends on tool reliability, configuration and maintenance policy.

If the module being maintained is a one-of-a-kind module in a serial process, the tool is down for that process until the maintenance is complete. In this case, blocking similar to that discussed in Section 3.1 occurs and all process modules will be idle until the repair is complete. The tool user can choose to have duplicate processes for the high and ultra-high vacuum modules on this tool to avoid one-of-a-kind modules. However the wafer orienter, pre-clean chamber, transfer chambers are necessarily one-of-a-kind modules that are in series with

any redundancy. Clearly the success of this tool relies heavily on the reliability of these modules.

If the module being maintained is one of several duplicate processes running in parallel, the situation is more complicated. For primarily safety reasons, some manufacturers follow a policy of not operating a cluster tool while maintenance is being performed on any of its modules. For the PVD tool this policy means that, while PM procedure is performed, the tool is down for all processes, even if unaffected chambers are capable of operating. Typically, this policy only applies to the hands-on portion of the maintenance procedure and other maintenance phases, such as pump down (which frequently represents the largest portion of maintenance time) do not interfere with other, operable modules.

The decision to use the operable modules during the non-interference phases is not a simple one. Typically, the tool is configured to balance the throughput of the various processes. For example, a typical metalization process may involve depositing a thin TiN layer in addition to a much thicker layer of aluminum. Because the aluminum layer will take longer to deposit than the TiN, the user may choose to configure the tool with one TiN deposition module and three aluminum modules. If one of the aluminum modules is taken down for a target replacement and the tool user decides to run the tool during the pump down phase, the tool will no longer be balanced and will require more time to process a batch of wafers. With three modules requiring PMs on approximately one week intervals, the tool may spend a significant amount of time running in a "handicapped" mode.

Alternatively, the tool user may decide to synchronize the PMs on all chambers of one type such that they happen simultaneously. In this scheme the entire tool would be taken down for the length of time required for replacement and pump down after which the tool will be fully functional. Thus, it is never necessary to run in handicap mode. However, taking an entire tool down will require coordination with maintenance schedules of similar tools in the factory to insure that some metal deposition capacity is always available. [9]

Additional risk is associated with taking several chambers off line at the same time. When ultra high vacuum chambers are opened, resealed and pumped down there is some chance that the chamber will not return to an operable state. If the chamber fails the re-qualification tests, it may have to be reopened and adjusted. If several chambers are opened at once, the likelihood of having all chambers pump down without a problem is decreased.

Complicating all these decisions is the reality that unscheduled maintenance will also be required. If one tool fails while a similar tool has been purposely taken off-line for scheduled maintenance, the negative effect on product flow through the fab could be severe.

3.5 SUMMARY

The case studies presented in this chapter served to illustrate issues that are unique to (or exaggerated for) cluster tools and which have a detrimental effect on availability for production. The first case illustrated how integrated processes can interfere with each other in ways that non-cluster tools do not. Common

components such as load locks may allow one process to contaminate another. Similarly, serial process on cluster tools experience blocking when one module fails.

In the second case, the difficulty of controlling complicated processes on cluster tools was discussed. More complicated processes result in more process control work and less time spent running actual product. Failure diagnosis time was identified as another cost of integrating multiple processes.

The third case study illustrated the difficulty of understanding and even defining availability of cluster tools. When tools are capable of different processes in different states the question of availability itself is obscured.

The fourth case study illustrated how maintenance planning is a greater challenge for cluster tools. Careful consideration must be given to tool configuration and the planning of cluster tool maintenance in order to maximize tool availability.

This chapter was not intended to be an argument against cluster tools in general. The benefits of cluster tools were stated in Chapter 2. However, understanding the costs, in terms of availability, associated with clustering is essential to making the appropriate decisions regarding clustering. Guidelines for integrating processes on cluster tools, based in part on the results of the case studies, are presented below.

3.5.1 Guidelines for Integrating Processes on Cluster Tools

The simplest and most general guideline for clustering is that processes should be clustered when there is a clear technological or economic benefit and the carefully considered costs in terms of availability are small. Although this statement may seem obvious, excitement over cluster tools and the difficulty of assessing the costs may lead to clustering that does not provide a net benefit. Although technological and economic benefits have been discussed separately, all technological benefits can be translated into the economic benefits associated with the competitive advantage that the technology provides. The recommendations below refer to situations where the net economic benefits of clustering are not clear.

- Clustering is most successful when applied to established, well defined processes. Developing new processes on separate tools is challenging. When the complexity of integrated processing is added, these challenges increase.[10] Also, as evidenced in section 3.2, clustering processes in the development phase can require excessive process control.

- Different processes which are not sequential production steps should not be clustered, as was done on the tool in Section 3.3. This form of clustering offers no clear technological benefit and the economic benefit of a smaller footprint and shared platform will probably not compensate for the availability costs.

- When compared to serial processing, parallel processing configurations are less sensitive to several of the issues identified in this chapter. Inter module interference, process control, and failure diagnosis are less of a problem for cluster tools with identical process modules.

- When volumes allow, it is preferable to dedicate each tool to a single process sequence. Doing so simplifies production and maintenance planning as well as the task of assessing tool availability.

Implicit in these guidelines is the ability to accurately assess the availability of cluster tools in order to compare different tool configurations. Successful methods for accurately quantifying cluster tool availability have not been developed. [11] Chapter 4 considers the use of computer models as a method of estimating cluster tool availability.

4. Modeling Cluster Tool Availability

Making the appropriate decisions regarding the purchase and utilization of semiconductor manufacturing tools often requires analysis of complicated systems. One method that manufacturers use to perform this analysis is computer-based models that simulate tool performance for a given set of circumstances. Computer models could also provide a means for quantifying the effects of clustering on tool availability.

4.1 Applications of Semiconductor Tool Models

Computer models are used by manufacturers to simulate the performance of individual tools or entire fabs. In the case of individual tools, custom software may be developed, often by the tool vendor, to capture many details of the tool's performance. For the purpose of modeling entire fabs, sophisticated software is available in the form of generic discrete event simulators, or programs designed specifically for modeling manufacturing lines.

4.1.1 Modeling Individual Tools

Individual manufacturing tools are modeled to analyze tool performance characteristics such as cycle time, production rate, availability, cost of ownership, etc. The purpose of such analysis is often to compare one tool to another or to compare various possible configurations of a tool in order to select the equipment that best suits the manufacturer's needs. These models contain a great

deal of detailed information about the specific tool, often including precise simulation of wafer movement through the tool.

4.1.2 Modeling Entire Fabs

Some semiconductor manufacturers use computer simulations to model their entire fab. The models include representations of every piece of equipment in the fab and product flow is simulated typically on a lot by lot basis. The models include databases with thousands of parameters which summarize key performance characteristics for every tool in the fab. The purpose of these simulations is to estimate fab capacity and production cycle time. Fab models can also be used for designing fabs. Manufacturers may use models to consider various combinations of tools and layouts when planning a new fab or the expansion of an existing facility.

4.2 CHARACTERISTICS RELEVANT TO ACCURATE AVAILABILITY MODELS OF CLUSTER TOOLS

The value of a model has much to do with the clarity of its purpose.[12] The intended purpose of the modeling being discussed here is to provide a basis for comparing different tools or different tool configuration with respect to availability. This section considers the characteristics of cluster tools which may result in reduced availability when compared to non-clustered tools.

4.2.1 Individual Process Module Reliability Characteristics

Accurate modeling of cluster tools requires detailed information regarding process module reliability characteristics. Mean time to failure and mean time to

repair affect tool availability. The distributions underlying those means will also have an effect on tool availability particularly when considering module interactions. Preventive maintenance frequency and duration will affect availability. If a tool requires a requalification process, such as pumping down a vacuum chamber, mean time to re-qualify and probability of successful requalification will also be relevant statistics.

Reliability characteristics for the central platform should also be included in cluster tool models. Because the platform is used by all process modules its reliability has a potentially large effect on tool availability. Experience has not demonstrated that all cluster tool platforms are more reliable than the processing modules they support[13].

4.2.2 Interferences Between Process Modules

Section 3.1 identified interferences between process modules as an issue affecting cluster tool availability. In the case of serial processing, product flow between modules stop when one of the processes fails. Cluster tool models should recognize these inter-module dependencies and consider the entire tool as being unavailable when a single module fails. Physical interferences will also reduce availability. An accurate model would recognize processes that are more prone to failure when they are clustered with other processes.

4.2.3 Multiple Processes for a Single Tool

Accurate availability models would include descriptions of the process sequences performed on the tool. The model should be able to determine which processes the tool is available for in any given state.

4.2.4 Failure Diagnosis Delays

The time required to diagnose a failure on a tool varies with the complexity of the tool, particularly as a function of the number of different processes integrated on the tool. In light of this, an accurate model should include the mean time to diagnose a failure for the tool.

4.2.5 Maintenance Policies

Maintenance policies can affect tool availability in several ways. The PM schedule for a tool including the frequency and typical duration of a PM procedure will be included in an accurate model. If PMs are done early in the event of a failure repair or if PMs are postponed when there is limited capacity for that process, tool availability will change in response to these policies. Highly clustered tools may have a complex set of maintenance policies depending on the tool configuration, the state of the tool at the time of maintenance, and the production demands on the tool. Models that do not contain the details of such policies will be approximate at best.

4.3 RECOMMENDATIONS FOR ACCURATE AVAILABILITY MODELS OF CLUSTER TOOLS

Historically, the computer models that have been used to simulate semiconductor fabs have considered each tool as a separate entity with a set of parameters that characterize its performance such as set up times, cycle times, throughput rates, mean time between failures (MTBF), mean time to repair (MTTR), etc. As cluster tools have been introduced, the same software has been used to model these tools as well. Unfortunately, the capability of the existing software does not allow for easily modeling tools with multiple processes and multiple sets of performance parameters.

The publishers of some of the existing software are currently working on adjusting their models to accommodate cluster tools.[14] Some adjustments have been made by deriving aggregate performance characteristics to model all the process modules on a tool as one single unit. Aggregating setup times and cycle times can be done with out much accuracy being lost, but deriving a single MTBF or MTTR for a tool with eight process modules or five process sequences is omitting considerable detail. Applying the existing modeling technology to cluster tools has resulted in models getting simpler while reality has become more complex. Accuracy has likely suffered in the process.

Making models is often a question of trade offs between accuracy and tractability[12]. If the model is too much a simplification, then the predictions made with it will be inaccurate and of little practical use. On the other hand, if the model contains too much detail it will be difficult to use and again of little

practical use, regardless of accuracy. Difficulty of use includes collecting the data necessary to determine model parameters. The recommendations in this section are intended to address the relevant issues identified above while keeping the detail and complexity to a manageable level.

4.3.1 Incorporate Availability Modeling Into the Existing Tool Models

To avoid duplicating much of the work previously done to model semiconductor manufacturing tools, the existing models should be modified to include adequate availability analysis, rather than create separate models for the sole purpose of modeling availability. Likewise the work of collecting data for availability should be incorporated into the tool state tracking systems that are already in place.

4.3.2 Model Process Modules and Platform as Separate Entities with Interactions

Rather than attempt to aggregate process module characteristics, the process modules should be modeled as separate entities and the potential interactions between them should be modeled. Likewise, the platform should be recognized as an additional entity upon which all process modules are dependent. For example if a portion of a cluster tool fails, the model should consider all other dependent portions unavailable for production.

Modeling the availability of a tool with more than one process sequence will require details about the individual modules and the tool states which can support each sequence. Tools that combine serial and parallel processing may have several combinations of modules capable of supporting a sequence. This

recommendation means an increase in model complexity. The complexity in this case is justified; a single availability number for a tool that performs more than one process sequence is nearly meaningless.

The recommendation above refers to logistic interactions. Modeling physical interactions between process modules may be impractical. Contamination of one process module by another happens in an unpredictable manner. Trying to quantify this type of interference would necessarily include approximations. It is possible that physical interactions would be adequately represented by reduced MTBF. However this approach includes the risk of not identifying the increase failure rate as a result of clustering.

4.3.3 Delays for Monitoring and Failure Diagnosis

Manufacturing simulations should include delays or unavailable time for each tool associated with process monitoring and failure diagnosis. Process monitoring activities are typically done on a regular basis and should not add greatly to the complexity of the modeling task. Failure diagnosis on the other hand presents problems related to data collection. Initial failure diagnosis begins when a tool fault is first identified. However, repair efforts often include cycling between trouble shooting and repair work as potential failures are identified and corrected, with the exact nature of the activity not always being clear. The job of recording tool state, typically performed by manufacturing and maintenance personnel, would be significantly complicated if failure diagnosis identification was included.

As mentioned in the case of physical interactions, failure diagnosis may be more tractable if considered as part of the time required to repair a failure; this is much the way it is handled currently. However if failure diagnosis is not separately tracked, the ability to predict increases in diagnostic time as a function of tool complexity is lost. In light of the difficulty of collecting failure diagnosis data as part of a tool tracking system, a study is recommended to evaluate the sensitivity of diagnostic time to tool complexity.

4.3.4 Include All Relevant Maintenance Issues

The frequency and duration of scheduled maintenance should be represented in the simulation model. Time to re-qualify a process module for production should be added to all repair activities. If the probability of successful requalification is not explicitly included in the model, the average time to re-qualify should be adjusted to reflect the reality that requalification may occasionally be repeated. The conditional nature of some maintenance policies will require sophisticated simulation models. The sensitivity of tool availability to such policies should be considered before they are incorporated into a model.

4.3.5 Recommendations for Additional Research

This research identified factors which affect the availability of cluster tools, it did not determine the significance of these factors. It is possible that some of the factors have only a minor influence on availability and therefore the effort of including them in tool model and collecting the requisite data would not be justified. Additional research should be conducted to assess the impact that each of these factors has on availability. Such research would include collecting data

relating to the issues identified here and simulating various tool configurations to analyze the sensitivity of availability to each factor.

