

Electromigration of Damascene Copper for IC Interconnect

William Kevin Meyer
B.E.E., University of Minnesota
M.S.Mat.Sci., Oregon Graduate Institute

A dissertation presented to the faculty of the
OGI School of Science & Engineering
at Oregon Health & Science University
in partial fulfillment of the
requirements for the degree
Doctor of Philosophy
in
Electrical and Computer Engineering

April, 2004

The dissertation "Electromigration of Damascene Copper for IC Interconnect" has been examined and approved by the following Examination Committee:

Raj Solanki

Professor

Thesis Research Advisor

David Evans

Adjunct Professor

Thesis Research Advisor

Anthony E. Bell

Associate Professor

Edward F. Ritz, Jr.

Adjunct Professor

5/1/2

Kudos are due to Raj Solanki and Dave Evans for insight, direction, stimulating discussions, and patience during this effort. I gratefully acknowledge wafer processing and financial support by Sharp Laboratories of America. I appreciate the many useful discussions with colleagues, including Hidayat Kisdarjono, Jinshan Hui, and Graham Tewkesbury. Previous employment at Intel Corporation gave a rich experience to draw from, where I gained much from mentors like Jose Maiz and Dwight Crook.

Dedicated with love to my family. Thank you for this time apart.

Table of Contents

Acknowledgement and Dedication	iii
Table of Contents	iv
List of Tables.....	vii
List of Figures	viii
List of Acronyms.....	x
List of Symbols	xi
Abstract	xiii
1. Introduction	1
1.1. Future Trends in Integrated Circuits	1
1.2. Copper and Supporting Roles	2
1.3. Electromigration.....	4
1.4. Process Monitor and Control.....	5
1.5. Dissertation Overview	6
2. Electromigration Theory	7
2.1. Vacancy Dynamics.....	7
2.1.1. <i>Equilibrium vacancy concentration</i>	7
2.1.2. <i>Vacancy Motion</i>	10
2.1.3. <i>Electromigration Wind Force</i>	12
2.1.4. <i>Migration Mechanisms</i>	13
2.1.5. <i>Diffusion Measurements</i>	15
2.2. Stress Evolution.....	16
2.2.1. <i>Blech Condition</i>	16
2.2.2. <i>Time dependence</i>	19
2.2.3. <i>AC EM</i>	20
2.2.4. <i>Instability</i>	21
2.2.5. <i>Concentration Waves</i>	22
2.2.6. <i>Dissipative Structures</i>	23
2.3. Failure.....	24
2.3.1. <i>Time-to-Fail and Black's equation</i>	24

2.4.	Summary	25
3.	Sample Preparation	26
3.1.	Test Structure Design	26
3.2.	Wafer Processing.....	28
3.3.	Copper Deposition.....	29
3.3.1.	<i>Electrolyte mixture</i>	29
3.3.2.	<i>Maintenance of electrolyte</i>	30
3.3.3.	<i>Equipment automation</i>	32
3.3.4.	<i>Current ramp</i>	34
3.3.5.	<i>Thickness and Resistivity</i>	34
3.4.	Summary	36
4.	Self-heat Characterization	37
4.1.	Heat Flow Model.....	37
4.1.1.	<i>Metal-Dielectric Stack</i>	38
4.1.2.	<i>2-D Isolated Infinite Line</i>	39
4.1.3.	<i>End Straps</i>	40
4.2.	Joule Heating.....	43
4.2.1.	<i>Via Resistance (Schottkey Barrier)</i>	45
4.2.2.	<i>Thermal Runaway</i>	46
4.3.	Regression	48
4.3.1.	<i>Measurements</i>	49
4.3.2.	<i>Parameters</i>	50
4.4.	Summary	52
5.	Electromigration Stress Measurements.....	53
5.1.	Comparison of Package and Wafer-level Methods.....	53
5.2.	Wafer Probe Stress System	53
5.3.	EM Stress Conditions.....	54
5.3.1.	<i>Stress Envelope</i>	58
5.4.	Time-dependent Resistance Increase	59
5.5.	Extraction of Diffusion Coefficient.....	62
5.6.	Time to Failure Distributions	63

5.7.	Threshold Estimate.....	64
5.8.	Multiple Activation Energy Model	67
5.9.	Summary	70
Chapter 6 VLSI Design Rules.....		71
6.1.	Supply Line Duty Factor.....	71
6.2.	Design limits	73
6.3.	Consequences.....	76
6.4.	Supply Tree vs Grid.	78
6.5.	Summary	80
Chapter 7 Conclusion.....		81
7.1.	Findings.....	81
7.2.	Application to Process Monitors.....	83
7.3.	Future Work	83
References		85
Biography.....		88

List of Tables

Table 1.1: Properties of Copper and Aluminum	3
Table 1.2 Expected Interconnect Technology Requirements (source: ISTR).....	4
Table 2.1: Activation Energy Measured by Radioactive Tracers	16
Table 3.1: Structure Dimensions	27
Table 3.2: Damascene Process Flow	28
Table 4.1: Typical Thermal Conductivities (in W/m/°K)	38
Table 4.2: Temperature Coefficient of Resistance for Representative Metals	45
Table 5.1: List of Acceleration Models.....	69

List of Figures

Figure 1.1: EM Failure due to voiding (left) and extrusions (right)	4
Figure 2.1: 2-D vacancy formation	8
Figure 2.2 Concentration of Defects	10
Figure 2.3 EM Characteristic Length (for $Z^* = -5$)	13
Figure 2.4: Mass Diffusion Coefficient Measured by Radioactive Tracers.....	15
Figure 2.5 Stress vs position in interconnect segment	18
Figure 2.6 Offset from Anode of Maximum Compressive Stress	19
Figure 2.7: Propagation of Vacancy Concentration Wave.....	23
Figure 3.1: Sketch of Test Structure Layout and Cross-section.....	26
Figure 3.2: View of Finished EM Test Structures	27
Figure 3.3: Potentiostatic Measurement of Metal 1 Plating Bath (a thru d are 4 consecutive passes)	30
Figure 3.4: Potentiostatic Measurement of Metal 2 Plating Bath	31
Figure 3.5: Potentiostatic Measurement after 151 days	32
Figure 3.6: Electoplating Equipment	33
Figure 3.7: Plating Current Steps	34
Figure 3.8: SEM Cross Sections of Patterned Lines	36
Figure 4.1: 2-D Fringing of Heat Flow	39
Figure 4.2: Via Fraction Parameter	42
Figure 4.3: Average Via Fraction Parameter	43
Figure 4.4: Resistivity Vs Temperature	45
Figure 4.5: Temperature Rise and Thermal Runaway	47
Figure 4.6: Temperature Dependence of Via Resistance	50
Figure 4.7: Measured Temperature Profile	51
Figure 5.1: Wafer Stress System Setup.....	54
Figure 5.2: EM Iso-acceleration Curves for T_m	55
Figure 5.3: EM Iso-acceleration Curves for T_j	56

Figure 5.4: EM Iso-acceleration Curves Above Threshold.....	57
Figure 5.5: EM Iso-acceleration for Various Length Structures.....	58
Figure 5.6: EM Stress Conditions and Stress Limits	59
Figure 5.7: Initial Stress Voltage at Start of EM Stress	60
Figure 5.8: Resistance vs Time Near Threshold	60
Figure 5.9: Resistance vs Time Near Runaway	61
Figure 5.10: Voltage Increase Before Catastrophic Open.....	62
Figure 5.11: Probability Plot of Fail Times.....	64
Figure 5.12: Variation in EM Thresholds	66
Figure 5.13: Statistical Distribution of EM Thresholds	67
Figure 5.14: Scatterplot of Acceleration Models	69
Figure 6.1: Simple Inverter Chain Circuit.....	72
Figure 6.2: Temperature Rise.....	74
Figure 6.3: I*R Drop	74
Figure 6.4: Comparison of Capacitance Limits	75
Figure 6.5: Temperature/Voltage Boundary	76
Figure 6.6: Long/Short EM Boundary	77
Figure 6.7: Map of Design Limits. (Hashed area is near-threshold region.).....	78
Figure 6.8: Supply Mesh (arrows show direction of current flow).....	79
Figure 7.1: Linear Plot of Inverse MTF (Threshold Dependence)	82
Figure 7.2: Log Plot of MTF (Black Dependence)	82

List of Acronyms

CISC	Complete Instruction Set Computer
CMOS	Complimentary Metal-Oxide-Semiconductor
CMP	chemical-mechanical polishing
ECD	electro-chemical deposition (electroplating)
EM	Electromigration
fcc	face-centered-cubic
GPIB	general-purpose interface bus
IC	integrated circuit
ILD	inter-layer dielectric
ISTR	International Semiconductor Technology Roadmap
LN2	liquid nitrogen (N ₂)
M1	metal 1 (first metal level)
M2	metal 2 (second metal level)
MTF	median time-to-failure
RISC	Reduced Instruction Set Computer
SH	Joule self-heating

List of Symbols

\bar{q}	vector heat flux
σ	hydrostatic tensile stress
Λ	atomic volume
ν	atomic vibration frequency
ρ	electrical resistivity
∇	three-dimensional del operator
α	thermal diffusivity
δ	thermal fringing correction factor
Ω or ω	'via fraction' parameter
η or N	'average via fraction' parameter
β, β_0, β_1	temporary variables to solve thermal differential equations
δ_{EM}	characteristic EM length
θ_{mj}	thermal resistance
a	atomic spacing
C	capacitance
C	vacancy concentration
c_p	specific heat ($d \cdot c_p$ is volumetric heat capacity)
D	diffusion coefficient
d	mass density
E	energy
E	electric field
F	force
G	vacancy generation (creation) rate
H	work done on individual vacancy
h	metal thickness
J	vacancy flux
J_e	electric current density

J_m	mass flux
k	thermal conductivity
k_B	Boltzmann's constant
L	structure segment length
n	quantity of vacancies
N	quantity of atoms ($n+N$ lattice sites)
P	pressure
P	power
q	electron charge
Q	heat flow through the via
Q_{via}	<i>electrical energy dissipated at the via</i>
R	resistance
r	vacancy jump rate
R	<i>vacancy recombination (extinguishment) rate</i>
R	electrical resistance
S	entropy
T	temperature
t	oxide thickness
$t_{.1}$	time-to-0.1% fail
t_{50}	time to 50% fail
T_j	junction temperature (silicon substrate)
U	work done on the system
V	volume
v_v	vacancy velocity
Z^*	effective valence

Abstract

Electromigration of Damascene Copper for IC Interconnect

William Kevin Meyer, B.E.E., M.S.

Ph.D., OGI School of Science & Engineering
at Oregon Health & Science University

April, 2004

Thesis Advisor: Dr. Raj Solanki

Copper metallization patterned with multi-level damascene process is prone to electromigration failure, which affects the reliability and performance of IC interconnect. In typical products, interconnect that is not already constrained by I-R drop or Joule selfheating operates at 'near threshold' conditions. Measurement of electromigration damage near threshold is very difficult due to slow degradation requiring greatly extended stress times, or high currents that cause thermal anomalies. Software simulations of the electromigration mechanism combined with characterization of temperature profiles allows extracting material parameters and calculation of design rules to ensure reliable interconnect. Test structures capable of demonstrating Blech threshold effects while allowing thermal characterization were designed and processed. Electromigration stress tests at various conditions were performed to extract both short-line (threshold) and long-line (above threshold) performance values. The resistance increase time constant shows immortality below $J_e \cdot L$ (product of current density and segment length) of 3200 amp/cm. Statistical analysis of times-to-failure show that long lines last 10^5 hours at $3.1 \text{ mA}/\mu\text{m}^2$ (120°C). While this is more robust than aluminum interconnect, the semiconductor industry will be challenged to improve that performance as future products require.

Chapter 1

Introduction

Electromigration (EM) is the motion of mass induced by an electric field and resulting electric current. In the case of solid valence metals (like copper) the random diffusion of individual atoms is 'skewed' in one direction by momentum transfer from the electron current. This effect is significant over extended duration in short fine interconnects, such as in integrated circuits (ICs). The resulting change in morphology causes failure of IC products and is a leading reliability problem, especially in high-power microprocessors.

1.1. Future Trends in Integrated Circuits

The microelectronic trend toward faster devices and smaller features places performance and reliability constraints on interconnects. For the first 40 years of ICs, product performance was primarily determined by the active devices while interconnects played a necessary supporting role. In the next 40 years, product performance will be increasingly dependent on the interconnect technology, with active devices playing a necessary supporting role. Portable entertainment devices were once called 'transistor radios' in honor of the key technology, while those in the future might be called 'nanowire boxes' – the label honoring the key technology enabling higher performance and longer battery life.

As each technology generation operates at increased operating frequency, time delays in interconnects have increased while time delays in transistors have decreased. The time delay of a typical CMOS inverter stage can be estimated by the product of resistance (R) and capacitance (C) in the interconnects (wire) and transistors (xtr) according to Elmore's rule:

$$T_{delay,50\%} = 0.4 \cdot R_{wire} \cdot C_{wire} + 0.7 \cdot (R_{wire} \cdot C_{xtr} + R_{xtr} \cdot C_{wire} + R_{xtr} \cdot C_{xtr}) \quad (1.1)$$

Of the four terms, the transmission line delay ($0.4 \cdot R_{wire} \cdot C_{wire}$) for global-length lines greatly exceeds the gate delay ($0.7 \cdot R_{xtr} \cdot C_{xtr}$), and both cross-terms are strongly dependent on the interconnects. Repeaters can maintain the performance trend, but consume power and chip area and thus degrade the speed-power product.

The capital cost of wafer fabrication equipment is evidence of the key role that interconnects already play in today's products. Interconnects are fabricated in the back-end processes (thin-film and etch), while active devices are made by the front-end processes (diffusion and implant). The ratio of back-end/front-end equipment cost in a typical fab was about 30/70 in the 1980's, but has become about 60/40 and is likely to increase further.

1.2. Copper and Supporting Roles

Continued reduction in the speed-power product requires minimum R-C delay of interconnects, focusing industry attention on lower resistance metal and lower dielectric constant insulators. Several industry associates have cooperated to develop a global consensus regarding future technology needs, as formalized in The International Semiconductor Technology Roadmap (ISTR),[1] which states: "New materials, structures, and processes create new chip reliability (electrical, thermal, and mechanical) exposure. Detecting, testing, modeling and control of failure mechanisms will be key."

Copper is replacing aluminum as metal-of-choice for future semiconductor processes, which brings advantages and disadvantages. The greatest advantage is 38% reduction in electrical resistivity, as shown in Table 1.1. The associated downside, however, is the longer mean-free-path of electrons, which implies greater surface scattering. Aluminum should maintain its' bulk resistivity to smaller feature sizes, and with continued scaling the carrier density suggests that Al might even outdo Cu.

Table 1.1: Properties of Copper and Aluminum

	resistivity ² ($\mu\text{Ohm-cm}$)	mean-free-path (nm)	electron density ($\times 10^{23} / \text{cm}^3$)	melting temperature ($^{\circ}\text{K}$)
Cu	1.637	39.	0.9	1357
Al	2.655	15.	1.8	934
at 293 $^{\circ}\text{K}$				

The greatest disadvantage of Cu may be lack of a dry metal-etch process, requiring other patterning methods. The favored alternative is the damascene process, which involves dry-etch of dielectric trenches, filling with a conformal Cu layer, and chemical-mechanical polishing (CMP) to remove overburden and electrically isolate lines. While dielectric etch provides excellent definition of small features, CMP is relatively expensive with low throughput and high consumables.

Another disadvantage of Cu is high diffusivity in oxide and silicon, especially influenced by a voltage bias. Copper in the dielectric increases electrical leakage, and formation of dendrites that can lead to shorting of adjacent lines. In silicon, copper forms a mid-gap trap state which reduces minority carrier lifetime and 'kills' semiconductor device performance. Degradation is prevented by cladding Cu on all sides in a diffusion barrier. Typically, Cu is deposited on a refractory metal layer which protects dielectric below and to each side, and Cu is covered by silicon-nitride to protect the overlying oxide. The refractory metal 'glue' layer improves the rather poor adhesion of copper to oxide, serves as an electromigration barrier between levels of a multilevel metal process, and carries current around copper voids (as a shunt layer) maintaining electrical continuity. The silicon-nitride cap also serves as via etch stopping layer.

Current density increases in successive IC product generations, as lines become smaller while carrying about the same current. Table 1.2 shows excerpts from the ISTR, demonstrating the aggressive scaling of future processes. The interconnect line current density, J_{max} , is expected to increase 4x over the next 15 years as lines scale to smaller dimensions. At sufficient current density, Joule self-heating (SH) and EM can lead to product failure and it is expected that Cu will be more robust.

Table 1.2 Expected Interconnect Technology Requirements (source: ISTR)

Year of introduction	2001	2004	2007	2010	2013	2016
local wiring pitch (nm)	350	210	150	105	75	50
barrier thickness (nm)	16	10	7	5	3.5	2.5
bulk diel. constant, k	2.7	2.4	2.1	1.9	1.7	1.6
RC delay (pSec/mm)	21	37	79	131	248	452
Jmax (MA/cm ²)	0.96	1.5	2.1	2.7	3.3	3.9
I _{max} (mA/via)	0.32	0.24	0.18	0.1	0.07	0.04
Blech length (um)	31.3	20.0	14.3	11.1	9.1	7.7
ratio pitch/Blech length	11.2	10.5	10.5	9.5	8.3	6.5

1.3. Electromigration

Electromigration, EM, is the motion of a solid material under the influence of electron current. High electric currents provide an electrodiffusion driving force which moves metal along the length of an interconnect line. As a pipeline analogy might suggest, metal density decreases at one end of the line, increases at the other end, and (barring defects) remains relatively constant at mid-line. Decreased density results in a void which can cause failure due to increased resistance in the line. Increased density results in hillocks or extrusions which can cause failure due to shorting to adjacent lines, as shown in Figure 1.1.