5. NEW TOOL ACQUISITION AND INTEGRATION

This chapter documents a procedure used at Digital Equipment Corporation to acquire a new semiconductor manufacturing tool and to integrate it into an existing CMOS production line. Digital's Fab 4 is a relatively small volume, pilot line fab, used for developing new technologies and delivering first revenue product shipments. Approximately 4 months prior to the acquisition, new capacity targets for Fab 4 were established which exceeded the maximum capacity of several portions of the fabrication line. These targets were to be met by the end of the following year, in time to begin production of the next generation of CMOS products.

One of the processing areas which required additional tools to meet the capacity targets was tungsten deposition and etchback. A new set of tools were purchased to augment the capacity of the existing tungsten tools. They were manufactured by a different tool vendor than the existing tool set, thus the integration activities included process development work to find a process for the new tools which would produce acceptable results. The author was a member of the process startup team assigned to these new tools. The team consisted of process and equipment engineers from Digital, and representatives from the tool vendors.

Tungsten deposition and etch tools were both purchased at this time. However, this chapter will focus on the acquisition and integration of the tungsten etching tool. Section 5.1 describes the decisions leading to the acquisition of the etchback

tool and Section 5.2 relates the experiments performed to integrate it into the production line. In section 5.3, a discussion of these events is presented.

Much of this chapter refers to specific details of tungsten processing. Readers unfamiliar with applications of tungsten in integrated circuit fabrication and plasma etch process are referred to Appendix A for background on this material.

5.1 ACQUISITION

Manufacturing tool acquisitions at Digital are managed by a team of employees including process engineers, equipment engineers, production supervisors, and purchasing representatives. The team is responsible for evaluating the tools which are currently available and selecting the tool which best suits Digital's needs. Team activities include collecting industry data relating to the tools, visiting tool vendors, reviewing tool demonstrations, and negotiating tool specification and price.

5.1.1 Selection

In considering possible alternatives for expanding tungsten deposition and etchback capacity, the tool selection team quickly narrowed the field to two viable choices. The first and most obvious alternative was to purchase additional equipment from the supplier (referred to below as Vendor 1) of the existing integrated tungsten deposition and etchback equipment. Because Digital was familiar with the process used on these tools, this alternative presented the lowest risk. Integrating this tool into the fabrication line would be fairly simple.

However the existing tool set had relatively low throughput rates and had a history of poor tool availability.

The second alternative, which the tool selection team considered, offered increased throughput and higher availability. This alternative consisted of separate deposition and etchback tools: a non-integrated process. The deposition tool, a relatively new system manufactured by Vendor 2, offered a high throughput rate and in recent trials had proven to be a very reliable tool. The etchback tool, manufactured by Vendor 3, also offered improved throughput rates over Vendor 1's tool and was based on Vendor 3's successful line of aluminum etchers which industry experience had shown to be reliable as well.

An additional advantage of this alternative was the corrosion prevention capabilities of Vendor 3's etch tool. The new integrated circuit manufacturing process that Digital was developing was more technically demanding than the previous process. Digital engineers felt that, to meet the requirements of the new process, the tungsten etch tools would have to provide helium backside cooling as a means of recess control (see appendix A). The existing tungsten etch tools did not have backside cooling capabilities and attempts by Vendor 1 to provide helium backside cooling had been fraught with corrosion problems. On the other hand, the Vendor 3's tool, with its rinse module, offered corrosion prevention.

The main drawback to this tool set was that there was little industry experience using the tools in a combined tungsten deposition and etchback process. Early attempts by Vendor 2 and Vendor 3 to demonstrate a process to the tool selection

team had failed to produce acceptable results. In the end, Vendor 2 and Vendor 3 convinced the tool selection team that they could work together and develop a process capable of meeting Digital's specification. The decision was made to purchase deposition and etchback tools from Vendor 2 and Vendor 3. The chronology of the tool integration processed will be referenced from the date the purchase decision was made.

5.2 INTEGRATION

Introducing a new tool into a semiconductor fabrication line occurs in several stages and involves the efforts of different functional groups at each stage. Facilities personnel put the tool in place and connect the gas, water, and electrical systems. Equipment engineers calibrate the tool and test its functionality. Process engineers develop the procedures and tool settings, or "recipes" that result in acceptable tool performance. Production personnel then use the tool to process product.

5.2.1 Typical Startup Schedule

Digital's procedure for testing a new tool is divided into three stages or "levels". The levels are characterized by the type of functionality being tested and by the parties responsible for the tool at that level. The levels are sometimes modified to meet the particular requirements of a tool startup and are often not distinctly separate levels. However they serve as a guideline of the major objectives and responsibilities of the startup procedure.

5.2.1.1. Level 1: Mechanical Verification

Within Level 1, equipment engineers from Digital and the tool vendor calibrate the tool and demonstrate that the various subsystems are mechanically functioning according to specification.

5.2.1.2. Level 2: Vendor Process

Within Level 2, process engineers from the tool vendor operate the tool according to a standard process and compare the actual results to the expected results. Level 2 is intended to demonstrate that the tool is functioning according to the vendor's standard performance specifications.

5.2.1.3. Level 3: Digital Process

Within Level 3, Digital process engineers run experiments to characterize the tool's operating performance and to identify a range of operation capable of meeting Digital's performance specifications as agreed upon at the time of purchase.

5.2.2 Process Development

At the time of the purchase, the understanding was that Vendor 2, Vendor 3 and Digital would work together to develop an acceptable tungsten deposition and etchback process by the time the tools were delivered and installed in month three (after the purchase decision). The development work utilized the deposition tool in Vendor 2's lab in California, the etch tool in Vendor 3's lab in California, and several wafer preparation and analysis tools in Digital's fab in Hudson.

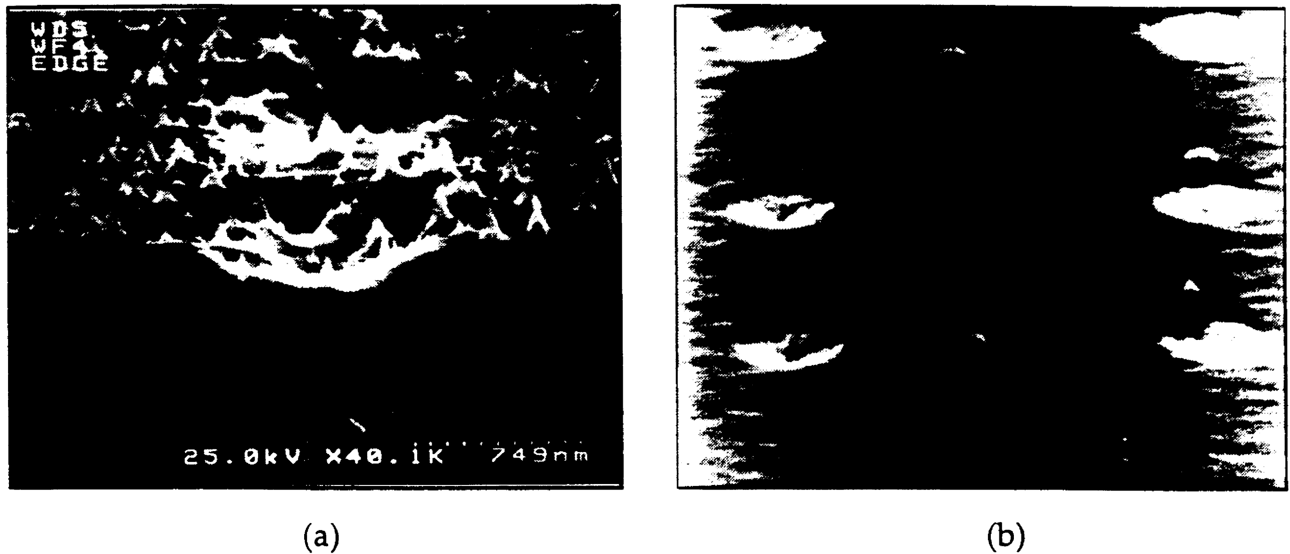
5.2.2.1. Tungsten Etch Development

The first process development experiments focused on the tungsten etch process. Initially, a two step tungsten etch process was proposed. The first step consisted of a high etch rate, high uniformity etch to remove the majority of the film (~80%) in a timed etch. The tungsten:TiN selectivity of this etch is relatively low, so the second step consisted of a slower etch which has a higher selectivity. In this way the tungsten film is quickly removed while exposing the tungsten/TiN interface to only a highly selective etch.

5.2.2.2. Residues

In the course of these experiments, scanning electron micrograph (SEM) images of etched wafers showed the presence of residues on the TiN film after the tungsten had been etched away. Figures 5.1 are SEM images showing samples of these residues. The residues were of unknown composition and were concentrated in a ring 5 to 15 millimeters from the edge of the wafer. Two fairly distinct types of residues were observed. One type was small, <0.1 micron, conical in shape, and were referred to as "pyramids", see Figure 5.1(a). The other was typically larger, ~0.5 micron, and less abundant. They were irregularly shaped and were referred to as "nodules", Figure 5.1(b).

Analysis of the composition of such small residues is difficult. EDX analysis of residue-containing samples has shown the presence of oxygen, titanium, fluorine, aluminum, sulfur, and tungsten. Speculation as to the composition of the residues has favored titanium-fluorine-oxygen compounds which form as a by-product of the etch process and deposit on the wafer.[15]



(a) (b)

Figure 5.1 Tungsten Etch Residues on TiN; (a) "pyramid" residues, (b) "nodule" residues.

The residues were resistant to the etch chemistry and appeared to be masking the tungsten etching process resulting in a severely roughened TiN surface, as shown in Figure 5.2. Digital engineers feared that the residues and the roughened surface would cause poor electrical contact between the tungsten plug and the next layer of metal. Whether or not the residues affected contact resistance, leaving unknown material on the wafer presented unacceptable risks to the process as a whole.

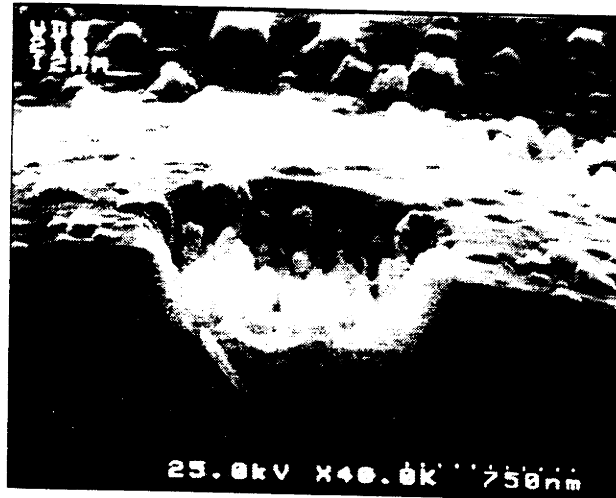


Figure 5.2 Residue and Roughened TiN Surface After Tungsten Etch

Attempts to remove the residues by extending the etch time resulted in excessively recessed plugs and roughened TiN surfaces as shown in Figure 5.3.

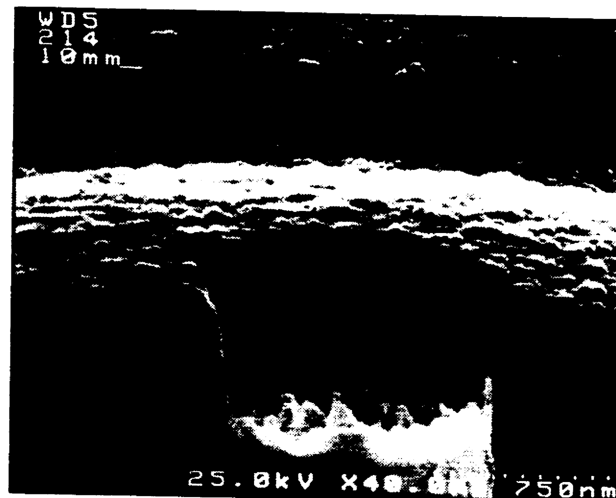


Figure 5.3. Tungsten Plug Surface After Extended Etching of Residues.

Efforts to eliminate residues then shifted to identifying the source of the residue constituent materials. The fluorine component of the residues was not surprising as it is the etchant species in the process. The presence of titanium would normally be suspected because the tungsten etch, though selective , attacks the TiN layer as the tungsten is clearing. However, the fact that the residues appeared to mask a portion of the tungsten etch suggested that the residues formed earlier in the etch process, before the TiN adhesion layer was exposed. The presence of oxygen was unexpected as well; typically oxygen is liberated from the underlying oxide only at the end of the TiN etch stage.

Further investigation showed that the titanium and oxygen present during the tungsten etch were artifacts of the combined deposition and etch processes. The Vendor 2 deposition process uses an edge exclusion technique to insure that tungsten was deposited only on the top surface of the wafer (tungsten deposited on the edges and back of the wafer would tend to flake off and form particles). The edge exclusion leaves approximately 3 millimeters of the TiN adhesion layer exposed at the edge of the wafer. The wafer clamp in the etch chamber covers 2.25 millimeters at the edge. This leaves approximately 0.75 millimeter of the TiN exposed to the tungsten etch. The TiN layer is so much thinner than the tungsten film that even with a selective etch, the TiN regions at the edge of the wafer can be completely etched away during the tungsten etch stage thus allowing the fluorine based etch to attack the oxide. Figure 5.4 illustrates the conditions at the edge of the wafer. The theory of residue generation at the edge of the wafer fit well with the observation that residues were concentrated in that region.

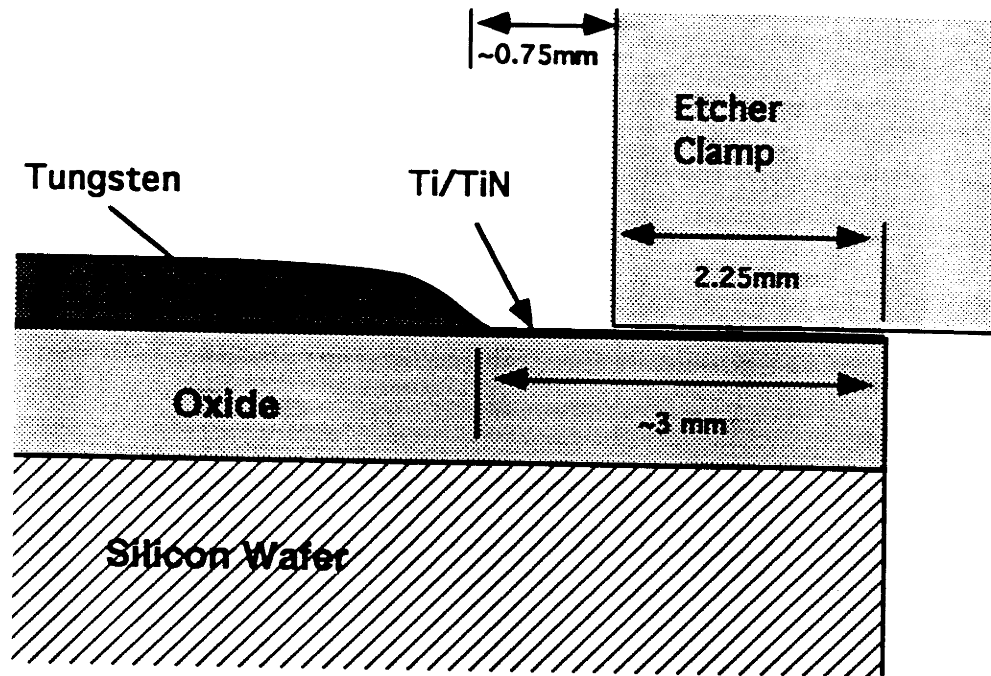


Figure 5.4. Cross Section of the Wafer Edge. The etcher clamp leaves a portion of the TiN exposed to the tungsten etch (diagram not to scale).

Eliminating the source of residue components proved to be difficult. Although it was possible to adjust both the size of the edge exclusion on deposition and the etch clamp ring, doing so did not present a viable solution. If the edge exclusion was controlled to be less than the area covered by the clamp ring, tungsten would lie under the clamp and would remain after etching. Digital's engineers were not willing to leave tungsten on the wafer after etchback because the effects of residual tungsten on successive processes were unknown. The positioning of the wafer varies slightly in both the deposition and etch processes. Consequently,

trying to match the size of the edge exclusion and the clamp ring was likely to produce both exposed TiN and residual tungsten.

Process development efforts continued with attempts to find a set of etch parameters that eliminated the formation of residues or prevented deposition of the residues on the wafer. As part of the efforts to isolate residue effects, the tungsten etch was changed to a single-step etch to reduce the number of variables in the process. From this point forward, all process development work involved a single-step tungsten etch.

Residue-reduction efforts were frustrated by the fact that residue formation occurred intermittently. One group of wafers would produce residues upon etchback while another group of wafers, identically etched, would not. Some evidence indicated a deposition dependence in the formation of residues, particularly pyramids - one group of wafers deposited at the same time consistently produced these residues. However what that dependence might be and how to deposit residue-free tungsten were unanswered questions at the time the tools arrived at Digital's fab in Hudson.

The process development work performed in California served to find a feasible operating region for the tungsten etch and identified the residue issue as a critical challenge to the process startup. The startup team decided that tungsten etch process development would continue at Digital after the tools were installed and development of the TiN etch process would continue at Vendor 3's lab in California.

5.2.3 Installation & Mechanical Startup

The tungsten etcher was placed in the fab at Hudson on the 10th day of month three. The facilities team performed the installation work of connecting gas lines, power systems, water systems and vacuum systems. On the 20th day of month 5, Digital's safety inspection team approved the installation and Level 1 of the startup began.

Equipment engineers from Digital and Vendor 3 performed the Level 1 work of characterizing and calibrating the tool's mechanical and electrical systems. The gas, rf and vacuum systems were calibrated to the vendor's specifications. System control and wafer handling functions were exercised to verify proper operation. Finally, wafers deposited with tungsten were partially etched to demonstrate overall tool functionality. On the 30th day of month 5, Digital's equipment engineers accepted the tool as operational and it was turned over to process engineering for process startup.

5.2.4 Process Startup:

Process startup activities were essentially a series of experiments performed on test wafers to identify a set of process parameters and sequences that would meet Digital's specifications for tungsten etchback. Initial experiments focused on etch rate and uniformity as the key characteristics of an acceptable process. Later, more detailed experiments were planned to examine the other important etch characteristics of the etchback process, (see Appendix A.)

Etch rate was determined by measuring the thickness of a tungsten film before and after etching for a fixed time. To determine film thickness, sheet resistivity was measured by 4-point probe and converted to thickness using the following formula:

$$\rho = T \times R_S$$

where: ρ = bulk resistivity of tungsten, assumed to be 9.5 $\Omega \bullet$cm,

T= film thickness, and

RS = measure sheet resistance.

The difference in pre- and post-etch measurements was then divided by etch time to determine etch rate.

Etch uniformity was calculated by measuring pre and post etch film thickness at 46 sites on the wafer (according to the process described above). From these measurements etch uniformity was expressed as the percent deviation (1 σ) from the mean among the 46 measurements, thus a lower uniformity percentage indicates a more uniform etch.

Three types of test wafers were prepared for use in the experiments. Blanket tungsten wafers were prepared for etch rate experiments. 3500 Å of thermal oxide was grown on bare silicon wafers which were then deposited with ~1200Å of TiN adhesion layer. The wafers were deposited as they were needed with ~10,000 Å of CVD tungsten on Vendor 2's tool.

Planar plug wafers were prepared for experiments intended to assess plug recess. 20,000 Å CVD oxide was deposited on bare silicon wafers. The wafers were then patterned with a mask including 1 μm vias. The vias were then etched into the oxide to a depth of approximately 2 μm. The wafers were deposited with a TiN adhesion layer and tungsten as described above.

A third type of wafer was prepared to simulate the most severe surface topography that would be present at plug layers in the actual product circuits. These wafers were to be used to test the ability of the etch to clear away stringers from steps in the underlying topography while maintaining minimal plug recess. These topography plug wafers were prepared in a similar process as the planar plug wafers with the addition of a layer of patterned metal (Al/Cu) under the oxide to produce topography.

5.2.4.1. Transfer of the California Process

The development work at Vendor 3's lab in California had identified a tungsten etch process with acceptable etch rate and uniformity. The etch used process parameters of

270 millitorr /320 watts / 140 sccm SF6 / 20 sccm Ar / 16 torr Helium

and temperature settings of 60 °C, 45°C and 60 °C on the upper electrode, lower electrode and chamber walls, respectively. This etch, referred to as the "California etch process" produced an etch rate of approximately 6000 Å/min. with uniformity better than 4% (1 sigma).

Plans for Level 2 of the tool startup included transferring the California etch process to Digital's tool in Hudson to demonstrate that it functioned in a manner comparable to Vendor 3's tool in California. Initial tests of this process on Digital's tool did not produce the same results as the tool in California. The etch rate was approximately the same, but the uniformity was worse than 6%. In addition, Digital's tool etched slightly slower at the center of the wafer, while the tool in California etched slightly faster in the center. This issue was important because the deposition process had been optimized to deposit tungsten slightly thicker in the middle of the wafer to compensate for the center fast etch process in order to achieve uniform plug recesses across the wafer.

Several differences between the two etchback tools were identified. The vacuum line between the pumping manifold (see Figure A.5) and the vacuum pump on Digital's tool was longer and had a smaller diameter than the tool in California. The rf power cable on Digital's tool was also longer. A baffle plate, situated above the upper electrode for the purpose of diffusing the gas flows was of a different configuration. Finally an orifice ring was identified in the chamber that was actually part of the aluminum etcher configuration (recall that this tungsten etch tool was a modified aluminum etcher) which was not required in the tungsten etch tool.

The pumping efficiency for a vacuum system varies, depending on the gas being pumped. In plasma etching, varying pumping efficiency can result in different residence times for the etchant, and by-product gases thereby changing the etch characteristics. Vendor 3's process engineers demonstrated that the tools in

Hudson and California had different pumping efficiencies using the following process. Individually, the process gases were flowed at a set rate with the throttle valve set fully open (maximum pumping capability). The chamber pressures measured under these conditions provided a means of comparing the pumping efficiencies of the two tools. For all process gases, the tool in California achieved a lower chamber pressure than the tool in Hudson for identical gas flows, indicating that the pumping configuration in Hudson was less efficient.

The most likely cause of the reduced pumping efficiency was the longer, smaller vacuum line. The vacuum line was longer because the facility at Hudson required that the pump be located on the building level below the fab rather than next to the tool, as was the case in California. The diameter of the vacuum line had been calculated to provide the required pressure at the required flow rates. Although capable of achieving the required pressure, the longer, smaller line could be expected to have a lower pumping efficiency.

This line is a stainless steel pipe that is specially fabricated and welded in place. Replacing it with a larger diameter vacuum line would be expensive and time consuming. Likewise, shortening the line by moving the pump closer to the tool would be expensive and difficult. The startup team decide to try to compensate for the existing line by making adjustments to the etch process.

The other differences between the two tools proved to be easier to correct. A new baffle plate which concentrated the process gas flows to the center of the chamber was installed with the expectation that this would produce a center fast

etch. The orifice ring which presented a possible gas flow restriction was removed. The rf power control system was re calibrated to adjust for the power loss in the longer cable.

After these modifications were made, the following procedure was used to compensate for any remaining difference in pumping efficiency due to the unchanged vacuum line. With the throttle valve fully opened as in the tests before, the process gases were separately flowed into the chamber and the flow rates were adjusted until the pressure in the chamber matched the pressure observed in the original test on the California tool. Due to the reduced pumping efficiency, the Hudson tool achieved a given chamber pressure for lower gas flow rates than the California tool. The ratio of the gas flow set points resulting in the same chamber pressure for both tools was used to adjust the process recipe gas flows. This adjustment, though recognized as approximate, was intended to achieve similar residence times for process gases and by-products for the two tools.

The adjusted tungsten recipe was as follows:

270 millitorr /320 watts / 100 sccm SF_6 / 14 sccm Ar / 16 torr Helium.

The backside helium flow was not adjusted as it has little affect on residence time.

The new etch, referred to as the "Hudson etch process" was tested and found to have an etch rate of ~6400 Å/min., etch uniformity of 2.3% and was slightly center-fast. These values were a moderate improvement over those observed for the California etch process. With these results, the process was considered to be transferred.

SEM photographs were taken of several of the wafers etched in the experiments described above. Nodule residues were observed near the edge of some of the wafers. Typically, transferring the standard process would conclude Level 2 of the tool startup process. However, because of the presence of residues after tungsten etch and the lack of a proven TiN etch stage, Vendor 3 process engineers continued to be involved in the process development efforts described below.

5.2.5 Process Refinement

With the tool functioning properly, process startup work shifted to refining the process. Experiments were performed to define tungsten and TiN etch processes that met all important characteristics of the etchback process. Attempts to understand and eliminate residue formation continued.

5.2.5.1. TiN Screening Experiment

TiN etch development experiments conducted in Vendor 3's California lab had indicated that nodule residues were etched away in the chlorine etch plasma. Based on the progress which had been made on the tungsten etch development

and the presence of residues, the startup team decided to proceed with a TiN etch experiment next.

The experiment was a 5 factor, fractional-factorial experiment of 10 runs (2 center points) designed to model main effects only. The experimental factors were pressure, power, backside helium pressure, total process gas flow (CL_2, Ar), and Cl_2/Ar ratio. Total flow and ratio were chosen as factors rather than separate flow rates for Cl_2 and Ar because Vendor 3's experience was that total flow and ratio modeled better as etch process factors. The TiN etch process developed in California was used as a center point for this experiment. This process was:

105 millitorr /450 watts / 16 sccm Cl_2 / 24 sccm Ar / 8 torr Helium.

Responses for the experiment were etch rate, and selectivity.

The experiment was performed by first etching blanket tungsten wafers in the Hudson etch process with no overetch and then etching in the TiN experimental set points for 100 seconds. Because the TiN layer is very thin the method used in the tungsten etch experiments of measuring pre- and post-etch film thicknesses was not applicable to the TiN etch. Instead, the optical emission endpoint traces were examined to estimate etch time and uniformity. (See Appendix A for a typical endpoint trace).

As the TiN film is cleared, the level of N_2 in the etch chamber drops indicating etch endpoint. The time to endpoint combined with a known TiN film thickness

was used to estimate etch rate. There is some error involved in this method as the tungsten etch process does not end precisely when all tungsten has been removed. Due to non-uniformity a small amount tungsten may remain in some areas while the TiN may be partially etched in others. Thus tungsten etch endpoint accuracy will affect the TiN etch time. Despite this known error, the TiN etch rate data modeled fairly well, as shown below.

Etch uniformity measurements were also difficult to obtain. The slope of the endpoint trace was used as a surrogate for etch uniformity. As mentioned above, the N_2 level drops as the TiN film clears. If the etch is uniform and the TiN clears from all areas of the wafer at the same time the endpoint trace will drop off sharply. Conversely, if the etch is non-uniform TiN will clear at different times resulting in a gradual endpoint slope. In addition to the error introduced by the tungsten etch, these uniformity measurements were rather subjective as to the portion of the trace used to estimate slope. Consequently the etch uniformity data modeled only approximately (R squared adjusted ≈ 0.5) as a function of power.