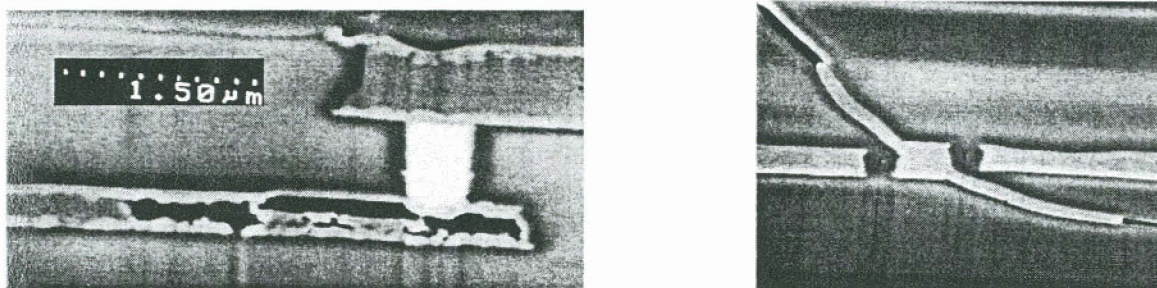


Figure 1.1: EM Failure due to voiding (left) and extrusions (right)

Higher melting temperature generally correlates to more robust EM reliability. The typical IC operating temperature (120°C) is 42% of aluminum's melting temperature, but only 28% that of copper. By operating further from a molten state, copper atoms should be less mobile and suffer less migration. While true for long lines where reliability is related to mobility, the situation is more complicated for short lines where mobility is effectively cancelled and thus may have little or no effect.

Understanding of both long-line and short-line EM performance is critical to modern designs. Long lines achieve acceptable failure rate at electric current density, J_e , below J_{max} . As predicted by Rent's Rule, most lines on VLSI products are short enough to benefit from the Blech effect, which greatly improves reliability. Short lines with $J_e \cdot L$ (product of current density, J_e , times segment length, L) below JL_{max} are immortal – even at $J > J_{max}$. This 'EM threshold' transition corresponds to the 'Blech length' shown in Table 1.2 (assuming $J = J_{max}$ and $JL_{max} = 3000$ amp/cm). As pitch scales 7x in 15 years, the expected Blech length increases 4x, extending the benefit to a greater fraction of all interconnects. To achieve this up-side potential, EM threshold must remain constant in subsequent technology generations.

The 4x increase in J_{max} shown in Table 1.2 (with constant JL_{max}) will require an interconnect process that achieves effective synergy between all materials that affect EM, including the copper, barrier, dielectric, and containment. The challenge will be to maintain robust EM performance with weaker materials. As Table 1.2 shows, RC delay will still increase 20x, in spite of an 84% reduction in barrier thickness and 41% decrease in dielectric constant. 'Making do' with stronger barrier or dielectric is not an option.

1.4. Process Monitor and Control

An aphorism states "Anyone can design a bridge that stands up. An *engineer* can design a bridge that *just barely* stands up." There are many reasons that society might prefer a weaker bridge, including reduced cost, environmental impact, and construction time. The allocation of resources to overbuilding a bridge takes away from other priorities. Just as bridges are made to be safe and cheap (in that order), interconnects must have high performance at low cost. (Reliability is simply performance over

extended duration.) The engineering problem is to make the many tradeoffs that result in the optimum compromise.

Manufacturing control is essential to maintaining the numerous material properties that affect long and short line EM performance. Since EM, like most reliability failure mechanisms, is susceptible to the 'weak-link', it is important to identify which few key parameters determine failure during normal use conditions. This can be difficult because testing is performed using accelerated conditions which may induce different failure mechanisms that depend on other key parameters. Practical process monitors must characterize additional parameters – those that affect acceleration in addition to those that affect real-life. Several mechanisms contribute to failure, and extrapolations to product-use conditions must employ acceleration factors appropriate for each mechanism.

Statistical analysis of resistance changes measured during accelerated stress testing provides estimates of times-to-failure. The large variability in time-to-failure between samples requires large sample sizes for reasonable precision. On the other hand, smaller variability in void growth kinetics allows measuring this step toward EM failure with reduced sample size. Understanding of the sources of variation can allow design of efficient experiments with minimum total sample size.

1.5. Dissertation Overview

Material and process parameters affect long and short line damascene Cu EM performance, which determine the 'do-not-exceed' limits of IC designs. The parameters measured by means of wafer-level EM stress testing will be reported. Practical considerations for design rules will be explored to estimate what portion of actual EM performance is actually useable and can be implemented in real-world chips.

This dissertation is divided into 7 chapters. After this introductory chapter, the next chapter will review the drift/diffusion theory of electromigration. The following chapter will describe test structure design, processing, and preparation. Chapter 4 will describe how Joule self-heating is predicted, followed by a description of the experimental procedure and measured results of accelerated EM stress testing. Consequences for product design rules are presented in Chapter 6. The concluding chapter reviews the major findings and outlines the significance to the industry.

Chapter 2

Electromigration Theory

Electromigration is the motion of a solid material under the influence of electron current. (While the term sometimes includes liquid-state migration, such as electrophoresis, this dissertation restricts EM to motion of solid metals below melting temperatures.) Atoms in solid metals are constrained in a polycrystalline structure (fcc for Al and Cu). Convection is frozen out, so motion requires the presence of point defects. Possible defects include several types of interstitials and n-vacancies, but for Al and Cu the monovacancy is by far the dominant diffusion mechanism.[3]

2.1. Vacancy Dynamics

The EM mechanism is a result of the dynamics of vacancies, including drift, diffusion, creation, and recombination.

2.1.1. Equilibrium vacancy concentration

Formation of a vacancy requires an amount of energy, H . It is helpful to visualize the creation of a vacancy in a simple square matrix in 2 dimensions, shown as the progression from points A to E in Figure 2.1. Each broken bond represents a shift ΔH_i to higher energy state. Extinguishing a vacancy recovers energy $H = \Delta H_3 + \Delta H_4$ and leaves the surface unchanged, except the step is shifted by one lattice site. A vacancy created at point X and extinguished at to point Y has the effect of moving an atom from Y to X while cooling X and warming Y.

Several diffusion paths are possible. A vacancy (E) moves with bulk mobility, a depression (D), moves with surface mobility, and in 3-D a 'notch' in the step moves with linear mobility along the step (C).

Passivation greatly reduces surface mobility and slows vacancy generation. Creation of a vacancy at a passivated surface requires that a metal atom shift into a vacant

site in the passivating material, but effective passivations have higher energy-of-vacancy-formation than the metal.

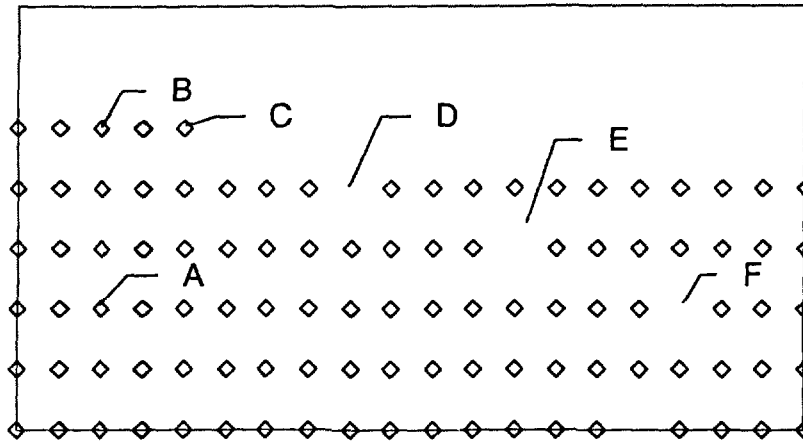


Figure 2.1: 2-D vacancy formation

Label	site	broken bonds	formula
A	Lattice	0	$A + \Delta H1 = B$
B	Surface	1	$B + \Delta H2 = C$
C	Step	2	$C + \Delta H3 = D$
D	Depression	3	$D + \Delta H4 = E$
E and F	Vacancies	4	

Grain boundaries are planes of crystal orientation mismatch, and can be modeled as sheets of strained bonds. Vacancies in a grain boundary have formation energies corresponding to fractional broken bonds, the fraction depending on the crystal mismatch and specific location. This implies that a vacancy in a grain boundary is an intermediate state between sites D and E in Figure 2.1, with a corresponding intermediate energy-of-formation, H.

The equilibrium vacancy concentration depends on the energy-of-vacancy-formation, H through the second law of thermodynamics (2.1), showing work done on the system, U; temperature, T; entropy, S; pressure, P; volume, V.

$$\partial U = T \cdot \partial S - P \cdot \partial V \quad (2.1)$$

Converting terms to per-vacancy units:

$$\frac{\partial U}{\partial n} = T \cdot \frac{\partial S}{\partial n} - P \cdot \frac{\partial V}{\partial n}, \text{ or } H = T \cdot \frac{\partial S}{\partial n} + \sigma \cdot \Lambda \quad (2.2)$$

where n is the quantity of vacancies, σ is hydrostatic tensile stress, and Λ is atomic volume. Entropy includes thermal and configurational components, the second term resulting from the number of ways that N atoms and n vacancies can be distributed among $n+N$ lattice sites, where N is the quantity of atoms.

$$S = S_{thermal} + S_{config} = n \cdot S_i + k_B \cdot \ln \left(\frac{(N+n)!}{N!n!} \right) \quad (2.3)$$

which simplifies by using Stirling's approximation and taking the derivative: [4]

$$\frac{\partial S}{\partial n} \equiv S_i + k_B \cdot \ln \left(\frac{N+n}{n} \right) \quad (2.4)$$

Substituting (2.4) into (2.2) and rearranging:

$$\frac{n}{N+n} = \exp \left(\frac{S_i}{k_B} - \frac{H - \sigma \cdot \Lambda}{k_B \cdot T} \right) \quad (2.5)$$

At infinite temperature, all energy states are filled with equal probability. This implies that $n=N$, or

$$\left(\frac{n}{N+n} \right)_{T=\infty} = \frac{1}{2} = \exp \left(\frac{S_i}{k_B} \right) \quad (2.6)$$

From (2.5) and (2.6), we derive vacancy concentration, C :

$$C = \frac{n}{V} = \frac{n}{(N+n) \cdot \Lambda} = \frac{1}{2 \cdot \Lambda} \exp \left(\frac{-(H - \sigma \cdot \Lambda)}{k_B \cdot T} \right) \quad (2.7)$$

Clearly, higher temperature and higher stress result in higher vacancy concentration. The stress dependence vanishes at infinite temperature, and the temperature dependence vanishes as σ approaches H/Λ . Figure 2.2 shows that vacancy concentration varies many orders of magnitude over possible conditions. Local differences in H will cause variation in vacancy concentration.

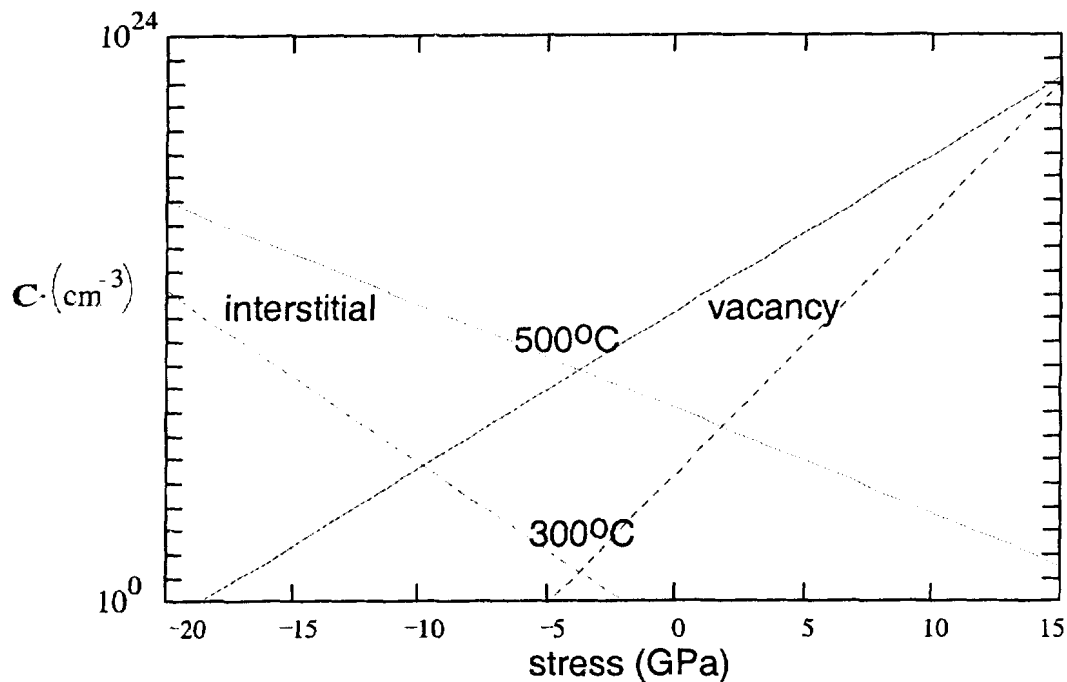


Figure 2.2 Concentration of Defects

Measurements of diffusion show a pressure dependence consistent with an 'activation volume' that is slightly less than the atomic volume. It is suggested that the neighboring atoms shift position toward the vacancy, thus reducing its effective volume by 10% for copper (29% for aluminum). This approximate treatment neglects the difference between atomic volume and activation volume of vacancies, and uses the same symbol, Λ , for both. Interstitials will have negative activation volume, suggesting that concentration increases with compressive stress, as shown in Figure 2.2.

2.1.2. Vacancy Motion

The velocity of an individual vacancy, v_v , is the product of atomic spacing, a , and jump rate, r . ($v_v = a \cdot r$) The combined flux of all vacancies is the product of concentration, C and velocity.

$$J_m = C \cdot v_v = C \cdot a \cdot r \quad (2.8)$$

If jumps are random and there is no correlation between successive jumps, the net flux is the difference between forward and backward fluxes:

$$J_m = J_{forward} - J_{backward} = C(x) \cdot a \cdot r_{forward}(x) - C(x+a) \cdot a \cdot r_{backward}(x+a) \quad (2.9)$$

Atoms vibrate at frequency, ν , and jump into an adjacent vacancy at a rate predicted by quantum mechanics:[5]

$$r = \nu \cdot \exp\left\{\frac{-E}{k_B \cdot T}\right\} \quad (2.10)$$

Application of a force, F , effectively decreases the forward barrier by energy $F \cdot a/2$ and increases the reverse barrier by the same amount. The resulting jump rates are:

$$r_{forward} = \nu \cdot \exp\left\{\frac{-E + F \frac{a}{2}}{k_B \cdot T}\right\}, \quad r_{backward} = \nu \cdot \exp\left\{\frac{-E - F \frac{a}{2}}{k_B \cdot T}\right\} \quad (2.11)$$

Rewriting (2.9):

$$\begin{aligned} J_m &= C(x) \cdot a \cdot \nu \cdot \exp\left\{\frac{-E + \frac{F \cdot a}{2}}{k_B \cdot T(x)}\right\} - C(x+a) \cdot a \cdot \nu \cdot \exp\left\{\frac{-E + \frac{F \cdot a}{2}}{k_B \cdot T(x+a)}\right\} \\ &= \frac{C(x) \cdot D(x)}{a} \cdot \exp\left\{\frac{F \cdot a}{2 \cdot k_B \cdot T(x)}\right\} - \frac{C(x+a) \cdot D(x+a)}{a} \cdot \exp\left\{\frac{-F \cdot a}{2 \cdot k_B \cdot T(x+a)}\right\} \end{aligned} \quad (2.12)$$

where diffusion coefficient, D , is defined as:

$$D = a^2 \cdot r_0 = a^2 \cdot \nu \cdot \exp\left\{\frac{-E}{k_B \cdot T}\right\} \quad (2.13)$$

Using the continuum approximation:

$$\begin{aligned} C(x+a) \cdot D(x+a) &= \left[C(x) + a \cdot \frac{\partial C}{\partial x} \right] \cdot \left[D(x) + a \cdot \frac{\partial D}{\partial x} \right] \\ &= C \cdot D + a \cdot D \frac{\partial C}{\partial x} + a \cdot C \frac{\partial D}{\partial x} + a^2 \frac{\partial C}{\partial x} \frac{\partial D}{\partial x} \end{aligned} \quad (2.14)$$

$$\begin{aligned} J_m &= \frac{C(x) \cdot D(x)}{a} \left[\exp\left\{\frac{F \cdot a}{2 \cdot k_B \cdot T(x)}\right\} - \exp\left\{\frac{-F \cdot a}{2 \cdot k_B \cdot T(x+a)}\right\} \right] \\ &\quad - \left(D \frac{\partial C}{\partial x} + C \frac{\partial D}{\partial x} + a \frac{\partial C}{\partial x} \frac{\partial D}{\partial x} \right) \cdot \exp\left\{\frac{-F \cdot a}{2 \cdot k_B \cdot T(x+a)}\right\} \end{aligned} \quad (2.15)$$

2.1.3. Electromigration Wind Force

Electron collisions induce a force, F , on vacancies. Conduction electrons are modeled as traveling waves, which deposit energy and momentum in the atomic lattice. A point defect interrupts the periodic potential surrounding each lattice atom, with the result that the electron scattering cross-section predicted by quantum mechanics is much larger for a point defect than for an atom at a lattice position. The momentum transfer biases the otherwise random direction of vacancy jumps, such that atomic motion is encouraged in the direction of current flow and discouraged against the current flow. Ohm's law predicts that electric current density, J_e , increases with the product of applied electric field, E , and resistivity, ρ . The force on an atom at a nearest neighbor site of a vacancy is modeled as that of a charged particle having an 'effective valence', Z^* , times the electron charge, q .

$$F = Z^* \cdot q \cdot E = Z^* \cdot q \cdot \rho \cdot J_e \quad (2.16)$$

For metals with negative charge carriers (as evidence by a negative Hall effect), the momentum transfer from the 'electron wind' directs atomic motion toward the anode, despite the electric field pulling the partially shielded ionic core to the cathode. This tug-of-war leads to a net negative valence for Ti, Ni, Cu, Ag, and Au.[6] For metals with positive carriers (positive Hall effect) such as V, Fe, Co, and Nb, the 'hole wind' increases the positive ionic valence moving matter to the cathode. (An interesting exception is Mo which migrates towards the anode despite a positive Hall effect.) Copper and aluminum have negative effective valence, variously measured from -4 to -10, corresponding to vacancy migration toward the cathode (metal motion toward the anode).

Energy is force acting through a distance, so a point defect moving distance L dissipates an 'EM energy' = $F \cdot L$. At the same time, thermal energy $k_B \cdot T$ moves defects in random directions, so EM is significant for $F \cdot L > k_B \cdot T$. It is useful to define a characteristic EM length, δ_{EM} , which is the shortest length for which EM effects become apparent above the thermal noise:

$$\delta_{EM} = \frac{k_B \cdot T}{F} = \frac{k_B \cdot T}{Z^* \cdot q \cdot \rho \cdot J_e} \quad (2.17)$$

Despite having temperature in the denominator, δ_{EM} is not significantly temperature-dependent (because the thermal coefficient of resistivity keeps the ratio T/ρ nearly constant), so δ_{EM} is primarily current dependent. Figure 2.3 shows the magnitude of δ_{EM} , which ranges from a few microns at high-current acceleration conditions to the entire structure length at typical low-current operation. Force, F , has direction and may be positive or negative (depending on the polarities of Z^* and J_e) but energy = $F \cdot \delta_{EM}$ must be positive, so we use a convention that δ_{EM} has the same sign as F . (Negative distance extends the other direction.)

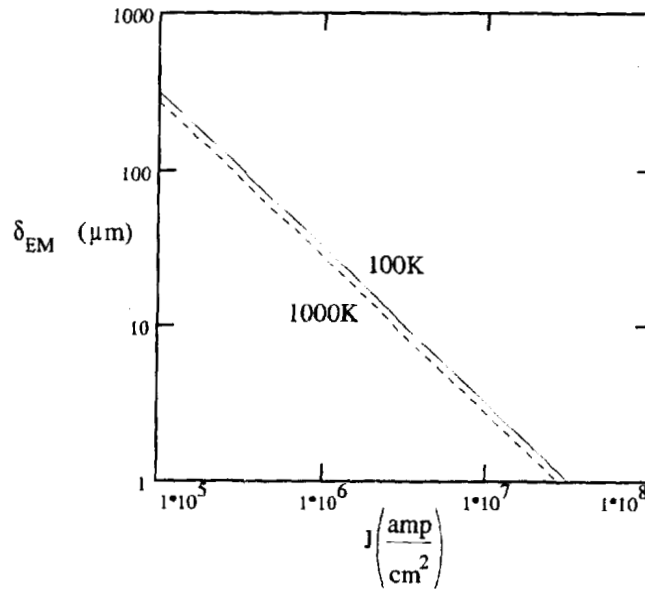


Figure 2.3 EM Characteristic Length (for $Z^* = -5$)

2.1.4. Migration Mechanisms

Equation (2.15) becomes:

$$J_m = \frac{C \cdot D}{a} \left[2 \cdot \sinh \left\{ \frac{a}{2 \cdot \delta_{EM}} \right\} \right] - \left(D \frac{\partial C}{\partial x} + C \frac{\partial D}{\partial x} + a \frac{\partial C}{\partial x} \frac{\partial D}{\partial x} \right) \cdot \exp \left(\frac{-a}{2 \cdot \delta_{EM}} \right) \quad (2.18)$$

Since $\frac{a}{2 \cdot \delta_{EM}} = y \ll 1$, $\sinh(y) \cong y$ and $\exp(-y) \cong 1$, and (2.18) simplifies to the

drift/diffusion equation:

$$J_m = \frac{C \cdot D}{\delta_{EM}} - D \frac{\partial C}{\partial x} - C \frac{\partial D}{\partial x} - a \frac{\partial C}{\partial x} \frac{\partial D}{\partial x} \quad (2.19)$$

The first term in (2.19) is the EM drift flux, while the second term accounts for Fickian diffusion in a concentration gradient. Taking partial derivatives of (2.7) and (2.13):

$$\frac{\partial C}{\partial x} = \left(\frac{\partial C}{\partial T} \right) \left(\frac{\partial T}{\partial x} \right) + \left(\frac{\partial C}{\partial \sigma} \right) \left(\frac{\partial \sigma}{\partial x} \right) = \frac{C \cdot (H - \sigma \cdot \Lambda)}{k_B \cdot T^2} \cdot \frac{\partial T}{\partial x} + \frac{C \cdot \Lambda}{k_B \cdot T} \cdot \frac{\partial \sigma}{\partial x} \quad (2.20)$$

$$\frac{\partial D}{\partial x} = \left(\frac{\partial D}{\partial T} \right) \left(\frac{\partial T}{\partial x} \right) = \frac{C \cdot E}{k_B \cdot T^2} \cdot \frac{\partial T}{\partial x} \quad (2.21)$$

equation (2.19) becomes the sum of various vacancy flux contributions:

$$J_m = \frac{C \cdot D}{\delta_{EM}} - \frac{C \cdot D \cdot (E + H - \sigma \cdot \Lambda)}{k_B \cdot T^2} \cdot \frac{\partial T}{\partial x} - \frac{C \cdot D \cdot \Lambda}{k_B \cdot T} \frac{\partial \sigma}{\partial x} + a \frac{\partial C}{\partial x} \frac{\partial D}{\partial x} \quad (2.22)$$

The second term describes the vacancy flux arising from a thermal gradient. This mass transport is known as thermo-migration,[7] and gives rise to the Soret effect.[8] The third term describes the vacancy flux arising from a hydrostatic stress gradient. This mass transport is known as stress-migration, and gives rise to diffusional creep. The last flux term is a cross-term and is may be significant at the extreme gradients around metal voids.

Equation (2.22) could be restated as the vacancy drift velocity, which is the sum of the various velocity contributions:

$$v_v = \frac{J_m}{C} = \frac{D}{\delta_{EM}} - \frac{D \cdot (E + H - \sigma \cdot \Lambda)}{k_B \cdot T^2} \cdot \frac{\partial T}{\partial x} - \frac{D \cdot \Lambda}{k_B \cdot T} \frac{\partial \sigma}{\partial x} + \dots \quad (2.23)$$

and (2.22) can be restated as the sum of inverse characteristic lengths:

$$\frac{J_m}{C \cdot D} = \frac{Z^* \cdot q \cdot \rho \cdot J}{k_B \cdot T} - \frac{(E + H - \sigma \cdot \Lambda)}{k_B \cdot T^2} \cdot \frac{\partial T}{\partial x} - \frac{\Lambda}{k_B \cdot T} \frac{\partial \sigma}{\partial x} + \dots \quad (2.24)$$

or as the sum of 'effective forces' on each vacancy (nearest-neighbor atom):

$$\frac{J_m \cdot k_B \cdot T}{C \cdot D} = Z^* \cdot q \cdot \rho \cdot J_e - \frac{(E + H - \sigma \cdot \Lambda)}{T} \cdot \frac{\partial T}{\partial x} - \Lambda \frac{\partial \sigma}{\partial x} + \dots \quad (2.25)$$

2.1.5. Diffusion Measurements

Diffusion is a thermally activated process with an exponential dependence on temperature. The diffusion coefficient (2.13) is defined by the activation energy for vacancy velocity, E . The vacancy concentration (2.7) is determined by the activation energy for vacancy creation, H . The total mass flux (2.8) is the product of concentration and velocity, with a combined activation energy $E+H$.

Self-diffusion of thin surface layers into a single crystal substrate have been measured using radioactive tracers[3] with results interpreted to report a 'mass diffusion coefficient', D_m , in terms of the combined $E+H$. The mass diffusion coefficients for a few elements are shown in Figure 2.4 and the combined activation energies ($E+H$) are listed in Table 2.1.

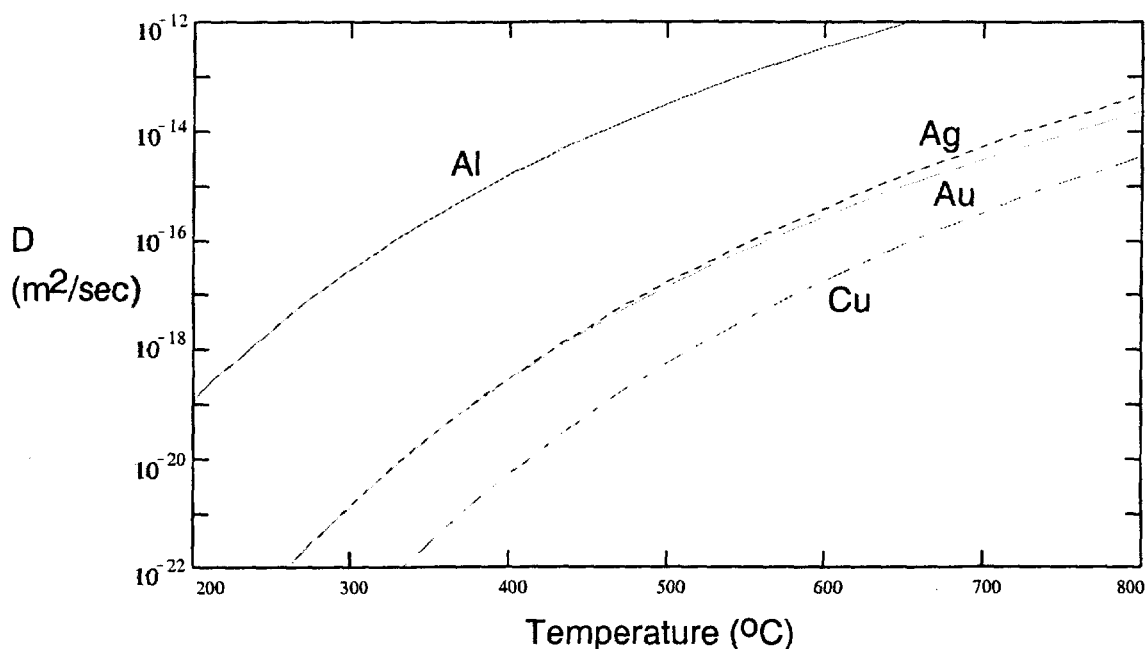


Figure 2.4: Mass Diffusion Coefficient Measured by Radioactive Tracers

EM activation energies measured on patterned interconnects are much less than the values for single crystal bulk samples in Table 2.1, showing that the mechanism is different. The dominant diffusion path in interconnects is likely to be along grain boundaries or along interfaces with surrounding materials (surfaces).

Table 2.1: Activation Energy Measured by Radioactive Tracers

Metal	Activation energy (eV)	Do (cm ² /sec)
Cu	2.07	0.217
Au	1.78	0.066
Ag	1.83	0.245
Ni	2.99	2.28
Al	1.33	0.183

2.2. Stress Evolution

The flux equation (2.22) can be used in combination with the continuity equation (2.26) to solve for concentration. The change in concentration is simply the difference between generation (creation), G , and recombination (extinguishment), R , and the net flux out of the region.

$$\frac{\partial C}{\partial t} = G - R - \nabla J_m \quad (2.26)$$

2.2.1. Blech Condition

Blech observed that EM drift velocity was proportional to current in unconstrained systems, but greatly reduced in passivated structures.[9] The drift velocity of constrained lines exhibited an initial value that decayed to zero, which he proposed as evidence of stress-induced reverse mass flow canceling the EM drift.

After sufficient stress time, concentration is time invariant everywhere ($dC/dt=0$). If we wait a bit longer, generation and recombination cancel in equation (2.26), so the gradient of flux also vanishes. If the gradient is zero and the boundaries maintain zero flux at the segment ends, then flux vanishes everywhere. Solving (2.25) for $J_m=0$,

$$Z^* \cdot q \cdot \rho \cdot J_e - \frac{(E + H - \sigma \cdot \Lambda)}{T} \cdot \frac{\partial T}{\partial x} - \Lambda \frac{\partial \sigma}{\partial x} + \dots = 0 \quad (2.27)$$

Diffusion coefficient drops out of the steady-state solution, so the E+H dependence is linear rather than exponential. At iso-thermal conditions ($\partial T/\partial x=0$), stress is linear and independent of E+H.

$$\frac{\partial \sigma}{\partial x} = \frac{F}{\Lambda} = \frac{Z^* \cdot q \cdot \rho \cdot J_e}{\Lambda} \quad (2.28)$$

For constant current density ($\partial J_e/\partial x=0$ implies uniform cross-section area), the solution for stress is found by integrating along the segment length, L:

$$\Delta \sigma = \sigma_{anode} - \sigma_{cathode} = \frac{Z^* \cdot q \cdot \rho}{\Lambda} \cdot (J_e \cdot L) \quad (2.29)$$

The stress change along the line is proportional to the $J_e \cdot L$ product. Longer lines will have larger tensile and compressive stresses in proportion to the line length, so the stress gradient limits the length of line that can support a steady-state solution. At some length, some material strength is exceeded and failure occurs (either voiding or extrusion). The stress change in (2.29) at failure defines the $J_e \cdot L$ product at failure, which is the product of current threshold times the Blech length, L_{Blech} . Solving (2.29) for L shows the relationship between Blech Length and the characteristic EM length:

$$L_{Blech} = \frac{\Delta \sigma_{max} \cdot \Lambda}{Z^* \cdot q \cdot \rho \cdot J_e} = \delta_{EM} \cdot \frac{\Delta \sigma_{max} \cdot \Lambda}{k_B \cdot T} \quad (2.30)$$

The $J_e \cdot L$ threshold is relatively constant (but possibly temperature-dependent) so Blech Length is inversely proportional to current threshold. Segments of longer length fail at proportionately lower current density, and the time-dependent solution takes much longer to reach the Blech equilibrium. Shorter segments tolerate higher current density, but increasing currents eventually causes Joule self-heating and significant thermal gradients.

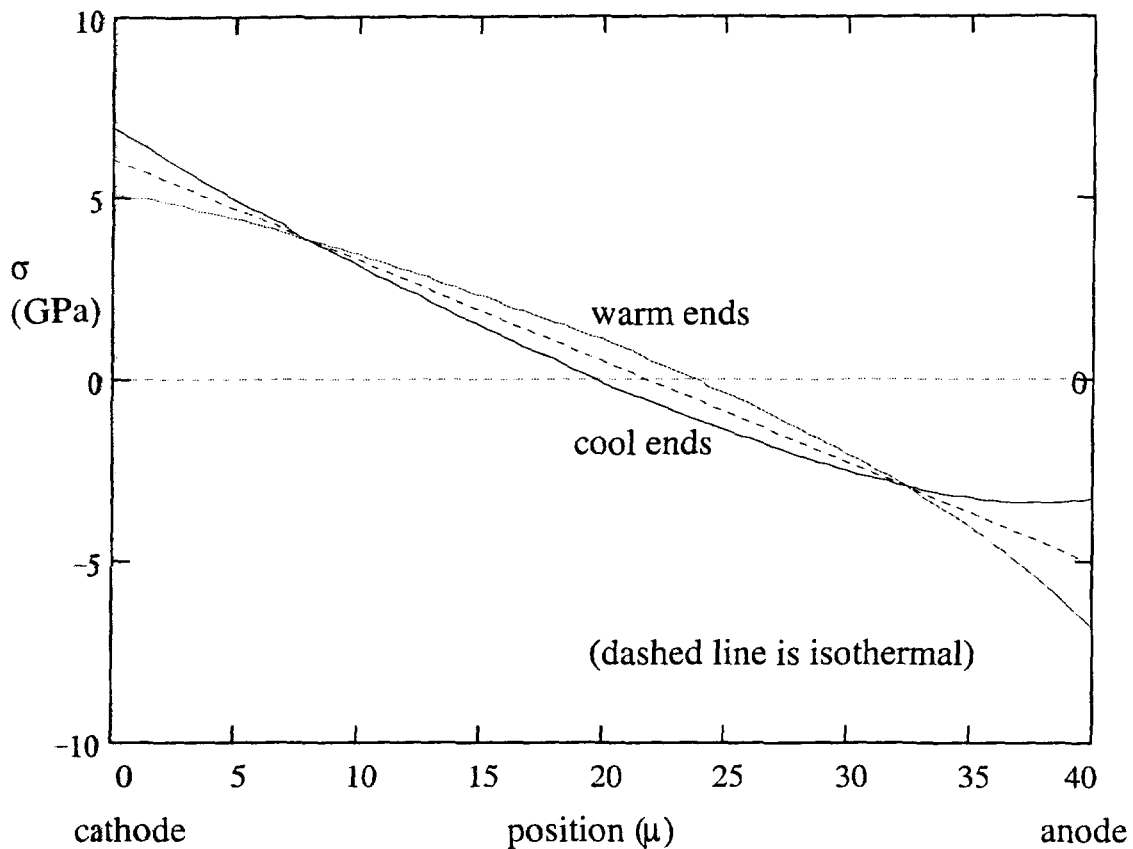


Figure 2.5 Stress vs position in interconnect segment

With the added driving force of a thermal gradient, equilibrium stress is nonlinear with distance. The solution to the differential equation (2.27) along a segment with uniform cross-section is shown in Figure 2.5 for representative examples of temperature vs. position. Nonlinear effects of thermal gradients cause the stress vs. position curve to become concave-down for interconnect segments with ends warmer than the middle, and concave-up when ends are cooler than the middle. For extreme thermal gradients, the stress curve may have an inflexion point, so the maximum compressive stress of a concave-down curve may occur some distance upstream from the anode.

To find the location of peak stress, solve (2.27) for $\partial\sigma/\partial x=0$:

$$\frac{1}{T} \frac{\partial T}{\partial x} = \frac{F}{E + H - \sigma \cdot \Lambda} = \frac{1}{\chi} \quad (2.31)$$

We consider a simple thermal model for Joule self-heating:

$$T(x) = T_{mid} + (T_{end} - T_{mid}) \cdot \exp\left(\frac{-x}{\lambda}\right) \quad (2.32)$$

Substituting (2.32) into (2.31) and solving for x :

$$x(J) := \lambda \cdot \ln \left[\left(1 - \frac{T_{end}(J)}{T_{mid}(J)} \right) \cdot \left(1 + \frac{x(J)}{\lambda} \right) \right] \quad (2.33)$$

which is plotted in Figure 2.6 for a representative structure with ends cooler than the middle. Only positive solutions are valid, so $x < 0$ indicates that the thermal gradient does not overcome the EM force, and peak stress occurs at the anode. The offset from the anode increases as the logarithm of current until self-heating everywhere is extremely high. Remember that random thermal fluctuations of a few $k_B \cdot T$ are expected, so offsets smaller than δ_{EM} are not significant, shown in Figure 2.6 by a shaded region.