Detailed experiment design and results are presented in Appendix B. The general trends from the experiment are given in Table 5.1 below. The arrows indicate the direction of the response variable for an increase in the factor. Dashes indicate factors that did not model well as contributing to the response data.

Table 5.1 TiN Etch Experiment Trends

Responses	**Factors**				
	Pressure	Power	Total Flow	Cl2/Ar	Helium
Etch Rate	↑	–	↑	↑	–
Selectivity	–	↑	–	–	–

SEM images were taken of several sites on each wafer. All wafers showed some residues and oxide film roughness, these being concentrated at the edge of the wafer. Interpretation of these observations favored a theory that the residues observed after tungsten etch were being only partially etched in the TiN etch process and were masking the TiN etch leaving residual TiN and an uneven oxide surface.

Based on the result of this experiment, the startup team selected the following recipe for the TiN etch:

90 millitorr /550 watts / 10 sccm Cl_2 / 20 sccm Ar / 8 torr Helium.

5.2.5.2. Tungsten Etch Experiment

To further characterize the Hudson etch process, a 5 factor, fractional-factorial experiment with 18 runs (2 center points) was designed and run. The experiment factors were chamber pressure, rf power, total process gas flow (SF6, Ar), SF6/Ar flow ratio, and backside helium pressure. Responses were etch rate and uniformity, measured as described above.

The responses for this experiment modeled well. The general trends from the experiment are given in Table 5.2 below. Detailed experiment design and results are presented in Appendix B.

This experiment identified an optimal process capable of an etch rate of 7200 Å/min. with better than 2% uniformity. However, that process proved to be unacceptable in terms of selectivity. Wafers etched to tungsten endpoint, with the intent of stopping on the TiN layer revealed that the TiN layer was completely etched away, or "punched through", on portions of the wafer. Similar tests with the center point, 6300 Å/min., process did not result in similar punch through.

SEM images of the wafers from the tungsten etch experiment showed substantial residues including the first observation of pyramidal residues on wafers deposited and etched on Digital's tools at Hudson.

Table 5.2 Tungsten Etch Experiment Trends

Responses	**Factors**					
	Pressure	Power	Total Flow	Power•Press.	Power•T.F.	Press.•T.F.
Etch Rate	↓	↑	↑	–	↑	–
Uniformity	↓	↑	↑	↓	–	↓

5.2.5.3. *Residue Control Experiments*

Several single wafer tests were run with varying chamber temperatures in an attempt to eliminate residue deposition. These tests consisted of etching wafers in the tungsten-etch process or combined tungsten and TiN processes with reduced upper electrode or chamber wall temperature settings.

The expectation was that residues would tend to deposit on cooler surfaces. With the previous settings, the wafer was the coolest surface in the process chamber and therefore the most likely place for residues to deposit. Increasing lower electrode temperature was not explored as a potential solution because doing so would likely affect etch characteristics significantly. Although cooling the other surfaces in the chamber below that of the wafer was a possible method of reducing residue deposition on the wafer, it presented other risks. If the residues were deposited on the upper electrode or chamber walls, over time these deposits might flake off and become a source of particles.

The results of these experiments were confounded by the intermittent nature of residue deposition. Upper electrode temperatures ranging between 35°C and 70°C, and chamber wall temperatures between 30°C and 45°C were tested. No conclusive results were obtained from these tests.

5.2.5.4. Deposition Dependence of Pyramidal Residues

A third experiment was designed to examine the trade off between etch rate and W:TiN selectivity which had been identified in the tungsten etch experiment. However, during the initial runs of this experiment, unusual amounts of residual tungsten were visually observed on the wafer after etching to tungsten endpoint (stopping on TiN). Residual tungsten had not been observed previously when running essentially the same process recipe.

As these results were being confirmed, Vendor 2's deposition tool experienced faulting in it's vacuum pump systems. Further investigation determined that the deposition chamber exhaust line had become clogged with deposition by-products. Exhaust line clogging on the deposition tool represented a potential source of deposition-dependent residues. A clogged exhaust line could increase residence time of deposition by-products, with the potential for these materials being incorporated into the film. The exhaust line was cleared and modified to prevent future clogging.

New wafers were deposited with tungsten after the deposition tool exhaust line was cleared and etched in the Hudson etch process, which had previously produced residues. No pyramidal residues were observed; only minimal nodule

residues were present. These results were confirmed by etching a wafer that had been deposited with tungsten before the exhaust line was cleared, in the same process. Pyramidal residues were observed on this wafer. These observations and the apparent deposition dependence of residues observed in California served to confirm the belief that pyramidal residues were caused by certain deposition conditions. The mechanism for the formation of these residues was not identified, however pyramidal residues were not observed after the exhaust line was modified.

Additional tests were run with reduced upper electrode and chamber wall temperatures. Wafers were etched in the Hudson etch process and the TiN recipe identified in the TiN etch experiment. The upper electrode temperature was set at 35 °C, the lower electrode was set at 45 °C, and the chamber walls at 30 °C. The chamber walls, being resistively heated, were not capable of achieving temperatures below the ambient environment; the 30 °C setting was achieved by essentially turning the heaters off. SEM images of these wafers showed no residues. Despite the risk mentioned before of residues depositing in the chamber and producing particles, the process startup team decided to continue using these settings as a residue-prevention strategy.

5.2.5.5. *Overetch Experiment*

After the elimination of residues, several planar plug wafers were etched in the Hudson etch process, including the baseline TiN etch process. 60 second overetch times were used for both etch stages. SEM images of these wafers showed acceptable plug loss and TiN trenching. Measurements from cross sectional SEM

images estimated plug recess at 3000 Å and TiN trenching extending approximately 1000Å below recess level.

To further investigate the capabilities of the etch process, an experiment was conducted using varying overetch times to etch topography-plug wafers. After etching, the wafers were visually inspected in the SEM tool to assess plug loss and to look for tungsten stringers on topography. Cross sectional images were not available on this SEM tool, so TiN trenching was not visible and only visual estimates of plug recess were possible.

The experiments included test runs with overetch times between 5 and 30 seconds for each of the etch stages. Stringers were observed on only the wafers with the shortest overetches, 5 seconds tungsten overetch and 5 seconds TiN overetch. Wafers etched with 20 second overetches on both stages showed excellent plug recess, < 1000Å and no visible stringers. Figure 5.6 is a SEM image of plugs produced with the 20 second overetch process.

5.2.5.6. *Corrosion Prevention Experiments*

Concurrent with the process refinement activities described above, several corrosion prevention experiments were conducted. For these experiments wafers were prepared by depositing blanket layers of oxide, aluminum, TiN, and tungsten to bare wafers in that order. The wafers were then processed through the Hudson etch process(including the new TiN etch), exposing the aluminum to chlorine at the end of the process. Some of the wafers were rinsed in the rinse module and others were left unrinsed as a control group.

Figure 5.6. SEM Image of Tungsten Plugs from Overetch Experiment

The wafers were then left exposed to the fab environment for several days. Periodically, the wafers were measured with a laser scattering surface particle counter to check for particles caused by corrosion. The control wafers quickly corroded before the first particle counts were taken, in a period of minutes. The wafers rinsed in the rinse module showed no signs of corrosion.

In the course of these experiments it was observed that the rinse module would occasionally add a large number (>100) of particles to the wafer surface. After numerous tests, equipment engineers identified bacteria in the de-ionized water system as the source of particles. The particles were eventually eliminated by installing additional filters to the incoming rinse water line, increasing the water pressure, and purging the tool's water system with hydrogen peroxide.

5.2.6 Process Verification

To verify the capability of the Hudson etch process, batches of wafers that were scheduled to receive the entire fabrication sequence (referred to as "product lots") were split between the old and new process. Half of the wafers in each lot were deposited with tungsten and etched back on Vendor 1's tool and the other half were processed through Vendor 2's and Vendor 3's tools. After the tungsten plugs were formed, the next layer of metal was deposited on the wafers and patterned. At this point, test structures on the wafer provided the ability to electrically test the plugs to indicate how well they would perform in the finished circuit.

Two split lots were processed at different plug layers within the process sequence. For both lots, the results indicated that the new process performed equivalently or better for all electrical test parameters. One lot, when processed through the entire fabrication sequence, showed significantly improved final circuit yield for those wafers that had been processed on the new tools versus those that had been processed on Vendor 1's tool.

5.2.6.1. Repeatability Experiment

As a final test of the Hudson etch process, a repeatability experiment was run. This experiment consisted of processing two blocks of 25 wafers with the intent of exercising the tool over an extended run to assess the consistency of the process characteristics. The blocks included tungsten etch rate wafers, etch particle wafers, machine particle wafers, residue monitor wafers, and dummy etch wafers. The etch rate wafers were processed and measured as in previous

experiments. The etch particle wafers were intended to assess the number of particles added to the surface of the wafer during the etch process. These were blanket tungsten wafers which had been measured for film particles using a laser scattering particle counting tool. After tungsten and TiN had been etched from the wafers, particles were measured again and the difference between pre and post-measurements was calculated. Machine particle wafers were bare silicon wafers that were simply cycled through the tool to measure the cleanliness of the tool's internal environment. These wafers were measured for particles before and after processing in the same manner as the etch particle wafers. Residue monitor wafers were blanket tungsten wafers that were etched to stopping on the TiN layer and analyzed in the SEM tool for residues. Etch dummies were blanket tungsten wafers that were processed through the entire etch process to exercise the tool between test wafers.

Between the two blocks of experimental wafers, 35 product wafers were processed through the entire etch process. This provided additional exercising of the tool between tests.

A detailed description of the repeatability experiment and results are presented in Appendix D. Across the 85 wafers that were processed as part of this experiment, process parameters remained stable. Etch particle and machine particle counts were within acceptable limits.

5.2.7 Release to Production

Following the positive results on the split lots and the repeatability experiment, the tungsten etcher was turned over to production. On the 9th day of month 8, the tool was released to production under process engineering control while the final startup activities were completed – process control procedures, operator training, and rinse module particle control. On the 20th of month 9, the tool was turned over to production.

5.3 DISCUSSION

The tool startup described in this chapter was unusual in several ways. The discussions in this section consider some of the unusual aspects of the startup as well as some more common lessons that were observed during the project.

5.3.1 A Case for Production Optimized Tools

In many ways the selection of the new tungsten tools was a choice for production optimized tools. While the final purchase decision was influenced more by Vendor 3's etcher's technical capability (helium backside cooling and corrosion prevention), members of the tool selection team initially considered alternatives to the existing tool set because of throughput and availability considerations. The existing tools had a long history of poor availability, low throughput and particle generation problems. They also had a high cost of ownership due to the need for a vendor service contract and high consumables cost. The new tools required comparatively low consumables and were simple and reliable enough that Digital could maintained the tools in-house and thereby avoid the cost of a service contract.

The decision to purchase an unproven tool less than one year before capacity targets had to be met was a remarkable one. The challenge of developing a new process includes considerable risk. Had process development efforts been unsuccessful, the capacity of the entire fab might have been limited by the throughput of the tungsten process. Arguably, the dissatisfaction among Digital's process engineers regarding Vendor 1's tools had much to do with this risk taking.

Due to significant differences between the new tungsten tools and Vendor 1's tool, this chapter does not present a comparison between a cluster tool and a non-integrated alternative. However, it is clearly an example of a case in which there was no clear technical advantage to a cluster tool[6]; a non-integrated set of tools proved capable of producing acceptable results.

5.3.2 Shifting Priorities

Shortly after the tool selection team had decided to purchase the tools from Vendor 2 and Vendor 3, Vendor 1 succeeded in making several upgrades to the existing tungsten tools. As a result of these upgrades, the tool was capable of processing tungsten without corrosion and with a dramatic increase in throughput. In fact, with the upgrades, the existing tool set might be capable of meeting the capacity targets without the new tools (capacity studies to confirm this have not been completed).

One result of these changes was that by the time the new tools arrived at Digital, the tool startup project was no longer a priority; tungsten capacity was no longer

a critical issue. Tungsten processing was not the only area of the fab that was starting up new tools; several other processing areas had purchased additional tools to meet the higher capacity targets. The startup of these other tools now had a higher priority than the new tungsten tool startups. This reduced priority status of the tungsten tools had negative effect on the startup efforts.

One example of the effect of low-priority status was reduced access to fab resources. Resolution of the residue problem required considerable time on an SEM tool as this was the only way to detect residues. However, SEM time is a limited fab resource and the reduced priority of the tungsten startup often meant very long wait times between performing a residue experiment and analyzing the results. Similarly, wafers are a limited fab resource, both because of the cost of the wafers and the time on other tools in the fab that test wafer preparation requires. A budget of wafers was initially allocated for the tungsten startup experiments. When resolving the residue and rinse module particle issues required more wafers than were expected, obtaining additional wafers for this relatively low- priority project was difficult. Some of the experiments initially planned to characterize the new tools were never completed due to a lack of both time and wafers.

Priorities shifted once again near the end of the startup period. The importance of starting up the tungsten tools increased after the initial split lots were analyzed. Because the tools offered a potential increase in yield, getting the tools into production became important once again. In fact, process development activities were cut short as soon as a production worthy process was demonstrated.

5.3.3 The Need to Cooperate

The circumstances at the time the tungsten tools were purchased led the two vendors and Digital to a level of cooperation that does not always exist between tool vendors and their customers. The vendors had been working together to introduce their tungsten tools to the industry as a combined tungsten solution. Their marketing efforts would be successful only if they could establish a customer base and demonstrate the performance of their tools. Digital was looking for an alternative to their existing tungsten tools. The new tools could provide that alternative if they could successfully develop a process that met Digital's needs. The situation presented the opportunity for all three companies to win, provided they could cooperate successfully.

The incentives to cooperate convinced Digital to risk purchasing the unproven tools and in the end these same incentives were instrumental in the success of the process development. Typically, vendors take little responsibility for each other's tools. In such typical situations, a problem like etch residues being caused by the deposition tool would lead to "finger pointing" and wasted time. In this case, the fact was that each company would benefit only if the entire process was successful. This fact led to the cooperative efforts necessary to resolve the residue problem and other development challenges.

All three companies have indeed benefited from the cooperative efforts of the process development and tools startup. Vendor 2 and Vendor 3 gained the opportunity to demonstrate to Digital as well as the rest of the industry that their tools offer a viable tungsten process. This success has strengthened their

relationship and will doubtless add to their ability to sell additional tools. Digital has gained an alternative tungsten process capable of meeting requirements with a possibility of improved quality and yield. An additional benefit to Digital is the improved strategic position of having a choice in tungsten processing tools. It is reasonable to speculate that Vendor 1 worked to improve their tools in response to competitive pressure presented by Vendor 2 and Vendor 3.

5.3.4 Developing a Process In Three Different Locations

Process development work for the tungsten tools was complicated by the fact that the work took place in three locations. Wafers used in the development experiments started out in Hudson where tools in Digital's fab were used to prepare the wafers as described in section 5.2. From there the wafers were boxed and shipped to Vendor 2 in California. After receiving tungsten deposition at Vendor 2 the wafers were sent to Vendor 3's laboratory in California for etching. Transporting wafers alone is problematic as it provides an opportunity for contamination, but the situation described above included other problems. The tools being used at Vendor 2's and Vendor 3's facilities were not dedicated to Digital's development work. Experiments and demonstrations for other customers were conducted on the tools. At times, this other work required that the tools be reconfigured to run other processes. The state of the tools at the time Digital's experiments were run varied from experiment to experiment. This additional source of variation limited the development team's ability to analyze and interpret experimental results.

The decision to conduct the process development in this manner was not a bad one; given the circumstances, it was probably the only way to develop a process within the required time constraints. However this example should serve to demonstrate the disadvantages of such an arrangement.

5.3.5 The Difficulty of Transferring a Process

In addition to the challenges mentioned above, developing the process on tools in California meant that the process had to be transferred once the tools became available in Hudson. The complexity and tight tolerances of semiconductor manufacturing processes make them sensitive to the slightest variation in tool configurations. Transferring a process from one semiconductor manufacturing tool to another is not a simple undertaking. In fact, for manufacturers with multiple fabrication facilities, transferring technology has become a key business issue.

Process sensitivity to hardware variation combined with the fact that some aspects of semiconductor processing are not completely understood, often means that process development is based as much on empirical evidence as it is on theoretical application. Tool characterization procedures often include complex response surface modeling (RSM) experiments to identify the subtle differences in the performance of individual tools. Minor adjustments in process recipes are typically required to obtain similar results from apparently identical tools. The differences in the configurations of the Vendor 3 etchers in Hudson and California dictated that much of the process development work be repeated.

Although there are considerable challenges to transferring a process as described above, the process development work done in California was beneficial and reduced the time required to perform the tool startup. This period of preliminary investigation located a region of operation likely to produce acceptable results and identified residues as a critical issue. It also provided an opportunity for the engineers from Vendor 2, Vendor 3 and Digital to develop a sound working relationship which was instrumental in the success of the later work at Hudson.

The experience gained from transferring the tungsten process will likely prove valuable to Digital in the future. Additional Vendor 2 and Vendor 3 tools have been purchased for use in Digital's volume production facility in South Queens Ferry, Scotland. Experience with the effect of variation in tool configuration and residue elimination will doubtless aid in the startup of those tools.

5.3.6 Planning in the Face of Uncertainty

A common theme to the preceding discussions has been the presence of uncertainty. As a result of this uncertainty, startup plans were revised several times. Initial expectations were that the process would be developed before the new tools arrived at Hudson. Level 3 work was expected to consist of a full integrated experiment, split lots, and a "marathon run". The integrated experiment would be a series of experimental runs carefully designed to map process variables to the resulting process characteristics for both etch stages. The results of the experiment would allow process engineers to select an optimum set of operating settings as well as provide the ability of making purposeful

adjustments to the process in the future. Split lots would provide confirmation that an acceptable process had been identified. The marathon run was to be an extended run of the new process to determine whether the results were stable over time. Initial plans for the marathon run included etching 300 wafers.

When the tools arrived in Hudson before a process had been developed, integrated experiment plans were modified to include separate tungsten and TiN etch experiments which were intended to locate a feasible operation region in which to design the full characterization experiment. The plans included following these preliminary experiments with the full integrated experiment, split lots and marathon on an accelerated, and rather aggressive, schedule. The presence of residues made it obvious that this schedule contained a great deal of uncertainty and would likely be revised as the work progressed.

As residue elimination work consumed startup time and resources, the startup team's goals shifted from determining the optimal process to finding a process that would produce acceptable results. By the time the etch experiments had located a feasible region of operation and the overetch experiment had demonstrated acceptable plug recess, very little time remained for refining the process before the split lots were scheduled to be processed.

Because split lots receive the entire fabrication process, they must be scheduled well in advance of tungsten deposition, which occurs fairly late in the process. Split lots are allocated to the various processing areas which require verification of experimental processes. Often, split lots are allocated to receive split

processing at more than one fabrication step. Such lots contain experimental data relating to more than one process and cannot be delayed without affecting other development efforts. Not processing a scheduled split lot would mean incurring the delay of waiting for other split lots to be allocated and processed. When the scheduled split lots were ready for tungsten processing, the startup team decided to process them through the new etch process although it had not been fully optimized.

After the split lot results indicated an increase in yield for the new process, plans changed once again. Increasing yield is an overall goal for the entire fab. With these tools offering a potential increase in yield, pressure was placed on the startup team to integrate the tools into the production process early. Plans to characterize the process fully in an integrated experiment were postponed indefinitely and a truncated marathon experiment (referred to as a repeatability experiment in the preceding section) consisting of 50 test wafers was run. The release of the tools to production control then marked the end of the acquisition and integration process.

5.4 SUMMARY

This chapter presents a record of the events surrounding the acquisition and integration of a new tool into a semiconductor manufacturing line. Although it was in some ways atypical, in a broader sense the startup was characterized by factors that affect all process development work.

Semiconductor manufacturing is extremely technical work. The people who develop semiconductor processes are highly skilled and employ sophisticated scientific techniques. In the final analysis, however, semiconductor manufacturing is a business. The goals of bringing the product to market must take precedence over a desire to obtain rigorous scientific results.

As scientists, the engineers managing the tungsten tool startup carefully planned a set of experiments to develop and understand the process. As semiconductor manufacturers, they adjusted those plans to compensate for uncertainty and to meet the time and resource constraints they faced. Through balancing their responsibilities as scientists and manufacturers they were successful in developing a process and integrating the new tools into the fab on time.

APPENDIX A: BACKGROUND, TUNGSTEN PLUGS

A.1. TUNGSTEN PLUGS

The semiconductor manufacturing tools described in Chapter 5 are used to create structures referred to as tungsten plugs. This appendix provides background on tungsten plugs and the process used to create them.

A.1.1. Purpose

Integrated circuits use thin traces of metal film as the "wiring" to connect individual devices (transistors, capacitors, resistors, etc.). In the more complex circuits, several layers of metal traces, separated by insulating layers of SiO_2, or "oxide", are required to make the connections. At the points in the circuit where it is necessary for the metal layers to be connected, tungsten plugs are used to make the connection.

A.1.2. Process

Figure A.1 illustrates the tungsten plug deposition and etchback process. After the underlying metal layer has been deposited and the circuit traces have been patterned into it, a blanket layer of SiO2 is deposited. Holes, or "vias" are then etched through the oxide down to the metal. These vias are filled with tungsten by chemical vapor deposition (CVD). Because CVD tungsten does not adhere well to oxide[16], the vias are first deposited with 150 Å of titanium followed by 1050 Å of titanium nitride (for simplification, this layer will be referred to as TiN) which acts as an adhesion layer between the oxide and the tungsten. The

tungsten film is deposited to a thickness at least equal to the diameter of the vias to insure that the vias are completely filled. The tungsten film is subsequently etched away from the wafer surface leaving tungsten "plugs" filling the via holes. When the next layer of metal is deposited, the tungsten plugs act as conducting paths providing electrical connections to the underlying metal.

The tungsten etch process is designed such that the etch rate of the TiN adhesion layer is less than that of tungsten, or in other words the etch is selective to tungsten. In this way, the TiN is used as a barrier layer to prevent the etch from attacking the underlying oxide. After the tungsten has been etched away, a second etch, which is selective to TiN, is used to remove the thin TiN layer. Thus the formation of tungsten plugs is a two stage process. First, the tungsten is removed leaving the TiN adhesion layer. Then the TiN is removed exposing the underlying oxide.

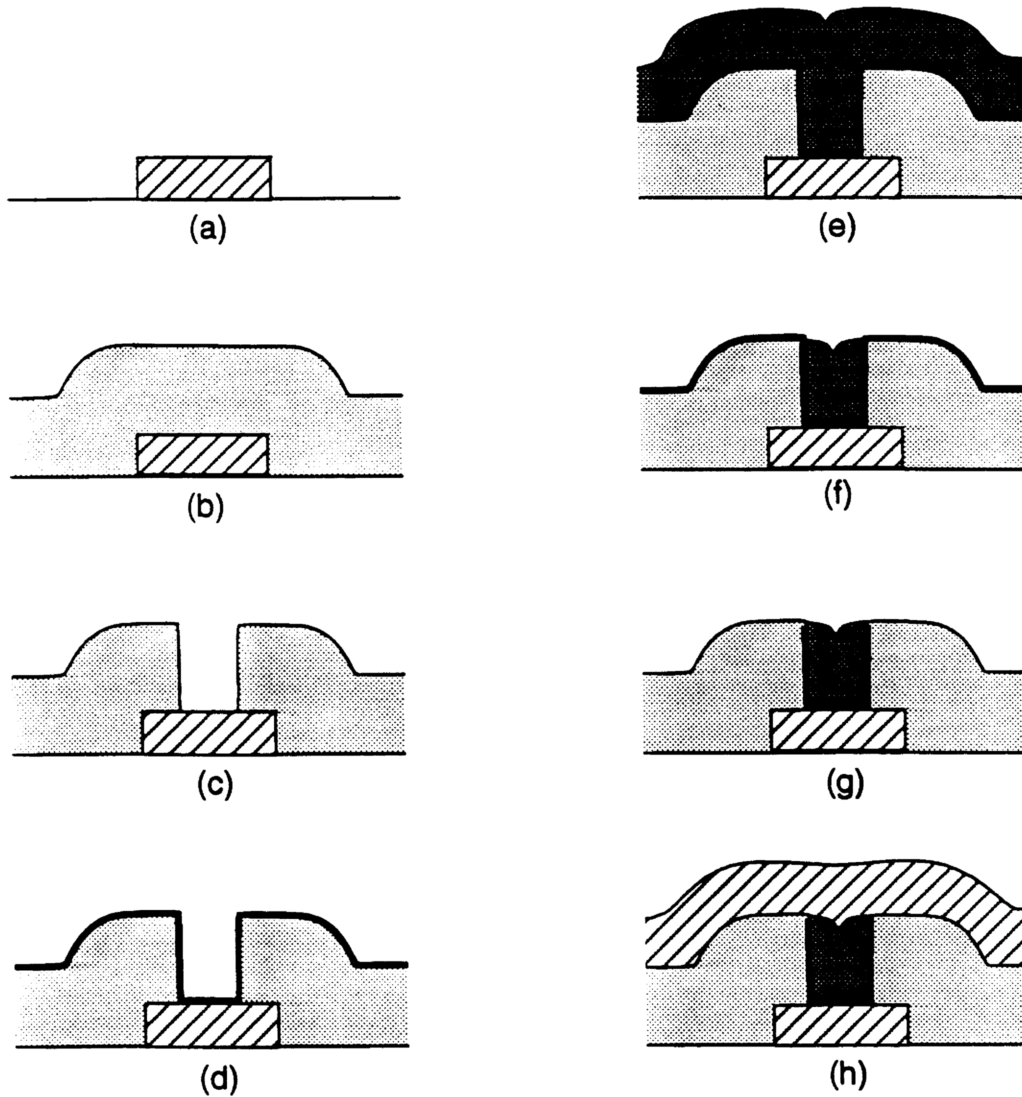


Figure A.1. Tungsten Plug Process Cross Sections.

(a) patterned metal line, (b) oxide deposition, (c) via etch, (d) TiN deposition, (e) tungsten deposition, (f) tungsten etch, (g) TiN etch, (h) second metal deposition.

A.1.3. Critical Plug Characteristics

Critical characteristics of the etched plugs are recess, trenching and coring. Recess refers to the level of the tungsten in the contact relative to the oxide. The metal which is deposited on top of the tungsten does not conform well to steps in the underlying topography. If the plug is recessed below the oxide, the metal deposited on top of the plug may become thin as it crosses the step, see Figure A.2, and as a result the metal will be a poor conductor of electricity.