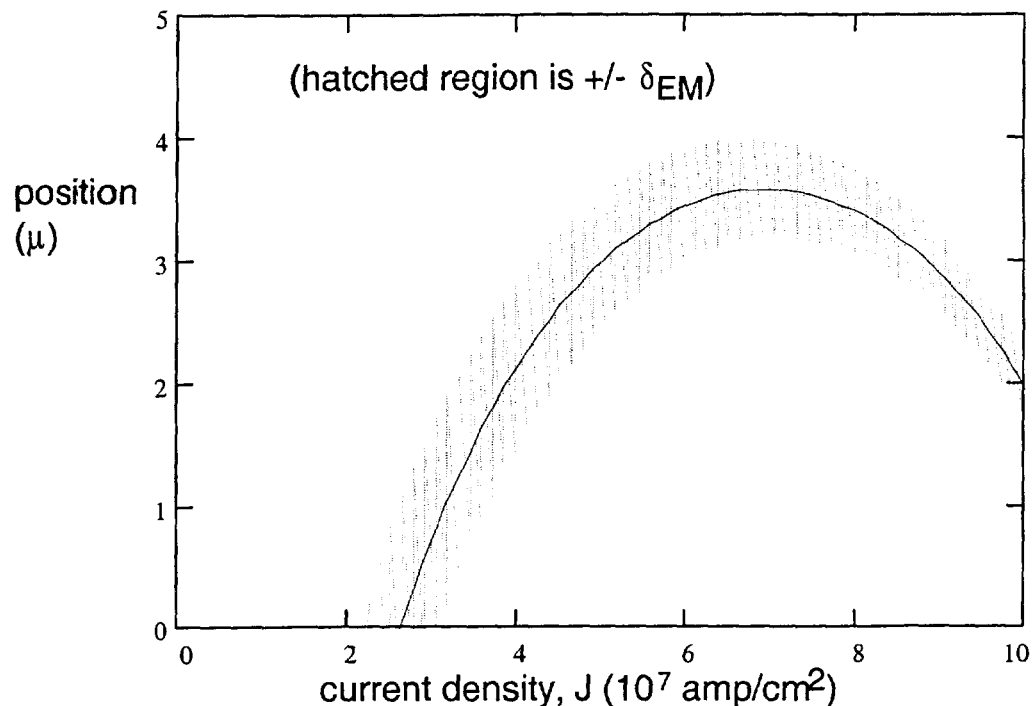


Figure 2.6 Offset from Anode of Maximum Compressive Stress

2.2.2. Time dependence

Using appropriate boundary conditions, (2.26) can be solved to determine characteristic time constants for approaching steady-state.[10] Interconnect typically has length much larger than width or thickness, allowing wires to be modeled as

1-dimensional segments. For constant current density (constant cross-sectional area), constant D (uniform temperature), and blocking boundaries (no flux in or out at the ends), the two time-constant components are:

$$\frac{1}{\tau} = \frac{1}{\tau_d} + \frac{1}{\tau_e} = \frac{\pi^2 \cdot D_{eff}}{L^2} + \frac{D_{eff}}{4 \cdot \delta_{EM}^2} = \frac{\pi^2 \cdot D_{eff}}{L^2} \left(1 + \left(\frac{L}{2\pi \cdot \delta_{EM}} \right)^2 \right) \quad (2.34)$$

2.2.3. AC EM

While (2.26) and (2.34) apply to constant electron current conditions, it is simple to consider time-varying current waveforms. The total amount of mass flow can be obtained by integration over time of equation (2.22) for vacancy flux. If the parameters including D and C don't vary significantly over the pulse period, then the average vacancy flux is that expected at the average electron current. The time average

(expectation value, mean) is defined: $\langle a \rangle = \frac{1}{duration} \cdot \int_{duration} a \cdot dt$

$$\langle J_m \rangle = \frac{1}{duration} \cdot \int_{duration} J_m [J_e(t)] \cdot dt \equiv J_m \langle J_e \rangle \quad (2.35)$$

Changes in C are expected over the range of a few minutes, while changes in D from temperature are expected over a few microseconds (see section 4.1). For the GHz clock frequencies in modern integrated circuits, the approximation in (2.35) is very good. The mean-time-between collisions (relaxation time) for electrons can be estimated from the drift velocity for electrons, which in turn depends on resistivity with the corresponding temperature dependence:[11]

$$v_e = \frac{-q \cdot E \cdot \tau}{m_e}, \quad J_e = \frac{-Z \cdot q}{\Lambda} \cdot v_e = \frac{E}{\rho}, \quad \tau = \frac{m_e}{\rho \cdot n \cdot q^2} \quad (2.36)$$

The relaxation time for Cu at typical EM temperatures is in the range $1-2 \times 10^{-14}$ seconds. This is about 100 collisions during a 1 picosecond current pulse (which is 1% of the period of a 10 GHz clock frequency). The number of atoms and conduction electrons are equal (Cu valence is 1), so the average Cu atom would experience 100 collisions per picosecond. A defect has a much larger collision cross-section than a lattice atoms so it is

reasonable to suppose that a vacancy experiences many thousands of collisions per picosecond.

If the current is switched off, the first term of (2.22) goes to zero, so the structure relaxes according to the second, third, and fourth terms. If these later terms are negligible, then the vacancy distribution is 'frozen' while the current is switched off. This results in an on-time correction to the time dependence. (If current is switched off half the time, the structure will take twice as long to fail.)

2.2.4. Instability

At early times before reaching the Blech equilibrium, vacancies move with drift velocity, $v_v = D/\delta_{EM}$, which causes metal to move with mass velocity, $v_m = C \cdot \Lambda \cdot D/\delta_{EM}$. The process is not smooth, however, due to several stochastic processes. Significant Brownian motion is expected over lengths of a few δ_{EM} , which causes random fluctuations in flux. There is also intrinsic instability of EM transport due to positive feedback between thermomigration and electromigration. [12,13]

Substituting (2.22) into (2.26):

$$\frac{d}{dt} C = (G - R) + \frac{d}{dx} \left[C \cdot D \cdot \left[\frac{E + H - \sigma \cdot \Lambda}{k_B \cdot T^2} \cdot \left(\frac{d}{dx} T \right) + \frac{\Lambda}{k_B \cdot T} \cdot \frac{d}{dx} \sigma - \frac{1}{\delta} \right] + a \cdot \left(\frac{d}{dx} C \right) \cdot \left(\frac{d}{dx} D \right) \right] \quad (2.37)$$

Expanding terms using subscript notation for derivatives: $Y_x = \frac{d}{dx} Y$, $Y_{xx} = \frac{d^2}{dx^2} Y$, and

removing the derivatives of diffusion coefficient (2.13):

$$C_t = (G - R) + C \cdot D \cdot \left[\begin{array}{l} \frac{\sigma_{xx} \cdot \Lambda}{k \cdot T} \dots \\ \frac{T_{xx} \cdot (E + H - \sigma \cdot \Lambda)}{k \cdot T^2} \dots \\ -2 \cdot \sigma_x \cdot \Lambda \cdot dT \\ \frac{-2 \cdot (T_x)^2 \cdot (E + H - \sigma \cdot \Lambda)}{k \cdot T^3} \dots \\ + \left(\frac{T_x}{k \cdot T^2} \right)^2 \cdot E \cdot (E + H - \sigma \cdot \Lambda) \dots \\ \frac{\sigma_x \cdot \Lambda \cdot E \cdot T_x}{k^2 \cdot T^3} \dots \\ - \frac{E \cdot T_x}{\delta \cdot k \cdot T^2} \end{array} \right] + C_x \cdot D \cdot \left[\begin{array}{l} \frac{T_x \cdot (E + H - \sigma \cdot \Lambda)}{k \cdot T^2} \dots \\ \frac{\sigma_x \cdot \Lambda}{k \cdot T} \dots \\ + \frac{-1}{\delta} \dots \\ + \left[\left(\frac{E \cdot T_x}{k \cdot T^2} \right)^2 + \frac{E \cdot T_{xx}}{k \cdot T^2} - \frac{2 \cdot E \cdot (T_x)^2}{k \cdot T^3} \right] \cdot a \end{array} \right] + C_{xx} \cdot D \cdot \frac{a \cdot E \cdot T_x}{k_B \cdot T^2} \quad (2.38)$$

The differential equation is nonlinear, so Fourier analysis is not useful, but we may gain understanding by considering sinusoidal variation in the underlying parameters:

$$T = T_0 + T_1 \sin\left(\frac{x}{b}\right) \quad C = C_0 + C_1 \sin\left(\frac{x}{b} + \theta\right) \quad \sigma = \sigma_1 \sin\left(\frac{x}{b} + \phi\right) \quad (2.39)$$

With the simplifying assumptions $\theta = \phi = 0$, the equation becomes

$$C_t = \frac{D}{b^2 \cdot k \cdot T} \left[\begin{aligned} & \sin\left(\frac{x}{b}\right)^3 \cdot C_1 \cdot \frac{T_1}{T} \cdot \sigma_1 \cdot \Lambda \dots \\ & + \sin\left(\frac{x}{b}\right)^2 \cdot \left[C_0 \sigma_1 \cdot \Lambda \cdot \frac{T_1}{T} - C_1 \cdot (E+H) \cdot \frac{T_1}{T} - C_1 \cdot \sigma_1 \cdot \Lambda \right] \dots \\ & + \sin\left(\frac{x}{b}\right) \cdot \left[-C_0 \frac{T_1}{T} \cdot (E+H) - C_0 \sigma_1 \cdot \Lambda \right] \end{aligned} \right] + (G-R) + \xi \cdot \cos\left(\frac{x}{b}\right) \quad (2.40)$$

where ξ includes the many out-of-phase terms. Positive feedback will increase the small signal variation in C , wherever C_t is in-phase:

$$\frac{C_t}{C - C_0} = \frac{C_t}{C_1 \sin\left(\frac{x}{b} + \theta\right)} > 0 \quad (2.41)$$

The resulting equation specifies the conditions necessary to sustain or increase any infinitesimal variation in the concentration.

2.2.5. Concentration Waves

For a given initial variation in concentration, (2.26) and (2.22) would predict the future concentration distribution. Figure 2.7 shows one such prediction based on an initial sinusoidal variation, clearly showing the nonlinear nature of the EM equations. Although the distribution began as a sine wave, the phase velocity of the crest is greater than that of the trough, so the leading edge steepens (something like ocean waves on the shore).

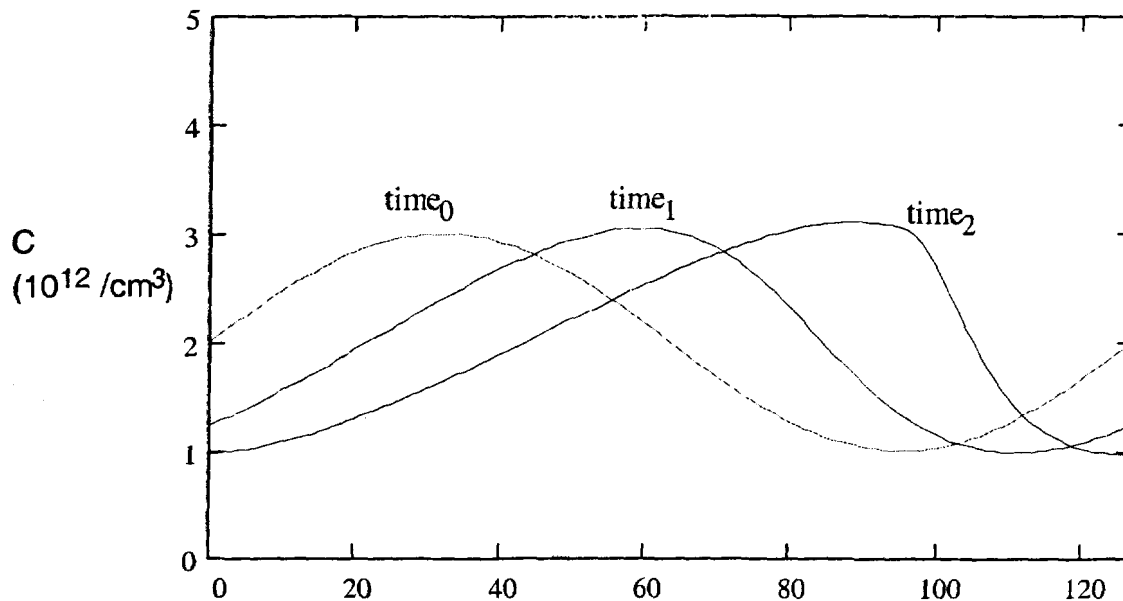


Figure 2.7: Propagation of Vacancy Concentration Wave

2.2.6. Dissipative Structures

A pattern that exists only while energy is dissipated is known as a 'dissipative structure.' One example is the Bernard instability observed when a liquid is heated from below. The updrafts and downdrafts form a somewhat stable pattern when seen from above, and the instability vanishes when the heat source is removed. Depending on power, viscosity, and depth, the pattern may be a regular hexagonal array or tubes or worms. Of many possible convection patterns, the most efficient pattern grows in magnitude – reducing the energy available to less efficient patterns, so they diminish in magnitude.

It is likely that EM stress is a breeding ground for dissipative structures, but we can only speculate what patterns might exist. In-situ microscope imaging shows that metal voids grow and shrink with no apparent rhyme or reason.[14] Since voids grow when vacancy concentration is high and evaporate (shrink) at low concentration, Figure 2.7 suggests that voids may grow as the wave crest passes and then shrink as the trough passes.

2.3. Failure

Successful interconnects maintain a low-resistance path where desired while insulating paths not desired, so failures are categorized as opens or shorts. The appropriate opens failure criteria depends on the circuit and type of line. Long lines where the $R \cdot C$ delay is dominated by line resistance could degrade product performance when electrical resistance increases by 10% or 30%, but the small resistance of short local interconnects might increase 3x or 10x without noticeable effect. During operation, resistance increases with time, so larger acceptable resistance increase will result in longer time-to-failure. In similar manner, different circuits have differing tolerance for shorts failures – some sense circuits might fail with gigohm leakage, while high fanout drivers may work fine with < 100 ohms.

Opens failures are due to growth of voids which increase resistance by forcing current through the shunt layer. Depending on processing conditions, interconnect is rarely void-free. As-produced wafers have tensile stress in the metal due to TCE mismatch and operating temperatures below the stress-free (encapsulation) temperature. Immortality to a resistance failure criteria requires either (a) an effective shunt layer to carry current around the void, or (b) a favorable void location and shape that does not open the line. Both conditions limit the maximum void size. A void forming at the cathode blocking boundary grows steadily at the rate of the net atomic drift velocity, resulting in monotonic resistance increase. Voids in mid-segment grow or shrink as they sink or source vacancies, with resulting noise in the resistance signal. The time-dependent resistance is modeled as noise superimposed on a steady resistance increase. Increased noise will result in earlier failure than the same steady increase with lower noise.

2.3.1. Time-to-Fail and Black's equation

EM times-to-failure exhibit a lognormal distribution, which is a Gaussian distribution in the logarithms of times-to-fail. The lognormal distribution is characterized by the median time-to-fail, MTF, and the variation in log-time, σ or sigma. Accelerated stress has an underlying premise that all structures experience equal acceleration, so that variation in times-to-fail (sigma) is constant, independent of use condition. The affect of stress current and temperature is predicted by Black's model, shown in equation (2.43).

This is an empirical model for median-time-to-failure, MTF, (t_{50} , or time to 50% fail), with current exponent, n , and thermal activation energy, E_a , determined by experiment. Generally, $n \cong 2$, and $E_a \sim 1$ eV.

$$\frac{MTF_1}{MTF_2} = \left(\frac{J_2}{J_1}\right)^n \cdot \exp\left[\frac{E_a}{k_B} \cdot \left(\frac{1}{T_1} - \frac{1}{T_2}\right)\right] \quad (2.42)$$

A typical acceptable failure rate is 0.1% fail thru 10^5 hours (the 10 year life of the product). Given the MTF and sigma, the time-to-0.1%fail at stress conditions, $t_{.1}$, is determined. Given the product operating temperature, Black's equation (2.41) is used to determine the current at which the acceptable failure rate is reached, known as the design rule current (discussed in Chapter Chapter 6).

2.4. Summary

Electromigration is metal motion due to vacancy dynamics, including generation/recombination, drift, and diffusion. This chapter presented the transport equations and solutions for a few simple cases. Real-life cases are extremely complicated, but simplifying approximations are useful for carefully designed test structures. The following chapter describes the design of suitable structures and the fabrication method used.

Chapter 3

Sample Preparation

The design and fabrication of the EM test structures used is presented.

3.1. Test Structure Design

Via chains were implemented with various dimensions to provide a suite of test structures. For the dual-damascene interconnect process, the upper metal level (metal-2) was selected to be critical for failure, while the lower metal level (metal-1) was made very robust with 10x the metal cross-section. Two bond pads were dedicated to each chain for high-current stress in two-terminal mode. Each chain consisted of 20 segments of the upper metal level (M2) with vias at each end (40 vias) connected to turn-around straps in the lower metal level (M1) of equivalent thickness and 10x width. Figure 3.1 is an approximate sketch of segment layout, showing that straight, parallel metal-2 segments connected with wide metal-1 straps.