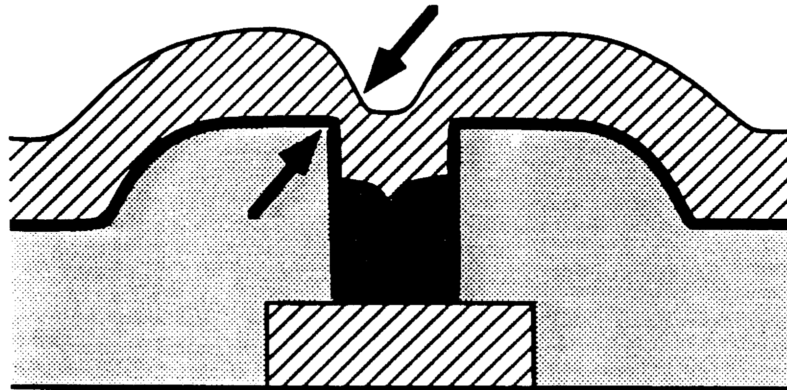


Figure A.2. Thinned Metal Over Recessed Plug.

Two mechanisms cause the finished plug to be recessed below the top of the via. The first recess mechanism is referred to as micro loading. At the end of the etch, as the tungsten is etched off the surface of the wafer and only the plugs remain, the area of exposed tungsten is greatly reduced. This results in a decrease in the rate of etchant consumption and a net increase in the amount of etchant available for etching. The increased level of etchant translates into an increased etch rate. So, as the tungsten clears and only the plugs are left, there is a tendency for the

plugs to etch at a rate faster than the bulk tungsten. Micro loading is expressed as the plug etch rate divided by the bulk tungsten etch rate.

The second mechanism causing recess is the need for an overetch. Plasma etching is anisotropic (directional) in the vertical direction. This means that it etches faster vertically than it does horizontally. As illustrated in Figure A.3(a), in areas where the underlying structures have a sharp step, there is more tungsten to etch near the base of the step (D2) than there is in other regions (D1) . If the etch stops when the bulk of the area has cleared, film depth D1 having been etched, there is a danger of leaving a fillet of tungsten (referred to as a "stringer") along the base of the step, as shown in Figure A.3(b). To insure that all stringers have been etched away, the etch is continued after the bulk has cleared. Unfortunately, this overetch also etches the plugs leading to additional recess, see Figure A.3(c).

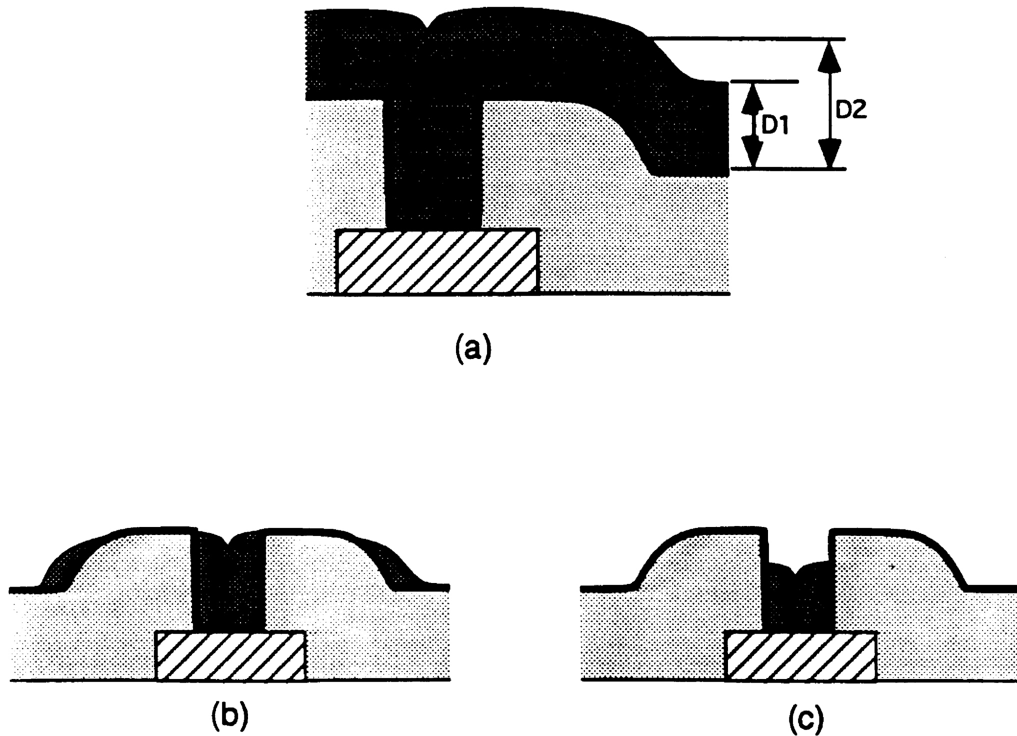


Figure A.3. Cross Sections of Tungsten Plugs.

(a) difference in tungsten thickness over step, (b) plug with stringers, (c) over etched plug showing excessive recess.

Trenching is an effect similar to plug recess except it affects the TiN layer. As with the tungsten etch, micro loading and the need for an overetch result in the TiN adhesion layer being etched away from around the plug (see Figure A.4 below). Excessive trenching can result in a undesirable electrical characteristics.

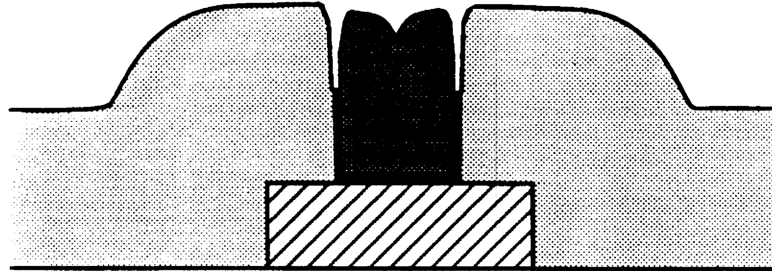


Figure A.4 Cross Section of Tungsten Plug Showing Trenching of TiN Adhesion Layer.

Coring refers to the tendency of the etch to open the seam in the center of the plug left by the tungsten deposition process. During deposition the tungsten builds up on all surfaces relatively evenly. As the tungsten accumulates on the via walls they eventually meet in the center of the via and the plug is formed. As the via closes off, the deposition of tungsten on the via walls slows down and the via may close stopping deposition before the entire plug has been filled. The resulting seam in the center of the plug may be reopened during etchback and the etch will attack the plug from the center out, as illustrated in Figure A.5. In the case of excessive overetch, the entire plug may be etched away leaving an empty via and no conductive path between the metal layers.

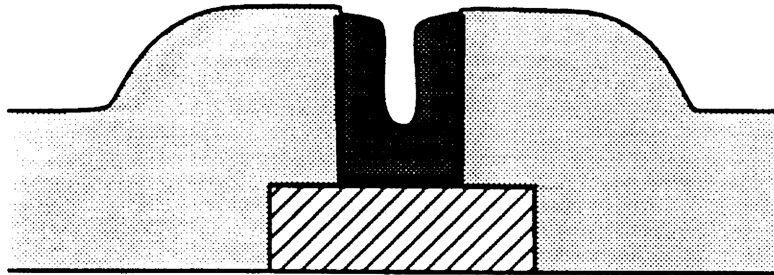


Figure A.5. Cross Section of Tungsten Plug Showing Coring.

A.1.4. Critical Etch Properties

Building on the desirable plug characteristics described above, there are several desirable properties for the tungsten and TiN etch stages. Desirable tungsten etch properties are:

- High tungsten etch rate
- High tungsten etch-rate uniformity
- High tungsten/TiN etch-rate ratio (selectivity)
- Low Micro loading

Because it is necessary to etch a tungsten film thickness equal to that of the via diameter, tungsten etch rate should be high to provide high throughput for the etch process. Etch rate uniformity should be high to insure that the tungsten clears from the wafer surface at approximately the same time in all areas. If the etch is not uniform, the areas that clear first will be excessively overetched, leading to plug recess, while the later areas are being cleared. The selectivity to

tungsten should be high to allow the TiN layer to be used as a barrier during the overetch. As described above, micro loading should be low to limit plug recess.

Desirable properties for the TiN etch stage are:

- High TiN etch rate
- High TiN etch-rate uniformity
- High TiN/Tungsten etch-rate ratio (selectivity)
- High TiN/oxide etch-rate ratio (selectivity)

As with the tungsten etch, the TiN etch rate should be high for throughput reasons. The etch rate uniformity and the selectivity of the etch to TiN over tungsten and oxide should be high enough to prevent excessive etching of the plugs and the surrounding oxide during the TiN etch.

For the entire etchback process, and for virtually all semiconductor fabrication processes, a critical factor is the number of particles added to the wafer surface. With device dimensions of less than one micron (1/millionth of 1 meter), very small particles on the wafer will destroy the product. The fabrication environment and tools are carefully designed to minimize particle contamination, however particles continue to be a major limiting factor to integrated circuit yield. Like all fabrication tools, tungsten etch equipment is monitored regularly for particle generation.

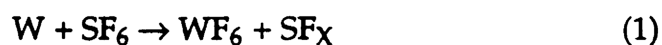
A.2. VENDOR 3'S TUNGSTEN ETCH PROCESS

Vendor 3's tungsten etcher is a single wafer plasma etch tool. It etches wafers one at a time in its reaction chamber using a process known as plasma etching. In this section plasma etching, the basic tool configuration and the tungsten etchback sequence are described.

A.2.1. Plasma etching

When an inert gas is exposed to a high energy electrical field, particles within the gas undergo a series of collisions and become accelerated resulting in high-energy atoms, ions and electrons, also known as a plasma. As some of the particles transition from higher to lower energy states, photons are emitted, causing the gas to glow. Under the proper conditions (pressure, rf energy) the plasma becomes self-sustaining and is referred to as a glow-discharge. A familiar example of such a glow-discharge is neon lighting.

In plasma etching a glow-discharge is used to produce reactive chemical species capable of etching the intended material.[8] In the case of tungsten etching, sulfur hexafluoride (SF_6) and argon (Ar) are used in a glow discharge to produce reactive fluorine. The fluorine then diffuses to the surface of the wafer, adsorbs onto the surface and reacts with the tungsten to form volatile WF_6 according to equation (1)[17]

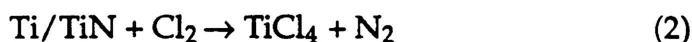


$$W + SF_6 \rightarrow WF_6 + SF_X \qquad (1)$$

The WF_6 then desorbs from the surface and diffuses into the bulk gases.

The etching process is sustained by continually adding Ar and SF_6 to the chamber and pumping out gases at a rate necessary to maintain a constant pressure. In this manner the inert and reactive gases are replenished while the etch by-products, WF_6 and SF_X are removed.

The Ti/TiN adhesion layer is etched in a similar plasma etch process using argon and chlorine gas (Cl_2) according to equation (2)[17]



$$Ti/TiN + Cl_2 \rightarrow TiCl_4 + N_2 \quad (2)$$

A.2.2. Etch Tool Configuration

A block diagram of the tungsten etcher is shown in Figure A.6. Wafers travel through the tool in the following sequence. The robotic handler select wafers from the entry cassette and moves them to the aligner which rotates each wafer to the same orientation for processing. After alignment, the wafers are transferred to the entrance load lock which serves to isolate the etch chamber from the external environment. The load lock is pumped down to process pressure and the wafer is transferred to the etch chamber where plasma etching takes place. After etching, the wafer is transferred to the exit load lock. The exit load lock is vented to atmospheric pressure and the wafer is transferred to the rinse module which is a de-ionized water rinse station (see next section). After rinsing, the wafer is transferred to the exit cassette and processing is complete.

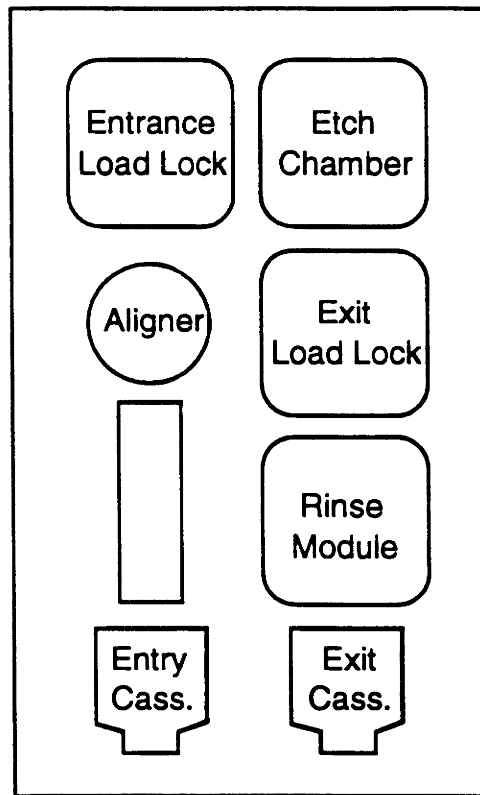


Figure A.6. Block Diagram of the Tungsten Etcher (plan view).

A.2.3. Etch Chamber Configuration

The basic configuration of the tungsten etch chamber is shown in Figure A.7. The upper and lower electrodes are used to generate the rf field in the chamber. Process gases enter the chamber through perforations in the upper electrode. During etching, the wafer rests on the lower electrode and is held in place by the clamp. Gases are pumped out of the chamber through the pumping manifold and the throttle valve at the bottom of the manifold provides pressure control. An optical emission detector is used to measure the level of certain gases in the chamber for etch endpoint detection. Temperature within the chamber can be controlled in several ways. Both the upper electrode and the lower electrode are

water cooled/heated and can be controlled separately. In addition, the walls of the chamber are resistively heated via a separate control circuit.

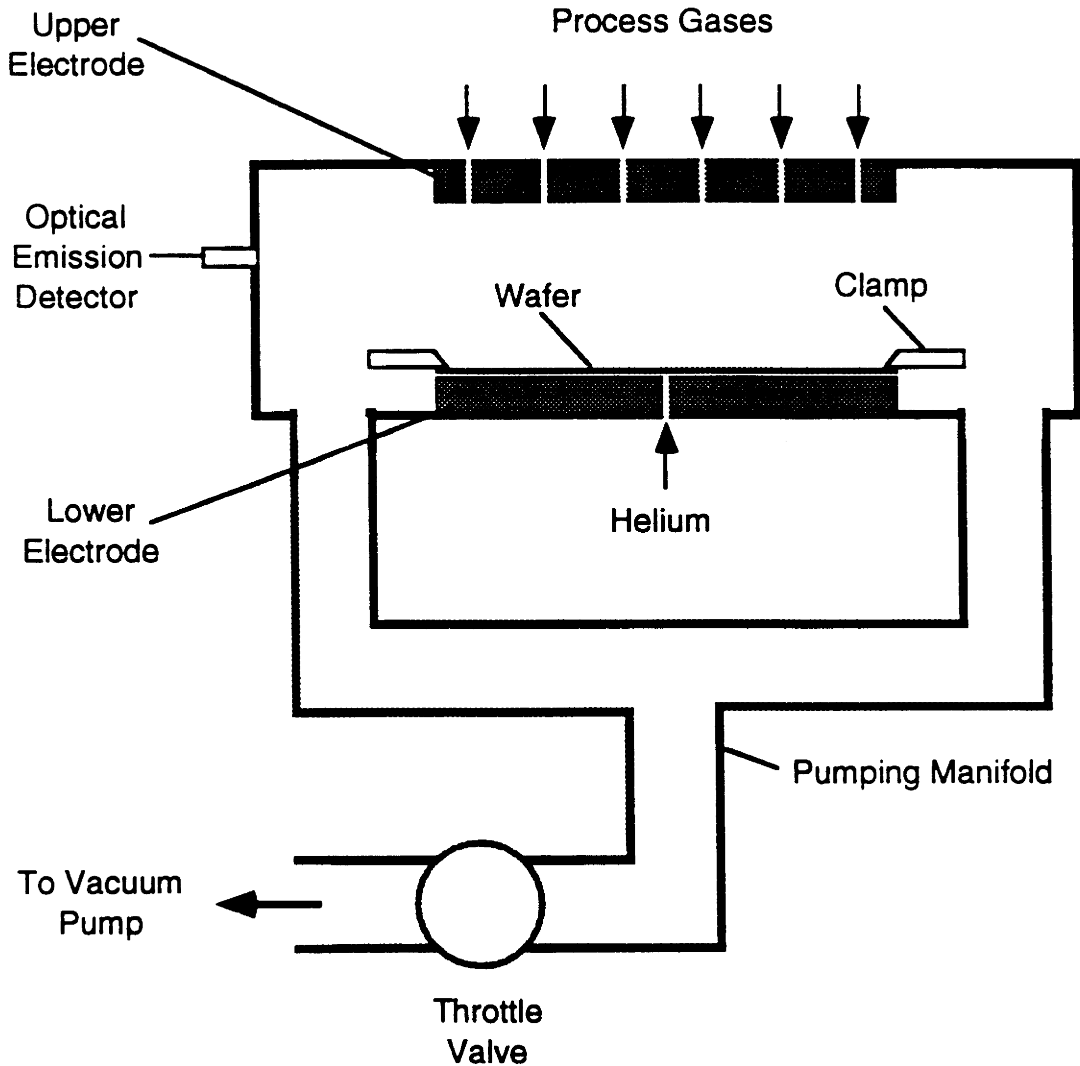


Figure A.7. Tungsten Etch Chamber Cross Section.

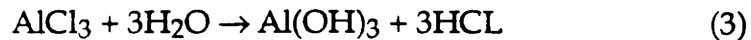
A.2.4. Temperature Effects and Helium Backside Cooling

The high-energy plasma tends to heat the wafer during the etch process. Without cooling the wafer can reach temperatures in excess of 200 °C. If the wafer is held at a lower temperature during the etch, the micro loading effect can be reduced and the selectivity of the etches can be increased[18]. At temperatures below 45° C an inverse loading effect can be achieved whereby the plug etches slower than the bulk film. This is a very desirable etch characteristic as it allows for removal of stringers without the danger of excessive recess.

As mentioned above, the lower electrode upon which the wafer rests is cooled, however the low pressure environment within the etch chamber does not provide good thermal conduction between the wafer and the electrode. The tungsten etcher provides wafer cooling by flowing helium gas across the back side of the wafer. The helium gas provides a thermal conductive medium between the wafer and the lower electrode. Varying the backside helium pressure provides additional control of cooling.

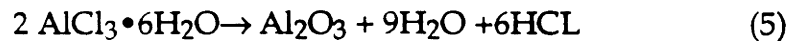
As mentioned above, the longer overetches provided by helium backside cooling allow for better stringer removal through extended overetches, however, an extended TiN etch can result in aluminum bond pad corrosion. Integrated circuits are connected to external electronics by wires attached to (relatively) large areas of aluminum on the wafer surface, referred to as bond pads. After tungsten and TiN etchback the bond pads are exposed. Upon prolonged exposure of bond pads to the TiN etch chemistry, chlorine can become embedded in the bond pad. When the bond pad is subsequently exposed to atmosphere the

chlorine and aluminum are believed to react with moisture in the following series of reactions[19]:



$$AlCl_3 + 3H_2O \rightarrow Al(OH)_3 + 3HCL \quad (3)$$

$$Al(OH)_3 + 3HCL + 3H_2O \rightarrow AlCl_3 \bullet 6H_2O \quad (4)$$



$$2\ AlCl_3 \bullet 6H_2O \rightarrow Al_2O_3 + 9H_2O + 6HCL \quad (5)$$

After several minutes of exposure to atmosphere, the result of this reaction is coils of reacted material extending out from the surface of the bond pad. The reacted material breaks off and becomes a source of particulates which can destroy the circuit.

The etcher provides corrosion prevention via the rinse module. Immediately after the wafer leaves the exit load lock it is placed on a pedestal in the rinse module. The pedestal spins and de ionized water is sprayed onto the surface of the wafer. The water rinses away the chlorine before the reaction can take place, thus preventing corrosion. After rinsing, the wafer, still spinning, is dried with a jet of nitrogen gas.

A.2.5. Etch Sequence

Tungsten etching is performed according to the following sequence. The wafer moves from the entry load lock into the etch chamber through a slit valve which then closes sealing the chamber. After the wafer is placed on the lower electrode,

the clamp lowers to hold the wafer in place. Backside helium and process gas (Ar, SF_6) flows are initiated and the throttle valve at the base of the pumping manifold adjusts to maintain process pressure. RF power to the top electrode is turned on and the plasma ignites initiating the tungsten etch process. The fluorine level in the chamber is monitored via the optical emission detector to determine etch endpoint. As the tungsten clears, the fluorine level in the chamber rises until endpoint, after which it remains stable. After endpoint is detected, the etch continues for the selected overetch time. A typical endpoint signal trace is shown in Figure A.6 below.

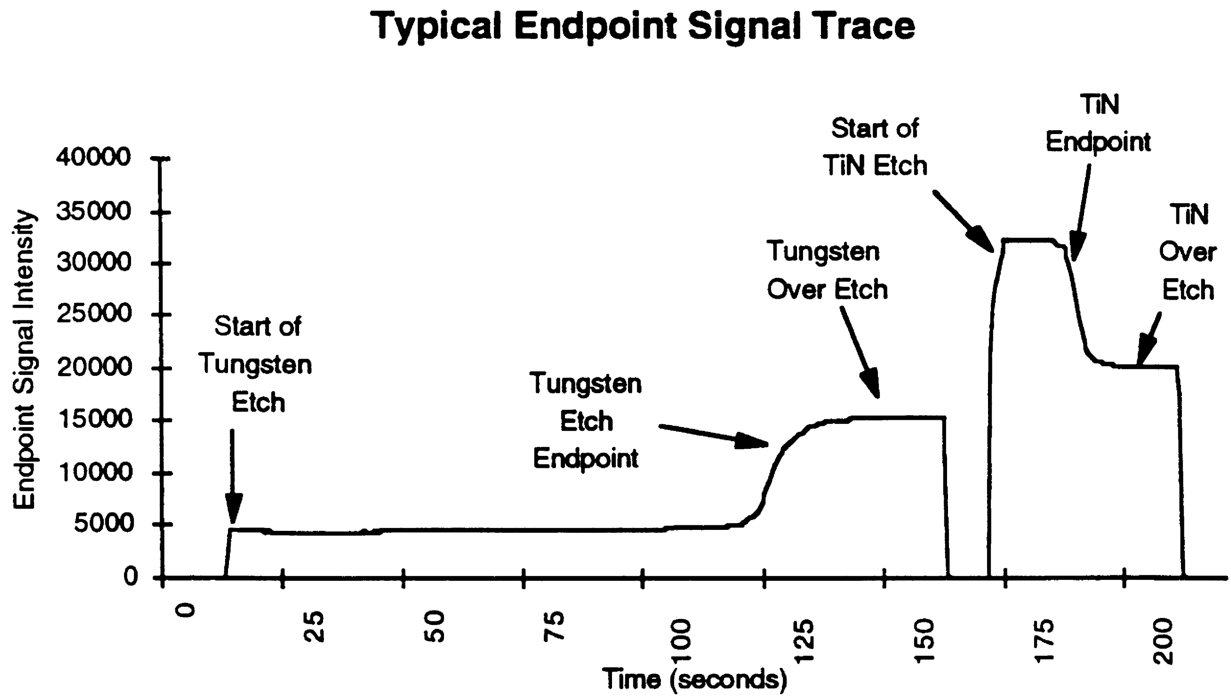


Figure A.8. Typical Endpoint Signal Trace for Tungsten Etchback Process

At the end of the tungsten etch the power is turned off extinguishing the plasma. The TiN etch gases (Ar, Cl_2) are flowed and the pressure is allowed to stabilize. The upper electrode is powered once again and the plasma ignites starting the TiN etch. Again the optical emission detector is used to detect endpoint, however, in the case of etching TiN, the detector measures the level of N_2 in the chamber. As the film is etched N2 is release into the plasma (see equation 2) and when the etch is complete the level of N2 decreases.

The variable process parameters for the tungsten etcher are:
Pressure [millitorr]
Power [watts]
SF_6 flow rate (tungsten etch only) [standard cubic centimeters per minute, sccm]
Cl_2 flow rate (Ti/TiN etch only) [sccm]
Ar flow rate [sccm]
Backside helium pressure [torr]
Upper electrode temperature [°C]
Lower electrode temperature [°C]
Chamber wall temperature [°C]
Overetch time [seconds]

The parameters listed above are controlled separately for the tungsten and TiN etch stages. In the following discussions etch recipes will typically be expressed as numerical values in the following format:

pressure/power/SF6(or Cl2)/Ar/helium

with temperatures and overetch times being listed separately.

APPENDIX B: TIN ETCH EXPERIMENT

Table B.1. Coded Experimental Design Matrix for TiN Etch Experiment

	Coded Factor Levels				
Trial	Pressure	Power	Total Flow	CL:AR Ratio	Helium
1	-1	1	-1	-1	1
2	1	1	-1	1	-1
3	-1	-1	-1	1	1
4	1	-1	-1	-1	-1
5	1	1	1	1	1
6	-1	1	1	-1	-1
7	1	-1	1	-1	1
8	0	0	0	0	0
9	-1	-1	1	1	-1
10	0	0	0	0	0

Table B.2. Actual Factor Levels Corresponding to Experimental Design Codes

	Actual Factor Levels				
Coded Level	Pressure (mtorr)	Power (Watt)	Total Flow (sccm)	CL:AR Ratio	Helium (torr)
1	120	550	50	1	12
0	105	450	40	.75	8
-1	90	350	30	.5	4

Etch Time was time from the beginning of the etch to endpoint detection. Slope was the change in endpoint intensity (counts is an arbitrary unit) per second of etch time during the endpoint portion of the etch. (see Figure A.6)

Table B.3. Results of TiN Etch Experiment

	Response	
Trial	Etch Time (sec)	Slope (count/sec)
1	27	3901
2	13	3015
3	19	434
4	23	501
5	10	1138
6	17	3039
7	10	1015
8	16	1446
9	13	1721
10	17	1389

Table B.4. Factor Analysis of TiN Etch Experiment Responses.

a. Response: Etch Time

Term	Least Squares Coefficient	Standard Error	T-Value	Signif.	Transformed Term (~ indicates trans. factor)
1	16.500	0.707107			
~PreSsure	-2.500	0.790569	-3.16	0.0195	((PS-1.05e+2)/1.5e+1)
~Total Flow	-4.000	0.790569	-5.06	0.0023	((TF-4e+1)/1e+1)
~Cl:Ar Ratio	-2.750	0.790569	-3.48	0.0132	((R-7.5e-1)/2.5e-1)

No. cases = 10 R-sq. = 0.8883 RMS Error = 2.236
Residual df = 6 R-sq-ad = 0.8324 Cond. No. = 1

b. Response: Slope

Term	Least Squares Coefficient	Standard Error	T-Value	Signif.	Transformed Term (~ indicates trans. factor)
1	1759.90	260.576534			
~POwer	927.75	291.333422	3.18	0.0129	((PO-4.5e+2)/1e+2)

No. cases = 10 R-sq. = 0.5590 RMS Error = 824
Residual df = 8 R-sq-ad = 0.5039 Cond. No. = 1

APPENDIX C: TUNGSTEN ETCH EXPERIMENT

Table C.1. Coded Experimental Design Matrix for Tungsten Etch Experiment

	Coded Factor Levels				
Trial	Pressure	Power	Total Flow	SF_6:AR Ratio	Helium
1	-1	1	1	1	-1
2	-1	-1	1	1	1
3	1	-1	-1	1	1
4	-1	1	-1	1	1
5	0	0	0	0	0
6	-1	1	1	-1	1
7	1	-1	1	-1	1
8	1	1	1	-1	-1
9	-1	-1	-1	1	-1
10	1	-1	-1	-1	-1
11	1	1	-1	1	-1
12	-1	1	-1	-1	-1
13	1	-1	1	1	-1
14	-1	-1	-1	-1	1
15	1	1	1	1	1
16	-1	-1	1	-1	-1
17	1	1	-1	-1	1
18	0	0	0	0	0

Table C.2. Actual Factor Levels Corresponding to Experimental Design Codes

	Actual Factor Levels				
Coded Level	Pressure (mtorr)	Power (Watt)	Total Flow (sccm)	SF_6:AR Ratio	Helium (torr)
1	320	370	140	8	18
0	270	320	115	7	16
-1	220	270	90	6	14

Experiment Description and Results

The wafers were deposited with 1.5 µm of tungsten and etched for 100 seconds during each trial. Tungsten film thickness was measured with a 4 point probe at 46 sites on each wafer, before and after etching. The post-etch thickness was subtracted from the pre-etch thickness to determine the amount of tungsten etched. Etch Rate was calculated as the amount of tungsten etched divided by etch time (time from the beginning of the etch to endpoint detection). Etch Uniformity was the percent standard deviation from the mean of the 46 etch measurements.