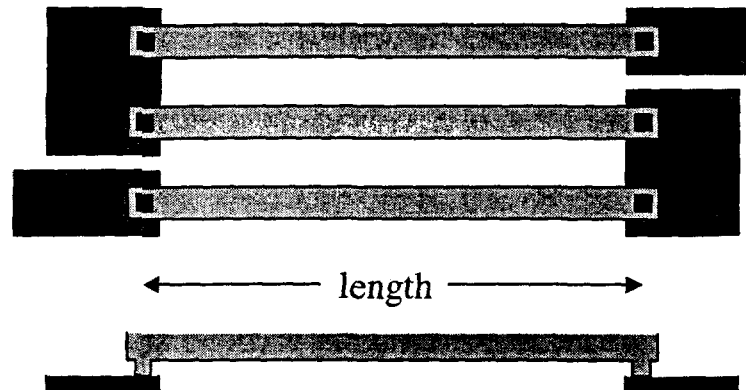


Figure 3.1: Sketch of Test Structure Layout and Cross-section

The 24 structures were organized in a matrix of two widths and six lengths in two types, as shown in Table 3.1. The variety of segment lengths covers the region of interest,

from lines much shorter than the Blech length to much longer. Two widths were provided, but only the narrower 0.5 μm structures were used in this study. Structures were available with and without adjacent lines to detect line-to-line leakage. (These extrusion monitors are sometimes referred to as ‘combs’ from single-layer serpentine/comb structures for monitoring continuity/shorts.) Examples of the two types of structures are shown in .

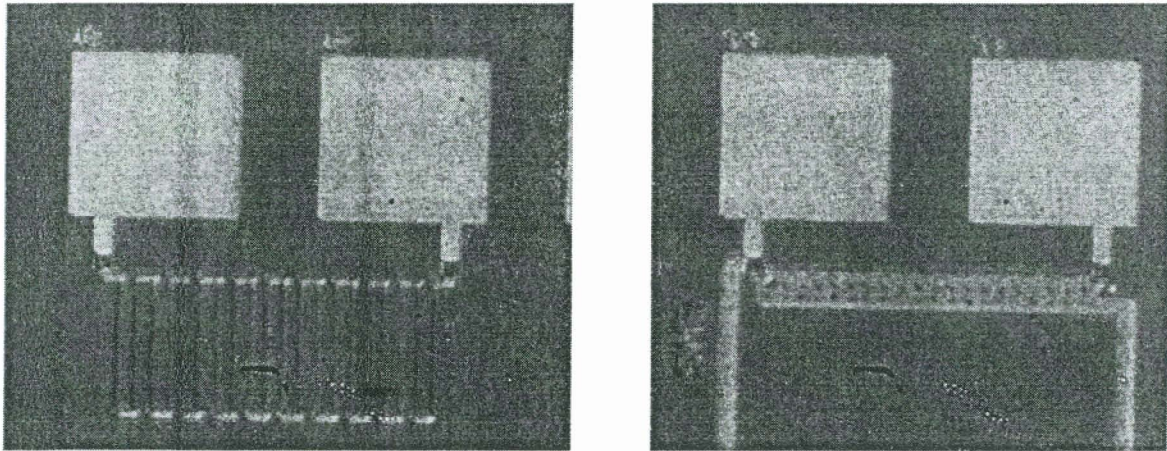


Figure 3.2: View of Finished EM Test Structures

Table 3.1: Structure Dimensions

Segment Widths	0.5 μm	0.8 μm
Thickness	0.5 μm	
Area	0.25 μm^2	0.4 μm^2

Segment Lengths
5 μm
10 μm
20 μm
40 μm
80 μm
200 μm

3.2. Wafer Processing

The damascene process involves etching a trench in dielectric, depositing a sandwich of barrier layers + electroplated copper, and polishing off the copper overfill with chemical-mechanical polishing.¹⁵ Dual-damascene is the use of vias to connect a sequence of two such damascene metal levels. The options for dual-damascene are either *via-first* or *trench-first* process modes. The process steps of the trench-first dual-damascene process used at Sharp Laboratories of America are listed in Table 3.2. As the name implies, the trenches are patterned and etched first, after which photoresist is re-applied over the topology and patterned to create via openings.

Table 3.2: Damascene Process Flow

1. bare 200mm silicon wafer
2. deposit 1.0 μm PECVD TEOS dielectric
3. etch trench
4. strip photoresist (vacuum break)
5. deposit TiN (Ti target magnetron sputtered with N plasma onto RF biased wafer)
6. deposit Cu: 20 nm strike layer (Ar-sputtered Cu target) + 1.5 μm fill (electroplate in high-acid at 2 volts)
7. Cu anneal 2 minutes at 400°C
8. abrasive-free CMP (neutral pH)
9. TiN strip in ammonia peroxide
10. blanket sputter clean
11. PECVD Silicon Nitride adhesion layer deposition
12. deposit 1.0 μm PECVD TEOS dielectric
13. trench etch
14. via etch
15. strip photoresist (vacuum break)
16. Ar sputter clean
17. repeat steps 5 through 12 for metal 2
18. etch bond pad openings
19. deposit Al layer
20. pattern and etch Al (leave Al covering pad openings)

After via etch and prior to TiN barrier deposition, oxygen exposure (vacuum break) and Ar sputter clean were efforts made to eliminate organic contamination and reduce via resistance. After deposition of TEOS over metal-2, bond-pad openings were etched and covered with sputtered aluminum patterned with the same bond-pad mask. This Al layer greatly reduced oxidation of the probe pads during hot measurements. Unprotected copper pads turned black and showed bubble delamination after just a few hours at 150°C in an oxygen-reduced atmosphere, while aluminized pads showed no visible degradation after 2000 hours at 250°C.

3.3. Copper Deposition

After sputter deposition of the Cu strike layer at Sharp Laboratories of America, copper fill was electroplated at Oregon Graduate Institute.

3.3.1. Electrolyte mixture

DI water must be added to the virgin makeup solution (VMS) to reduce the acid concentration to the final level needed in the electrolyte. The mass balance equations for concentration, C, and volume, V, are:

$$V_{\text{final}} = V_{\text{VMS}} + V_{\text{DI}} \quad C_{\text{final}} = \frac{V_{\text{VMS}} \cdot C_{\text{VMS}} + V_{\text{DI}} \cdot C_{\text{DI}}}{V_{\text{VMS}} + V_{\text{DI}}} \quad (3.1)$$

Since the concentration in DI is zero, this is simplified:

$$V_{\text{VMS}} = V_{\text{final}} \cdot \frac{C_{\text{final}}}{C_{\text{VMS}}} \quad (3.2)$$

Allowing calculation of the required volumes for the supplied and required concentrations:

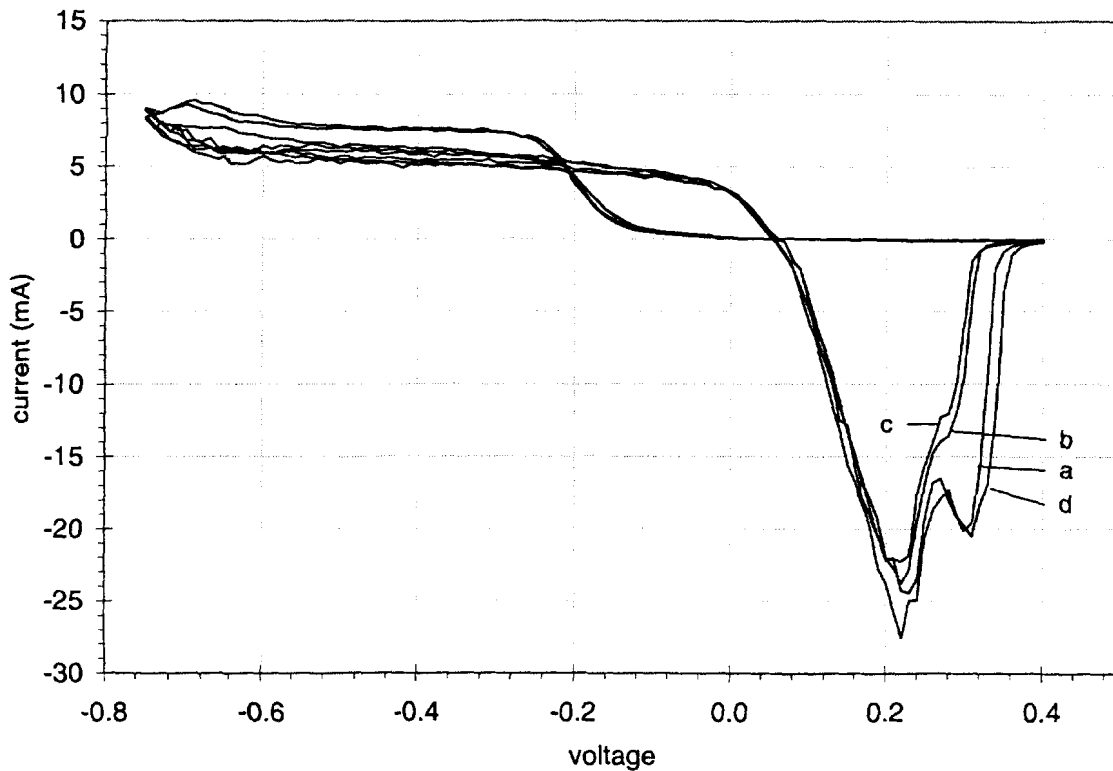
$$\begin{pmatrix} \text{final} \\ \text{VMS} \\ \text{DI} \end{pmatrix} \quad C = \begin{pmatrix} 175 \\ 240 \\ 0 \end{pmatrix} \cdot \frac{\text{gm}}{\text{liter}} \quad V = \begin{pmatrix} 10.9 \\ 7.95 \\ 2.95 \end{pmatrix} \cdot \text{liter} \quad (3.3)$$

In addition, additives are needed for optimum aspect-ratio filling capability. These additives are CUBATH® Viaform™ accelerator (sulfuric acid solution), Viaform™ Suppressor, and CUBATH® Viaform™ Leveler (non-hazardous) with the following amounts:

$$\begin{pmatrix} \text{accel} \\ \text{supp} \\ \text{lev} \end{pmatrix} \text{ add} := \begin{pmatrix} 2 \\ 8 \\ 1.5 \end{pmatrix} \frac{\text{mL}}{\text{liter}} \quad \text{add} = \begin{pmatrix} 0.2 \\ 0.8 \\ 0.15 \end{pmatrix} \% \quad \text{add} \cdot V_{\text{final}} = \begin{pmatrix} 21.8 \\ 87.2 \\ 16.35 \end{pmatrix} \text{mL} \quad (3.4)$$

3.3.2. Maintenance of electrolyte

The shelf-life of the additives is greatly reduced when mixed in the electrolyte bath. A fresh batch was mixed up for deposition of metal 1, but 40 days had elapsed before the wafers were ready for deposition of the second metal level. Degradation of the bath occurred during the interim, especially the long-chain polymer of the leveler. To characterize the bath performance and allow adding additional leveler to restore the plating characteristics, potentiostatic measurements were performed at the time of M1 deposition (Figure 3.3) and again just prior to M2 deposition (Figure 3.4). The aged and restored plating bath is very similar to the initial bath.



**Figure 3.3: Potentiostatic Measurement of Metal 1 Plating Bath
(a thru d are 4 consecutive passes)**

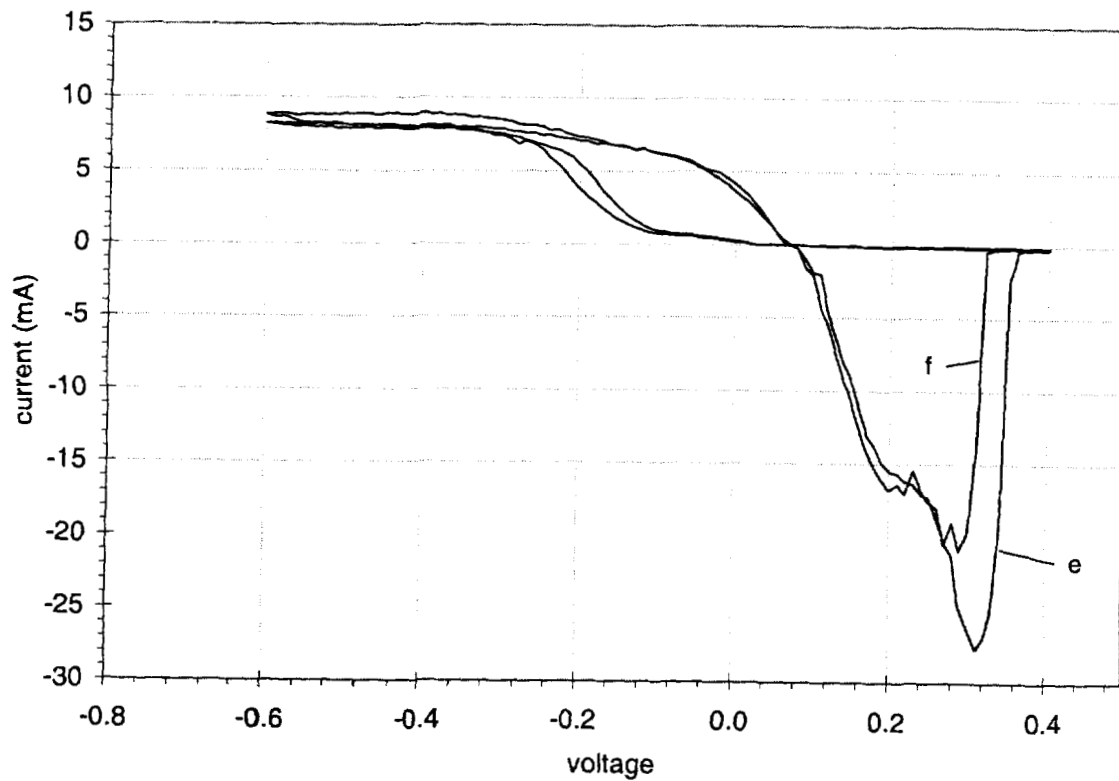


Figure 3.4: Potentiostatic Measurement of Metal 2 Plating Bath

After 5 months however, attempts to restore the bath were not as successful.

Figure 3.5 shows that while the plating curves (positive current) are qualitatively similar, the electro-etch curves (negative current) have very large variation from one pass to the next due to reduced buffering by the solution.

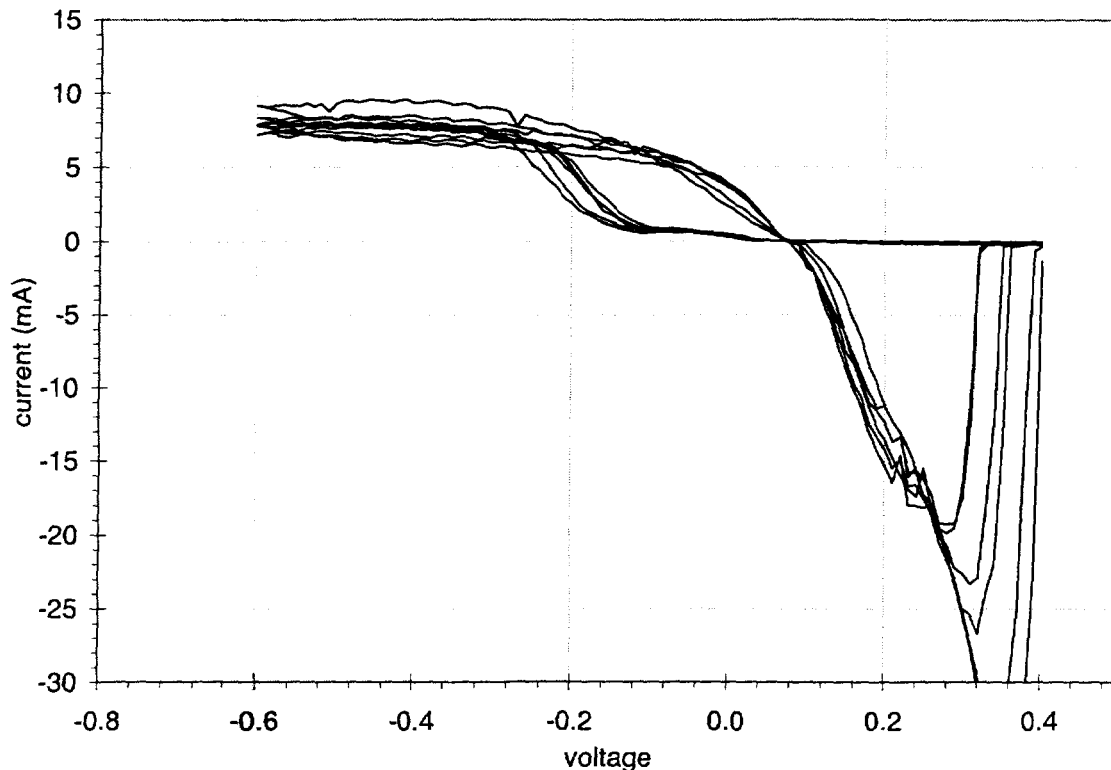


Figure 3.5: Potentiostatic Measurement after 151 days

3.3.3. Equipment automation

While the wafer is immersed, Cu is exposed to the corrosive electrolyte. When the wafer is first immersed, the thin Cu seed layer is at risk until the plating current is applied and cathodic protection stops corrosion. The timing of immersion is critical – current must start after the surface oxide is removed, but before the Cu seed layer develops pinholes. The earlier deposition system relied on manual operation of the lift motor and power supply. The operator used push button controls to lower the wafer into the electrolyte, and visually estimate the proper depth. The operator then initiated the computer-controlled ramp of plating current. Using this manual method, wafers showed defects and significant variation.

To better control the timing of these critical early steps, computer software was implemented for robotic control of the wafer lift with timing linked to the plating current power supply. Software written using the HP Vee programming application controlled

the lift stepper motor (through a serial link) and the power supply (through GPIB cable).

The improved system performs the following steps:

1. Begin slow wafer spin
2. Apply low voltage
3. Lower the wafer to prescribed depth
4. Apply each current step in succession
5. Partially raise wafer and apply fast wafer spin
6. Stop spin and completely raise wafer.

The wafer is held at an angle to the horizontal, so as the wafer is lowered, the outside radius is wet before the center. The across-wafer variation in time delay from immersion to plating is reduced by the second step. The low voltage provides cathodic protection of the seed layer but current is limited to ensure negligible plating. Electro-polishing is performed after the entire wafer is immersed, rather than accidentally during immersion.