Table C.3. Results of Tungsten Etch Experiment

Trial	Response	
	Etch Rate (Å/min)	Uniformity (%, 1σ)
1	7690	8.72
2	6071	5.86
3	4855	3.69
4	6713	6.66
5	6412	2.78
6	7482	8.62
7	5612	2.68
8	7210	3.07
9	5492	3.65
10	4873	3.34
11	6253	2.30
12	6611	7.79
13	5652	2.81
14	5257	3.67
15	7254	2.15
16	6109	5.60
17	6148	2.52
18	6411	2.71

Table C.4. Factor Analysis of Tungsten Etch Experiment Responses.

a. Response: Etch Rate

Term	Least Squares Coefficient	Standard Error	T-Value	Signif.	Transformed Term (~ indicates trans. factor)
1	6227.316113	25.227508			
~PreSsure	-222.379714	26.757813	-8.31	0.0001	((PS-2.7e+2)/5e+1)
~POwer	715.660898	26.757813			((PO - 3.2e+2)/5e+1)
~Total Flow	430.431618	26.757813			((TF-1.15e+2)/2.5e+1)
~PO*~TF	58.601053	26.757813	2.19	0.0474	

No. cases = 18 R-sq. = 0.9877 RMS Error = 107

Residual df = 13 R-sq-ad = 0.9840 Cond. No. = 1

b. Response: Uniformity

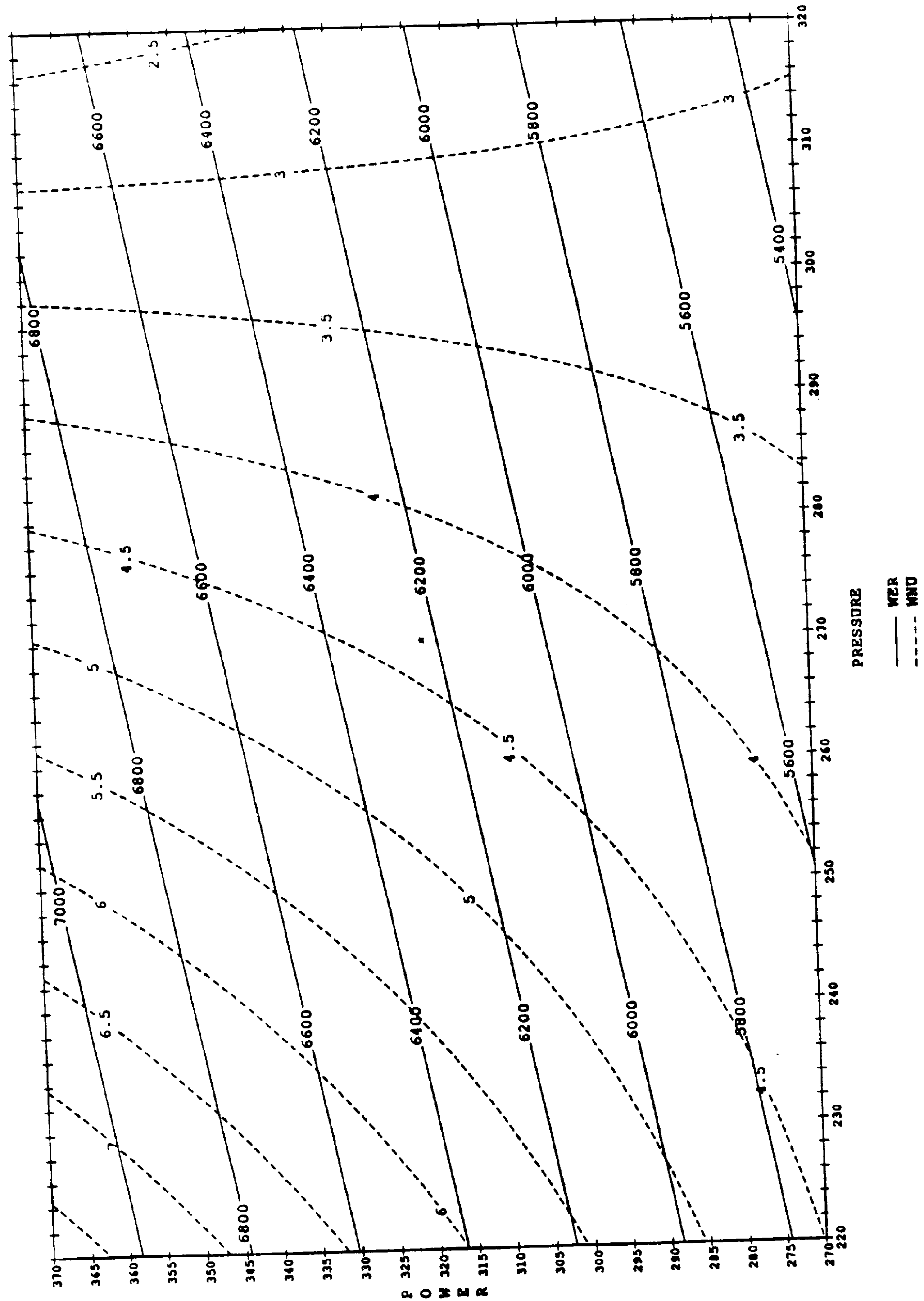
Term	Least Squares Coefficient	Standard Error	T-Value	Signif.	Transformed Term (~ indicates trans. factor)
1	4.367778	0.189882			
~PreSsure	-1.750625	0.201400			((PS-2.7e+2)/5e+1)
~POwer	0.658125	0.201400			((PO - 3.2e+2)/5e+1)
~Total Flow	0.368125	0.201400			((TF-1.15e+2)/2.5e+1)
~PS*~PO	-0.968125	0.201400	-4.81	0.0004	
~PS*~TF	-0.51[illegible]625	0.201400	-2.54	0.0262	

No. cases = 18 R-sq. = 0.9085 RMS Error = 0.8056

Residual df = 12 R-sq-ad = 0.8703 Cond. No. = 1

Figure C.1. Contour Plot of Tungsten Etch Rate (WER) and Uniformity(WNU) as Functions of Power and Pressure

Total Flow = 115 sccm, SF6:Ar Ratio = 7, Helium = 16 torr



APPENDIX D: REPEATABILITY EXPERIMENT

The repeatability experiment consisted of two blocks of 25 experimental wafers. The experimental blocks were designed as follows:

Test	Wafer Type	Quantity
Etch Rate (45 sec. tungsten etch)	1μm blanket tungsten	1
Etch Particles (full tungsten & TiN etch)	1μm blanket tungsten	1
Etch Dummies (full tungsten & TiN etch)	1μm blanket tungsten	9
Machine Particles (gases on, no rf power)	bare silicon	1
Etch Rate (45 sec. tungsten etch)	1μm blanket tungsten	1
Etch Particles (full tungsten & TiN etch)	1μm blanket tungsten	1
Etch Dummies (full tungsten & TiN etch)	1μm blanket tungsten	9
Machine Particles (gases on, no rf power)	bare silicon	1
Residue Monitor (tungsten etch only)	1μm blanket tungsten	1
	Total	25

35 product wafers were processed through the complete etch process after the first repeatability block was processed. Thus the repeatability test collected data across the processing of 85 wafers in all.

In addition to the etch rate, etch particle, machine particle tests, and residue monitors, data was collected regarding etch time and chamber wall temperature (the chamber walls are resistively heated and are subject to drift at the low temperature settings used). The results of the various tests are given below.

Etch Rate

Table D.1. Etch Rate and Uniformity Test Results

Experiment Block	Wafer Number	Etch Rate (Å/min)	Uniformity (%, 1 σ)
1	1	6111	3.48
1	13	6215	3.90
2	1	6188	3.50
2	13	6205	3.96

Particles

Table D.2. Etch Particle Test Results

Experiment Block	Wafer Number	Pre-Dep Particles	Post-Dep Particles	Post-Etch Particles	Particles Added
1	2	13	10	80	67
1	14	14	6	28	14
2	2	7	7	184	177*
2	14	32	20	69	37

*Wafer number 2 in block 2 appeared to have several scratches on it after etching. There was no known mis-handling of this wafer. Scratches were not apparent at pre-deposition or post deposition measurements.

Table D.3. Machine Particle Test Results

Experiment Block	Wafer Number	Pre-Etch Particles	Post-Etch Particles	Particles Added
1	12	2	5	3
1	24	12	15	3
2	12	0	10	10
2	24	1	12	11

Etch Times

Etch times were recorded for the tungsten and TiN etch processes. Etch times were observed to be very consistent over these observations. Summary statistics of the observations are given below.

Table D.4. Repeatability Experiment Etch Times (seconds)

	Tungsten Etch	TiN Etch
Number of observations	44	17
Minimum etch time	129	17
Maximum etch time	132	20
Mean etch time	130.5	17.6
Standard deviation	0.79	0.87

Residues

No residues were observed on the residue monitor wafers.

Chamber Wall Temperature

The chamber wall temperature drifted up during the experiment from 33.1 °C to 34.2 °C. The temperature appears to stabilize near 34 °C.

References

[1] M. Splinter, VP and General Manager of Manufacturing, Intel Corporation, in an address to MIT Leaders for Manufacturing Fellows, Cambridge Massachusetts, July, 1993.

[2] Krishna C Saraswat and Samuel C. Wood, "Adaptable Manufacturing Systems for Microelectronics Manufacturing: Economic and Performance Issues", Presented at "Strategies for Innovation and Changes in the US and Japan" , an IBEAR Research Conf., Univ. of S. Calif., Los Angeles, May, 1992.

[3] Chisato Mori, President, JATCO, from an address to MIT Leaders for Manufacturing Fellows, Cambridge Massachusetts,September 1993.

[4] Gary Castleman, as quoted by Pete Singer, "The Thinking Behind Today's Cluster Tools", Semiconductor International, August 1993.

[5] Pieter Berggraaf, "MESC: Is it Working", Semiconductor International, October, 1991

[6] Pieter Burggraaf, "The Thinking Behind Today's Cluster Tools" Semiconductor International, August 1993.

[7] Pieter Burggraaf, "Integrated Processing: The 1990s Trend", Semiconductor International, June 1989.

[8] S. Wolf and R. N. Tauber. Silicon Processing for the VLSI Era, Volume 1, Lattice Press, Sunset Beach, California 1986.

[9] Randall T. Nickel, "Improving Productivity of Semiconductor Manufacturing Cluster Tools", Master's Thesis, Massachusetts Institute of Technology, Cambridge, MA, June 1994.

[10] Albert S. Bergendahl, David V. Horak, Paul E Bakeman and David J. Miller, "Cluster Tools Part 2: 16 Mb DRAM Processing, Semiconductor International, September, 1990.

[11] M. DeMartino, Manufacturing Manager, Digital Equipment Corporation, from a conversation in December 1993.

[12] Stanley B. Gershwin, Manufacturing Systems Engineering, PTR Prentice Hall, Englewood Cliffs, New Jersey, 1994.

[13] T. Keith McNab, Ph.D., "Cluster tools, Part 3: Technical Issues", Semiconductor International, October 1990.

[14] Richard Kitson, Industrial Engineer, Digital Equipment Corporation, Hudson. MA, from an interview conducted in April, 1994

[15]U.C. Sridharan, Dennis Hartman, Rex Wright, James Carter and Marty kent, "Mechanism of Residue Formation in Low Temperature Tungsten Interconnect Etch", Proceedings of the Ninth Symposium on Plasma Processing, Volume 92-18, The Electochemical Society, Pennington, NJ 1992.

[16]John E. J. Schmitz, "Chemical Vapor Deposition of Tungsten and Tungsten Silicides", Noyes Publications, Park Ridge, New Jersey, 1992.

[17]Marion Garver, Joseph Bulger, Thomas Clark, Jamshed Dubash, Lorain M Ross, and Daniel J. Welch, "CMOS-4 Backend Process Development for a VLSI 0.75 μm Triple-level Interconnection Technology", Digital Technical Journal, Vol. 4 No. 2, Digital Equipment Corporation, Maynard, MA, Spring 1992.

[18]Ivo Miller, Barbara Haselden, Robert Frazier, May Su, "Tungsten Etch Technology for Submicron Devices", Lam Research Corporation, Fremont, California, 1991.

[19]G. Cameron and A Chambers, "Successfully Addressing Post-Etch Corrosion", Semiconductor International, May 1989.