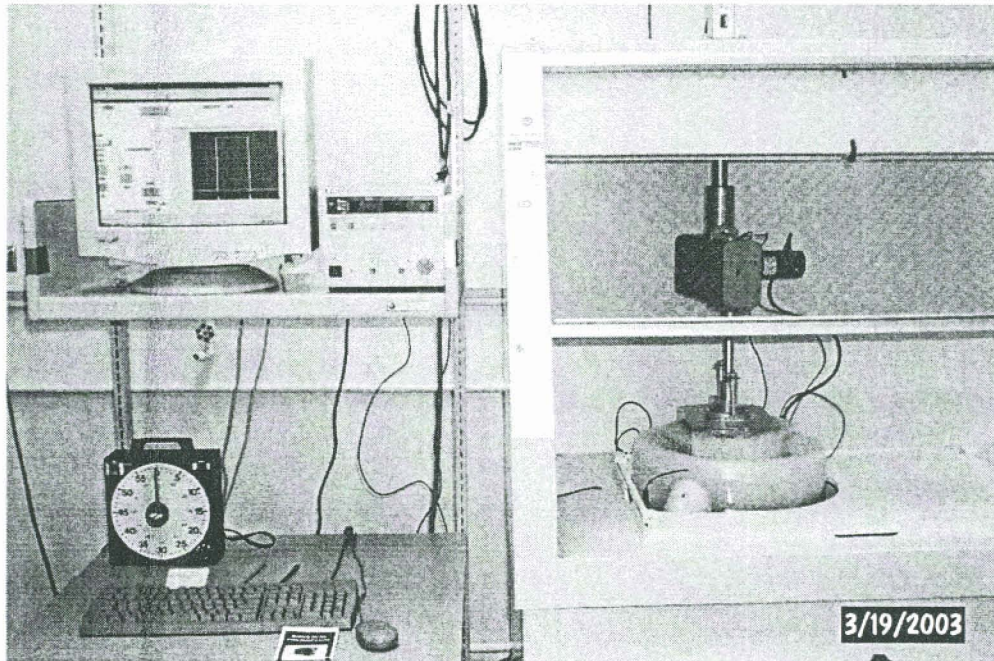


Figure 3.6: Electroplating Equipment

3.3.4. Current ramp

Uniformity of Cu thickness requires consistent deposition rate across the wafer. Ti fingers around the wafer edge provide electrical connection to the Cu layer, resulting in radial current flow and I-R drop. To avoid voltage variation across the wafer, the plating current starts small and increases as Cu becomes thicker and resistance drops. A constant voltage difference implies that current would increase exponentially with time, but in practice the current increases in discrete steps. Figure 3.7 shows the vendor's suggested current steps for their high-acid solution, compared to the current steps used to prepare samples for this study. The least-squares exponential fit shows that both initial current and exponential time constant are reduced for lower I-R drop and improved uniformity.

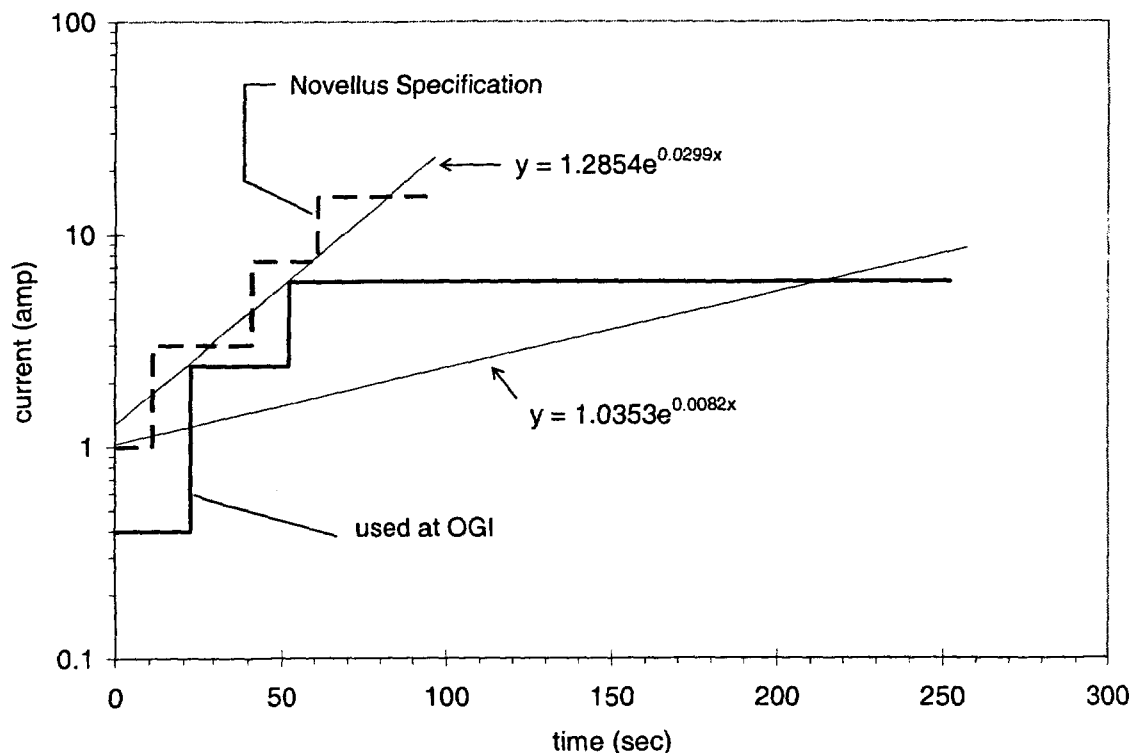


Figure 3.7: Plating Current Steps

3.3.5. Thickness and Resistivity

The resistance was measured by 4-point probe at room temperature (25°C). The sheet resistance calculated using the equation for an infinite sheet on an insulating bottom boundary is:

$$\frac{V}{I} = 2.5 \text{ m}\Omega, \rho_s = \frac{\pi}{\ln(2)} \cdot \frac{V}{I}, \rho_s = 11.3 \text{ m}\Omega/ \quad (3.5)$$

The total charge exchanged with the electrolyte is

$$Q := \left[\begin{array}{c} 22 \\ 30 \\ 200 \end{array} \right] \cdot \text{sec} \cdot \left[\begin{array}{c} 0.401 \\ 2.41 \\ 6.019 \end{array} \right] \cdot \text{amp} \quad \text{area} := \frac{\pi}{4} \cdot (200 \text{ mm})^2 \quad \frac{Q}{\text{area}} = 4.09 \cdot \frac{\text{coul}}{\text{cm}^2} \quad (3.6)$$

The thickness expected from mass balance equation is

$$\text{thick} := \frac{Q}{\text{area}} \cdot \frac{1 \cdot \text{mole}}{2 \cdot \text{eq}} \quad \text{thick} = 1.503 \cdot \mu \quad (3.7)$$

The bulk resistivity from measured sheet resistance and calculated thickness is very close to the bulk value.

$$\rho_s \cdot \text{thick} = 1.703 \cdot \mu\Omega \cdot \text{cm} \quad \rho_s \cdot \text{thick} = 95.7\% \cdot \rho_{\text{bulk}}$$

By peeling off an area of the Cu film, the resulting step was measured with a Dektak IIA profilometer made by Sloan Corp. The measured step height implies a much thicker film, resulting in a much higher estimate for bulk resistivity.

$$\text{thick} := \frac{1.8 \text{ cm}}{\left(\frac{6.25 \text{ cm}}{8 \text{ div}} \right)} \cdot 10000 \frac{\text{A}}{\text{div}} \quad \text{thick} = 2.304 \cdot \mu \quad (3.8)$$

$$\rho_s \cdot \text{thick} = 2.611 \cdot \mu\Omega \cdot \text{cm} \quad \rho_s \cdot \text{thick} = 146.7\% \cdot \rho_{\text{bulk}}$$

It is possible that either the Dektak measurement is in error or ECD creates a porous metal layer with lower density but the same quantity of metal. The most likely explanation is that tearing the Cu film distorts the metal and delaminates Cu from the substrate near the step. These effects exaggerate the step height, reducing the usefulness of profile measurements. Cleaving at LN2 temperatures might reduce the distortion.

Annealed and patterned samples were cleaved and imaged in cross section, as shown in Figure 3.8. Both metal layers were patterned with equal dimensions, and both showed thickness of $0.38 \mu\text{m}$ and width of $0.72 \mu\text{m}$. The section area of $0.27 \mu\text{m}^2$ agrees with the electrical cross-section shown on page 50 (within $\pm 15\%$ accuracy of SEM images).

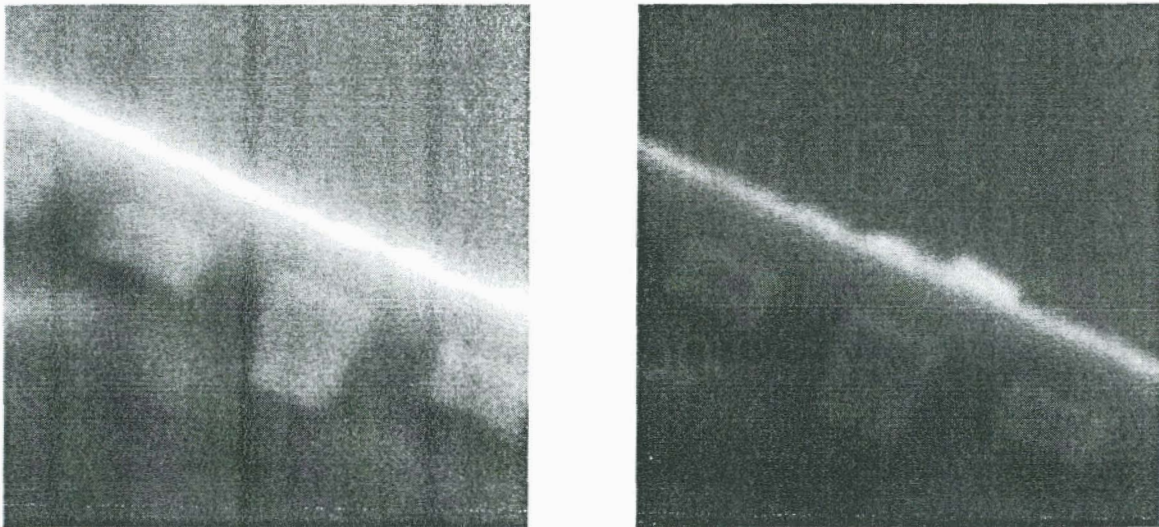


Figure 3.8: SEM Cross Sections of Patterned Lines

3.4. Summary

Via chain structures with various segment lengths were designed and fabricated. These structures are optimized to allow EM measurements which apply to the theory in the previous chapter. The measurement conditions cause Joule selfheating, which greatly affects measured performance. The following chapter describes the method used to carefully account for selfheating.

Chapter 4

Self-heat Characterization

The strong thermal dependence of the EM mechanism requires precise knowledge of absolute metal temperature as well as thermal gradients in the test structure. A 10°C hotter test structure will degrade 1.66x faster (200°C to 210°C at 1 eV). Even if the average temperature is known, the temperature gradient affects both the Blech equilibrium and the time dependence to reach equilibrium. Heat generation and heat-flow in the test structures were modeled and characterized to achieve the necessary thermal accuracy.

4.1. Heat Flow Model

Thermal conduction (heat transfer by diffusion) is predicted by the conduction rate equation (Fourier's law):

$$\vec{q} = -k \cdot \nabla T \quad (4.1)$$

where \vec{q} is the vector heat flux, T is temperature, ∇ is the three-dimensional del operator, and thermal conductivity, k , is a material property. The continuity equation allows derivation of the heat equation, also known as the heat diffusion equation. [16]

$$\frac{\partial}{\partial x} \left(k \frac{\partial T}{\partial x} \right) + \frac{\partial}{\partial y} \left(k \frac{\partial T}{\partial y} \right) + \frac{\partial}{\partial z} \left(k \frac{\partial T}{\partial z} \right) + q = d \cdot c_p \cdot \frac{\partial T}{\partial t} \quad (4.2)$$

for mass density, d , and specific heat, c_p ($d \cdot c_p$ is termed the volumetric heat capacity), and q is the heat generation rate (power) per unit volume. For constant thermal conductivity, the heat equation is

$$\nabla^2 T + \frac{q}{k} = \frac{1}{\alpha} \frac{\partial T}{\partial t}, \quad \alpha = \frac{k}{d \cdot c_p} \quad (4.3)$$

where α is the thermal diffusivity.

The time dependence of the heat equation is not of interest in typical use. For typical microelectronic geometries, the exponential time constant for temperature response is a few microseconds (μs). DC measurements allow adequate time to reach equilibrium, and gigahertz frequencies in modern microprocessors are so fast that RMS current can be treated as DC. Thus, we are only interested in the steady-state condition ($\partial T/\partial t=0$).

4.1.1. Metal-Dielectric Stack

For the region of interconnects far from vias, heat dissipated in the metal is conducted down through the dielectric (along the z axis) to the silicon substrate. If fringing is ignored, the 1-dimensional conduction rate equation in the dielectric is

$$q \cdot h = -k_{\text{diel}} \frac{dT}{dz} \quad \Delta T_{\text{diel}} = \frac{q \cdot h \cdot t}{k_{\text{diel}}} \quad (4.4)$$

where h is metal thickness, and t is oxide thickness. In the metal:

$$q \cdot (h - z) = -k_{\text{metal}} \frac{dT}{dz} \quad \Delta T_{\text{metal}} = \frac{q \cdot h^2}{2 \cdot k_{\text{metal}}} \quad (4.5)$$

Metal conducts heat much better than the dielectric, so the temperature gradient in the metal is much less than that in dielectric at the same heat flux. Thermal conductivities are listed in Table 4.1 for typical metals and dielectrics. The ratio of $k_{\text{Cu}}/k_{\text{silica}} > 200$, so for $t \geq h$, ΔT_{metal} is relatively insignificant.

Table 4.1: Typical Thermal Conductivities (in W/m²K)

Metal	k_{metal}	Dielectric	k_{diel}
Copper	401	Porous silica	$\ll 1$
Aluminum	237	Fluorinated silica	< 1
Tungsten	174	Silica, Quartz	1.48
Titanium	21.9	Sapphire	41.

The total temperature rise is:

$$\Delta T = \Delta T_{\text{diel}} = T_{\text{metal}} - T_j = \frac{q \cdot h \cdot t}{k_{\text{diel}}} = q \cdot h \cdot \theta_{\text{mj}} \cdot w \cdot L \quad (4.6)$$

where the 'junction temperature', T_j , is the temperature of the silicon substrate, and the thermal resistance, θ_{mj} , is defined as

$$\theta_{mj} = \frac{\Delta T}{q \cdot \text{volume}} = \frac{t}{k_{\text{diel}} \cdot w \cdot L \cdot \delta} \quad (4.7)$$

where L is the length of the structure, and the correction factor, $\delta=1$ for this simple 1-D case. The wafer-level experiments in this investigation were performed with controlled wafer temperature. Package-level measurements must account for the additional temperature change from wafer to ambient by adding an additional θ_{jc} .

4.1.2. 2-D Isolated Infinite Line

An infinitely long isolated metal line has no thermal gradient along its' length, so it reduces to a 2-dimensional problem. Figure 4.1 shows a typical rectangular metal cross-section surrounded by insulator some distance above the silicon wafer (heat sink). The left edge is a symmetric boundary (reflection) while the right side extends to infinity (in the y-axis). The direction of heat flow is indicated by fringing field lines.

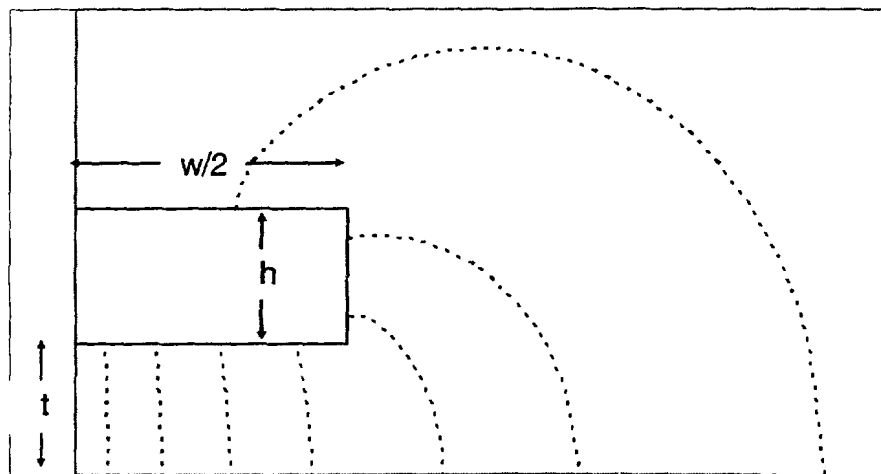


Figure 4.1: 2-D Fringing of Heat Flow

To account for fringing, several approximate steady-state solutions have been reported which generally reduce thermal resistance by increasing correction factor δ . [17,18] (For $w < h$, $\delta > 5$.)

$$\text{if } w > t/2, \delta = 1 + \frac{h}{w} \left[\frac{2 \cdot \pi}{\ln \left[1 + \frac{2 \cdot h}{t} \cdot \left(1 + \sqrt{1 + \frac{t}{h}} \right) \right]} - \frac{t}{2 \cdot h} \right] \quad (4.8)$$

$$\text{if } w < t/2, \delta = 1 + \frac{h}{w} \left[\frac{\pi \cdot \left(1 - \frac{0.0543 t}{2 \cdot h} \right)}{\ln \left[1 + \frac{2 \cdot h}{t} \cdot \left(1 + \sqrt{1 + \frac{t}{h}} \right) \right]} + 1.47 \right] \quad (4.9)$$

The EM test structures described earlier have several finite segments in proximity to each other. The neighboring lines compress the lateral fringing by imposing a reflection boundary condition on the right side of Figure 4.1. Reduction of the fringing field increases thermal resistance by reducing δ to less than that of (4.8) or (4.9) (but in any case $\delta > 1$ as the 1-D limit).

4.1.3. End Straps

Since the lines are finite and not perfectly terminated, there will be non-uniform temperature along the x axis. For longitudinal heat flow with metal and the dielectric in parallel, the dielectric is ignored. Similarly, for vertical heat flow through the metal and dielectric in series, temperature change from top to bottom of metal is ignored. Heat flow can be modeled in 2-dimensions:

$$\frac{d}{dx} \left(k_{\text{metal}} \cdot \frac{d}{dx} T \right) + q = \frac{k_{\text{diel}}}{h} \cdot \left(\frac{d}{dz} T \right) \quad (4.10)$$

$$k_{\text{metal}} \cdot \frac{d^2}{dx^2} T + q = \frac{k_{\text{diel}}}{h} \cdot \frac{T - T_{\text{substrate}}}{t}$$

For uniform heat dissipation, solve by change of variable:

$$\frac{d^2}{dx^2} T = \frac{k_{\text{diel}}}{k_{\text{metal}}} \cdot \frac{T - T_{\text{substrate}}}{h \cdot t} = \frac{q}{k_{\text{metal}}} = \frac{T - T_{\text{substrate}}}{\lambda l^2} = \beta \quad \lambda l = \sqrt{h \cdot t \cdot \frac{k_{\text{metal}}}{k_{\text{diel}}}} \quad (4.11)$$

For which the general solution is:

$$\frac{d^2}{dx^2} \beta = \frac{1}{\lambda l^2} \cdot \frac{d^2}{dx^2} T = \beta \quad \beta(x) = \beta_0 \cdot \exp\left(\frac{x}{\lambda l}\right) + \beta_1 \cdot \exp\left(\frac{-x}{\lambda l}\right) \quad (4.12)$$

$$T - T_{\text{substrate}} = \left(\frac{q}{k_{\text{metal}}} + \beta \right) \cdot \lambda l^2 = \Delta T + \left(\beta_0 \cdot \exp\left(\frac{x}{\lambda l}\right) + \beta_1 \cdot \exp\left(\frac{-x}{\lambda l}\right) \right) \cdot \lambda l^2 \quad (4.13)$$

where ΔT is given by (4.6). Boundary conditions are applied to derive the complete solution. The structure is symmetric about $x=0$, with vias at both ends $x = \pm L/2$. From symmetry, both ends are at the same temperature, so

$$T_{\text{via}} = T\left(\frac{L}{2}\right) = T_{\text{substrate}} = T\left(-\frac{L}{2}\right) = T_{\text{substrate}} \quad (4.14)$$

Substitution of (4.13) into (4.14) shows that for symmetric boundary conditions, $\beta_0 = \beta_1$, so (4.13) becomes:

$$T - T_{\text{substrate}} = \Delta T + 2\beta_0 \cosh\left(\frac{x}{\lambda_1}\right) \cdot \lambda_1^2 \quad \frac{dT}{dx} = 2\beta_0 \sinh\left(\frac{x}{\lambda_1}\right) \cdot \lambda_1 \quad (4.15)$$

Heat conduction through the via at each end of the line is in parallel with conduction through the dielectric. The thermal resistance of the via alone, θ_{via} , is defined:

$$T\left(\frac{L}{2}\right) - T_{\text{substrate}} = \theta_{\text{via}} \cdot Q = \theta_{\text{via}} \cdot \left[Q_{\text{via}} - w \cdot h \cdot k_{\text{metal}} \cdot \left(\frac{dT}{dx}\right)_{x=\frac{L}{2}} \right] \quad (4.16)$$

The heat flow through the via, Q , is determined by the energy balance between the energy dissipated in the via, Q_{via} , and the heat conducted down the metal line. Substituting (4.15) into (4.16) and solving:

$$\beta_0 = \frac{1}{2 \cdot \lambda_1} \cdot \frac{\theta_{\text{via}} \cdot Q_{\text{via}} - \Delta T}{\lambda_1 \cdot \cosh\left(\frac{L}{2 \cdot \lambda_1}\right) + \lambda_2 \cdot \sinh\left(\frac{L}{2 \cdot \lambda_1}\right)} \quad (4.17)$$

where we define another characteristic thermal length:

$$\lambda_2 = \theta_{\text{via}} \cdot w \cdot h \cdot k_{\text{metal}} \quad (4.18)$$

(4.15) becomes

$$T - T_{\text{substrate}} = \Delta T + \frac{(\theta_{\text{via}} \cdot Q_{\text{via}} - \Delta T) \cdot \cosh\left(\frac{x}{\lambda_1}\right)}{\cosh\left(\frac{L}{2 \cdot \lambda_1}\right) + \frac{\lambda_2}{\lambda_1} \cdot \sinh\left(\frac{L}{2 \cdot \lambda_1}\right)} = \Delta T + (\theta_{\text{via}} \cdot Q_{\text{via}} - \Delta T) \cdot \Omega \quad (4.19)$$

which determines the temperature profile based on the 'via fraction number', $\Omega(x)$, given in (4.20). This is restated using dimensionless parameters in (4.21) which is plotted in Figure 4.2.

$$\Omega = \frac{\cosh\left(\frac{x}{\lambda 1}\right)}{\cosh\left(\frac{L}{2\lambda 1}\right) + \frac{\lambda 2}{\lambda 1} \sinh\left(\frac{L}{2\lambda 1}\right)} = \omega\left(\frac{x}{L}, \frac{L}{\lambda 1}, \frac{\lambda 2}{\lambda 1}\right) \quad (4.20)$$

$$\omega(p, r, s) := \frac{\cosh(p \cdot r)}{\cosh\left(\frac{r}{2}\right) + s \cdot \sinh\left(\frac{r}{2}\right)} \quad (4.21)$$

As the via fraction number, Ω or ω , increases from 0 to 1 the metal temperature scales linearly from the 1-D metal temperature to the via temperature. To help understand (4.19) and (4.20) it is useful to consider some special cases:

1. For $\Omega = 0$, the temperature simplifies to the 1-D metal temperature: $T - T_j = \Delta T$. This happens far from vias ($x \ll L$) with $L \gg \lambda 1$.
2. For $\Omega = 1$, the temperature is dominated by the via: $T - T_j = \theta_{via} \cdot Q_{via}$. This happens near the vias: $x = \pm L/2$ (for $\lambda 1 \gg \lambda 2$) and everywhere for the shortest structures ($L \ll \lambda 1$).
3. For ideal termination, $\theta_{via} \cdot Q_{via} = \Delta T$ (by coincidence), temperature is uniform everywhere.

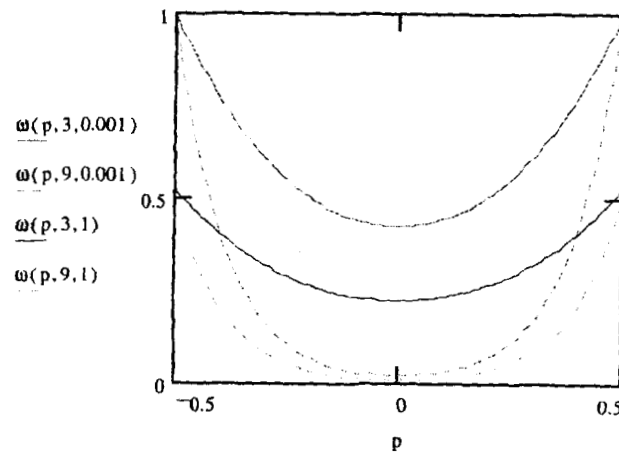


Figure 4.2: Via Fraction Parameter

It is useful to derive the average temperature along the segment.

$$T_{\text{avg}} - T_{\text{substrate}} = \frac{1}{L} \int_{-\frac{L}{2}}^{\frac{L}{2}} \left[\Delta T + (\theta_{\text{via}} \cdot Q_{\text{via}} - \Delta T) \cdot \Omega \right] dx = \Delta T + (\theta_{\text{via}} \cdot Q_{\text{via}} - \Delta T) \cdot \eta \quad (4.22)$$

where the 'average via fraction' parameter is

$$\eta = \frac{1}{L} \int_{-\frac{L}{2}}^{\frac{L}{2}} \Omega dx = \int_{-\frac{1}{2}}^{\frac{1}{2}} \omega(p, r, s) dp = N\left(\frac{L}{2\lambda_1}, \frac{\lambda_2}{\lambda_1}\right) \quad N(r, s) := \frac{\sinh(r)}{r \cdot (\cosh(r) + s \cdot \sinh(r))} \quad (4.23)$$

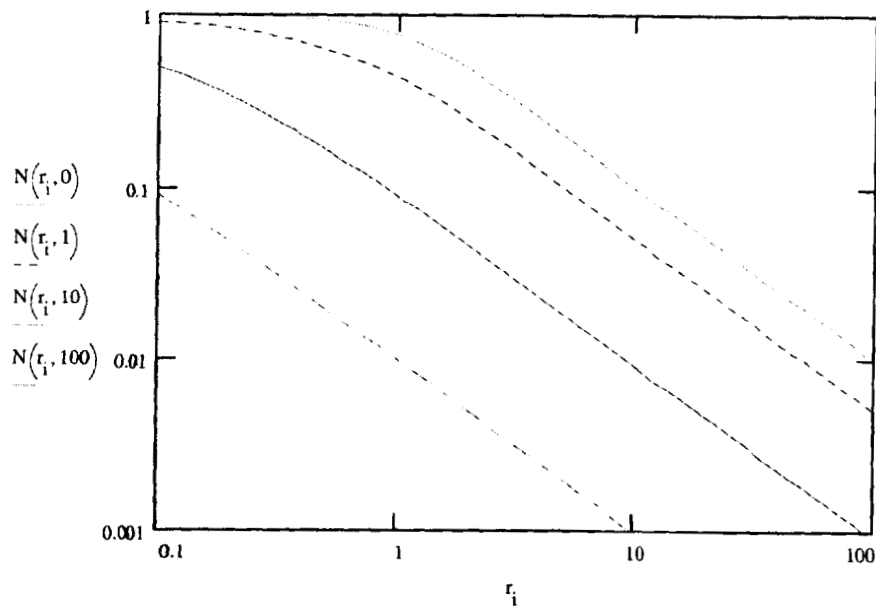


Figure 4.3: Average Via Fraction Parameter

As before, short structures with $L \ll \lambda_1$ have $N = 1$, so average temperature is dominated by via temperature. For long structures or weak coupling ($\lambda_1 \gg \lambda_2$), $N = 0$, and the average temperature simplifies to the 1-D metal temperature.

4.2. Joule Heating

Electrical current through a resistance dissipates electrical power as heat in the process known as 'Joule Heating'.

$$q = J^2 \cdot \rho \quad (4.24)$$

Electrical resistance is a consequence of electron scattering. Matthiesen's Rule assumes that different scattering sites are independent (much like an ideal gas), so the scattering rates are simply additive. It follows that total resistivity is the sum of resistivities caused by each scattering mechanism. These mechanisms exhibit various temperature dependences:

- Impurity concentration is constant with temperature, and the resulting electron scattering causes the low temperature 'residual resistivity' after other mechanisms are frozen out.
- Above the Debye temperature, phonon concentration (and phonon scattering) is about proportional to temperature.
- Interstitial or vacancy self-defects scatter electrons like impurities, but equilibrium concentration is negligible at typical temperature and stress. Extreme conditions cause a resistivity contribution, which remains if quenched to a super-saturated condition.

The resistivity, ρ , of metals increases (nearly) linearly with temperature, and measured data is fit to the linear equation (4.25). The material properties are slope (b) and intercept (either zero-resistance temperature, T_0 , or zero-temperature resistance, ρ_0).

Figure 4.4 shows the electrical resistivity of several metals.¹⁹

$$\rho = \rho_0 + b \cdot T = b \cdot (T - T_0) \quad (4.25)$$

The equation is modified as follows to define the 'thermal coefficient of resistance', TCR or α :

$$\rho = \rho_1 \cdot [1 + \alpha_1 \cdot (T - T_1)] \quad (4.26)$$

The TCR is commonly mistaken for a material constant, but as (4.26) shows, a change in the reference temperature, T_1 , will affect the reported TCR. Unless otherwise noted, all TCR values in this dissertation are based on a reference temperature of 25°C.

$$b = \rho_1 \cdot \alpha_1 = \rho_2 \cdot \alpha_2 \quad \frac{1}{\alpha_2} - T_2 = \frac{1}{\alpha_1} - T_1 \quad (4.27)$$

The best-fit parameters over the range 273°K < T < 600°K for the data in Figure 4.4 are listed in Table 4.2, which also shows the good correlation between x-intercept, T_0 , and the Debye temperature.

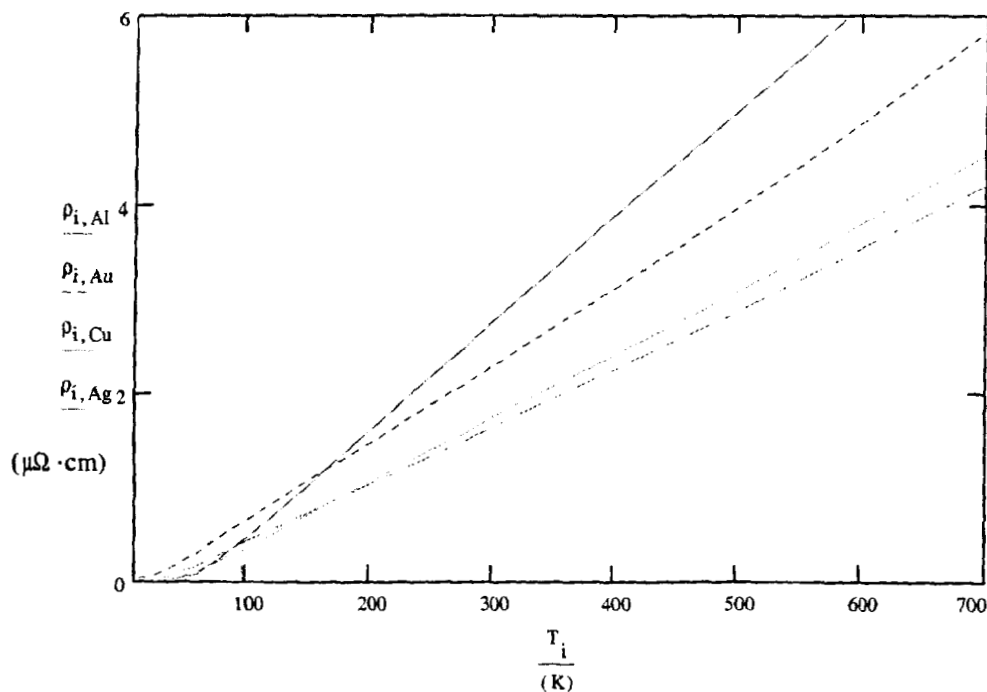


Figure 4.4: Resistivity Vs Temperature

Table 4.2: Temperature Coefficient of Resistance for Representative Metals

Metal	Slope, b ($\mu\Omega\cdot\text{cm}/^\circ\text{K}$)	x-intercept, T_0 (K)	Debye Temp (K)	Resistivity, ρ_1 ($\mu\Omega\cdot\text{cm}$)	TCR, α_1 (K^{-1})
Al	0.01133	59.1	394	2.709	0.00418
Cu	0.00686	48.8	315	1.712	0.00401
Ag	0.00629	41.1	215	1.617	0.00389
Au	0.00859	36.0	170	2.255	0.00381
at 25°C					

4.2.1. Via Resistance (Schottky Barrier)

For vias, the damascene metal layers are in series, and the interfaces between layers are significant. In addition to the simple resistances of the via current path, the transition from metal types introduces a Schottky barrier. The idealized current-voltage

dependence of thermionic-field emission through that barrier is given by (4.28).[20] Low-current measurements are dominated by this exponential dependence, while the series resistance dominates high-current measurements, especially when avalanche runaway reduces the thermionic emission voltage.

$$I = I_{sl} \left[\exp \left[\frac{q}{k_B \cdot T} \cdot \left(\frac{V}{n} - \phi_B \right) \right] - 1 \right] \quad (4.29)$$

The temperature dependence of via resistance is rather complicated, making analysis of data very difficult.

4.2.2. Thermal Runaway

From (4.24), power, q , increases with resistivity, (4.25) shows that resistivity increases with temperature, and (4.4) indicates that temperature increases with power. In constant current conditions, this feedback mechanism will cause a runaway condition where temperature increases until the structure opens by melting or vaporization.

From (4.5), (4.9) and (4.28), the temperature rise is:

$$T_{rise} = I^2 \cdot R \cdot \theta \quad (4.30)$$

Substituting (4.25) gives the transcendental form of the temperature equation:

$$T_{rise} = I^2 \cdot \left[R_1 \cdot \left[1 + \alpha_1 \cdot (T_{sub} + T_{rise} - T_1) \right] \right] \cdot \theta \quad (4.31)$$

Solving explicitly for temperature rise:

$$T_{rise} = \frac{\frac{1}{\alpha_1} - T_1 + T_{sub}}{\frac{1}{I^2 \cdot R_1 \cdot \alpha_1 \cdot \theta} - 1} \quad (4.32)$$

The temperature reaches infinity when the denominator goes to zero at runaway current, I_{run} :

$$I_{run} = \frac{1}{\sqrt{R_1 \cdot \alpha_1 \cdot \theta}} \quad (4.33)$$

Re-writing (4.31) in terms of I_{run} :

$$T_{\text{rise}} = \frac{\frac{1}{\alpha_l} - T_l + T_{\text{sub}}}{\left(\frac{I_{\text{run}}}{I}\right)^2 - 1} \quad (4.34)$$

Figure 4.5 shows the rapid increase in temperature as I approaches I_{run} , and compares it to temperature rise assuming constant resistance (zero TCR, or $\alpha = 0$). The comparison suggests that a constant-resistance approximation is accurate for self-heating temperatures below 50°C ($J/J_{\text{run}} < 0.3$).

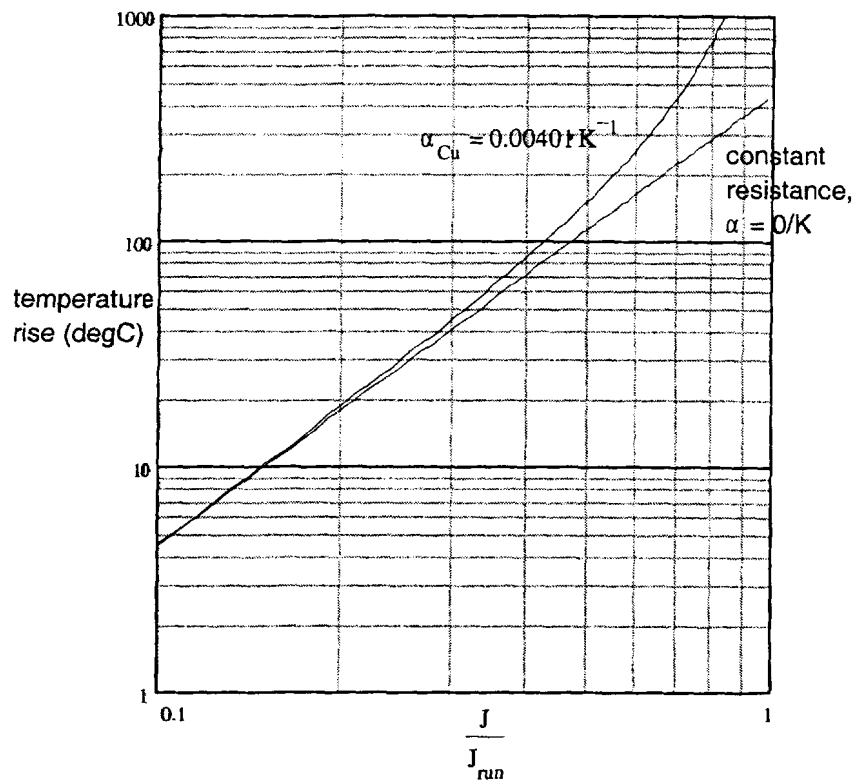


Figure 4.5: Temperature Rise and Thermal Runaway

The thermal model for these structures assumes the power dissipation, q , is uniform throughout the test structure. When temperature variation exceeds about 50°C , the resistance variation causes a corresponding power variation. The accuracy of the constant power approximation is improved by averaging q over a small distance. The dependence of temperature at a point, x , depends on local q

4.3. Regression

The total electrical resistance of the EM test structure includes the narrow metal lines, wide metal straps, and the via barrier. For the 20 segment via chain structure,

$$R_{\text{total}} = \frac{V}{I} = 20 \cdot (2 \cdot R_{\text{via}} + R_{\text{line}}) \quad (4.35)$$

Including the temperature-dependence of both via and line resistances:

$$\frac{V}{I} = 40 \cdot R_{v0} \cdot [1 + \alpha_v \cdot (T_{\text{via}} - T_1)] + 20 \cdot L \cdot R' \cdot [1 + \alpha_1 \cdot (T_{\text{avg}} - T_1)] \quad (4.36)$$

Substituting the via (4.14) and average (4.22) temperatures:

$$\frac{V}{I} = 40 \cdot R_{v0} \cdot (1 + \alpha_v \cdot T_{\text{via}}) + 20 \cdot L \cdot R' \cdot [1 + \alpha_2 \cdot [P' \cdot \theta' \cdot (1 - \eta) + P_{\text{via}} \cdot \theta_{\text{via}} \cdot \eta + T_a]] \quad (4.37)$$

Expansion of this equation produces 6 additive terms, which can be written as a dot product of two matrices, shown in (4.37). In this matrix notation, $R = A \cdot B$, where A includes all unknown terms, and B specifies the known or measured parameters.

$$\frac{V}{I} = \begin{bmatrix} R_{v0} \\ R_{v0} \cdot \alpha_v \\ R' \\ R' \cdot \alpha_2 \cdot \theta' \\ R' \cdot \alpha_2 \cdot \theta_{\text{via}} \\ R' \cdot \alpha_2 \end{bmatrix} \cdot \begin{bmatrix} 40 \\ 40 \cdot T_{\text{via}} \\ 20 \cdot L \\ 20 \cdot P_{\text{line}} \cdot (1 - \eta) \\ 20 \cdot L \cdot P_{\text{via}} \cdot \eta \\ 20 \cdot L \cdot T_a \end{bmatrix} \quad (4.38)$$

The solution of linear equations requires 6 equations in 6 unknowns, which by linear algebra can be determined by 6 measurements at various currents, voltages, and segment lengths. The experiments would create square matrices B and R (6 rows by 6 columns - one column per measurement). The solution of n equations in n unknowns is accomplished by matrix inversion: $A = R \cdot B^{-1}$. (While simple in notation, the calculation is numerically intensive.)[21]

In practice, it is difficult to choose truly independent conditions, and it is preferable to perform a regression fit to more than 6 measurements. A least-squares solution of an overdetermined system (n equations in m unknowns) is performed by a

similar operation on rectangular matrices: $A = R \setminus B$. The individual parameters are then extracted from the elements of A . The units of the B matrix are length L in μ , power P in watts, and temperature T in $^{\circ}\text{K}$, so units of the parameters extracted from the A matrix are:

$$\begin{aligned} R_{\text{via}} &:= A_1 \cdot (\text{ohm}) & R' &:= A_3 \cdot \left(\frac{\text{ohm}}{\mu\text{m}} \right) \\ \alpha_{\text{via}} &:= \frac{A_2}{A_1} \cdot \left(\frac{1}{\text{K}} \right) & \alpha_2 &:= \frac{A_6}{A_3} \cdot \left(\frac{1}{\text{K}} \right) \\ \theta_{\text{via}} &:= \frac{A_5}{A_6} \cdot \left(\frac{\text{K}}{\text{watt}} \right) & \theta' &:= \frac{A_4}{A_6} \cdot \left(\frac{\text{K} \cdot \mu\text{m}}{\text{watt}} \right) \end{aligned} \quad (4.39)$$

The B matrix in (4.37) includes quantities that are unknown prior to the experiment. T_{via} , P_{via} , and eta (η) must be estimated beforehand, and compared to parameters extracted after regression fit. This results in an iterative method which quickly converges to a self-consistent solution.

4.3.1. Measurements

Temperature rise of the various structures was characterized by voltage/current measurements on a heated wafer chuck at various wafer temperatures. Initially, current and substrate temperature were ramped from low to high values, but annealing at the high temperatures caused permanent resistance change, so the low temperature measurements could not be repeated. To reduce microstructure changes and make the best comparison between high and low temperatures, temperature and current were ramped down (from highest to lowest).

The difference between the observed and predicted data is called the *residual*. Better models with better goodness of fit will result in smaller residuals, and examination of correlation between residuals and parameters can lead to an improved model.

A regression fit was performed to understand the temperature dependence of via resistance. The self-heat model was modified to set $\alpha_v = 0$, which is equivalent to deleting the second rows of the matrices in (4.37), leaving 5 rows. The iterative regression was performed separately for each wafer temperature, the residuals were

analyzed to estimate the mean and RMS variation. Since the residuals are in units of ohms, the mean residual is simply R_{via} , which is plotted in Figure 4.6 along with \pm confidence intervals corresponding to the RMS variation in the residual.

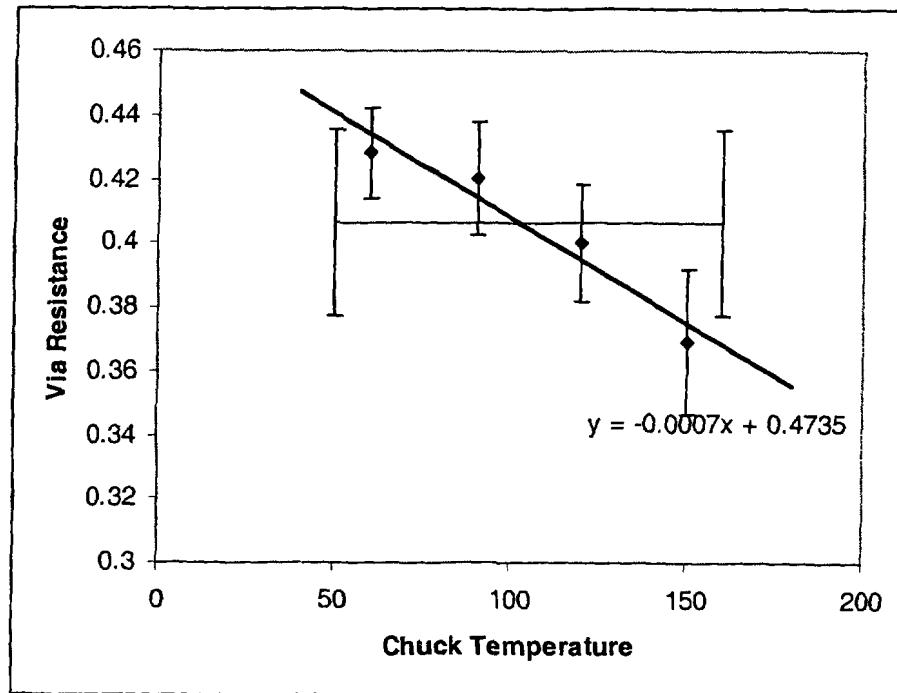


Figure 4.6: Temperature Dependence of Via Resistance

The horizontal line in Figure 4.6 is the mean R_{via} for the combined data from all wafer temperatures, and confidence intervals are repeated at each end. Some of the observed residual is die-to-die variation, but the confidence intervals of each temperature point are smaller than that of the overall data, because the improved model can predict more of the variation. The linear fit shows a negative TCR which is consistent with the confidence intervals. The temperature trend is apparently concave down suggesting the influence of nonlinear conduction from (4.27), but the curvature is not statistically significant.

4.3.2. Parameters

As an example, characterization of die J16 showed:

$$\begin{aligned}
 R_{\text{via}} &= 0.483 \cdot \text{ohm} \\
 \alpha_{\text{via}} &= -0.001 \cdot \text{K}^{-1} \\
 R' &= 59.8 \cdot \frac{\text{ohm}}{\text{mm}} \\
 \alpha_2 &= 0.00334 \text{K}^{-1} \\
 \theta' &= 0.562 \cdot \frac{\text{K} \cdot \text{m}}{\text{watt}} \\
 \theta_{\text{via}} &= 0.01 \cdot \frac{\text{K}}{\mu \text{ watt}}
 \end{aligned}$$

$$\begin{aligned}
 \text{area} &:= \frac{\rho}{R'} \quad \text{area} = 0.286 \cdot \mu \text{m}^2 \\
 \begin{pmatrix} \lambda 1 \\ \lambda 2 \end{pmatrix} &:= \begin{pmatrix} \sqrt{\text{cmet} \cdot \text{area} \cdot \theta'} \\ \text{cmet} \cdot \text{area} \cdot \theta_{\text{via}} \end{pmatrix} \\
 \begin{pmatrix} \lambda 1 \\ \lambda 2 \end{pmatrix} &= \begin{pmatrix} 8.029 \\ 1.128 \end{pmatrix} \cdot \mu \cdot \text{m}
 \end{aligned}$$

$$A := \begin{bmatrix} 0.4827 \\ -0.0007 \\ 0.0598 \\ 112.3 \\ 1.966 \\ 0.0002 \end{bmatrix}$$

As expected, vias are cooler than mid-segment, and shorter lines are cooler than longer lines, as shown by the example in Figure 4.7.

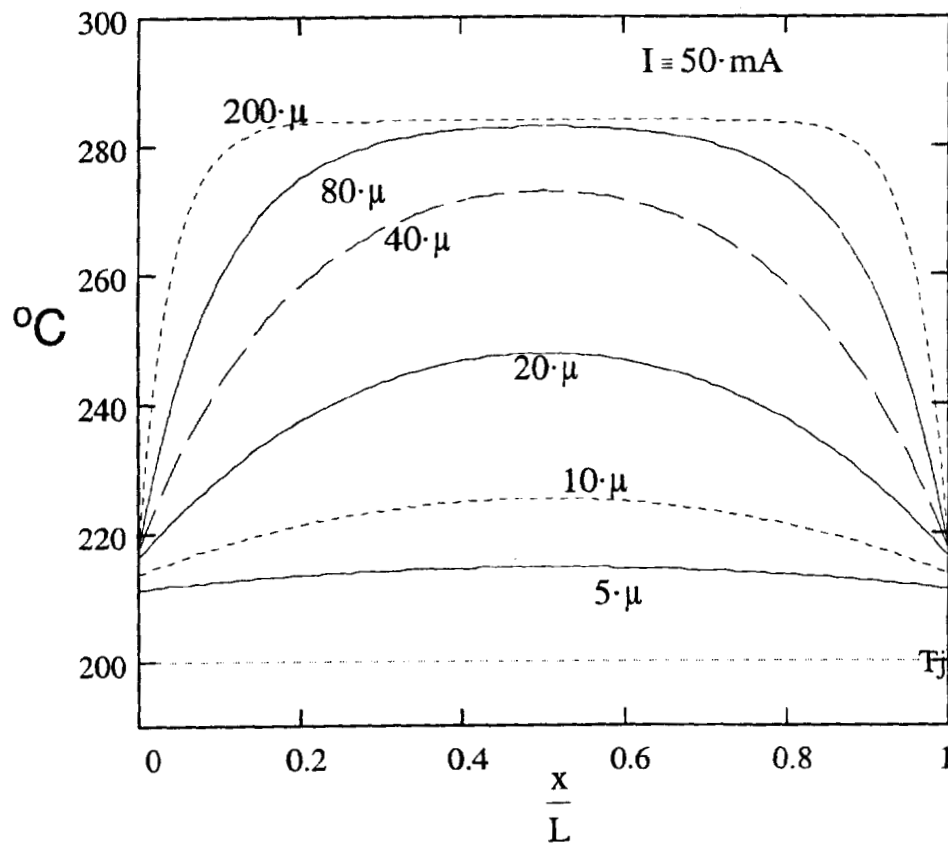


Figure 4.7: Measured Temperature Profile

4.4. Summary

Temperature and thermal gradient have a strong effect on EM, so considerable care was taken to model the heat flow and characterize test structures. The analytical equations allow use of matrix methods with some iteration to measure the important parameters and develop a complete understanding of average temperature and spot temperature along the length of each test structure. EM measurements at carefully controlled metal temperature are the subject of the next chapter.

Chapter 5

Electromigration Stress Measurements

Accelerated stress conditions and selected results are shown. Various methods are used to apply elevated currents and temperatures.

5.1. Comparison of Package and Wafer-level Methods

The choice of electrical connection greatly affects the thermal anomalies that accelerated stress must account for. Packaged structures on printed-circuit boards in an oven ambient allow simultaneous stress of many structures, but temperature control suffers due to variability of air flow velocity. Probing a flat wafer resting on a heated metal chuck improves temperature control but structure quantity is limited by available probes. To achieve high throughput with few probes, it is tempting to use highly accelerated conditions at high current, but this causes significant selfheating as discussed in the previous chapter. This tradeoff between accuracy, sample size, throughput, and cost can be optimized separately for wafer-level and package-level stress. Simultaneous stress of 24 packaged samples typically requires about 24 days (including 10 days setup time and 14 days in the oven). Wafer stress with one sample at a time can provide equivalent throughput at 24x higher acceleration (one day per sample). We shall show that the higher acceleration is not a problem, and for monitor purposes, wafer stress provides early feedback with a partial sample size.[22]

5.2. Wafer Probe Stress System

The wafer-level stress system used for this study consists of a Micromanipulator Model 6400 probe station with Model HSM heated 6-inch wafer chuck inside a Model 7000-LTE enclosure. Control and measurement of current and voltage is performed by a Hewlett-Packard Model 4145B Semiconductor Parameter Analyzer with GPIB control by a personal computer running 'Interactive Characterization Software' (ICS) by Metrics Technology.

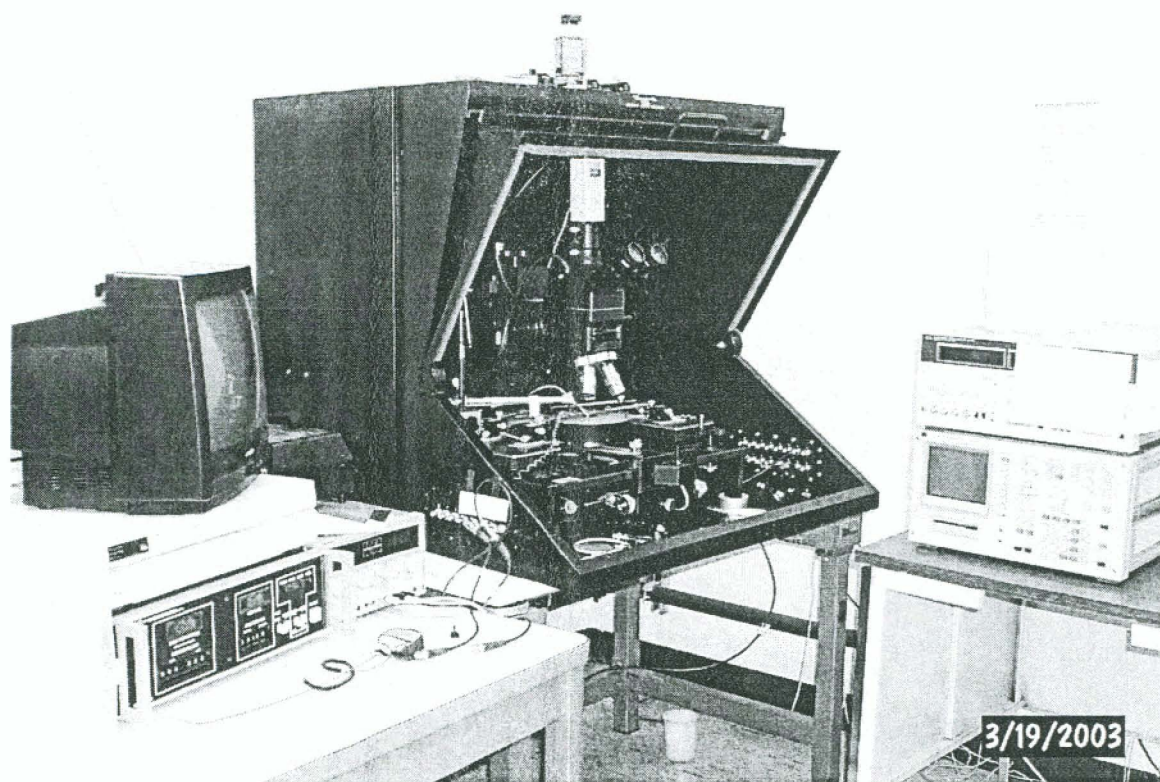


Figure 5.1: Wafer Stress System Setup

5.3. EM Stress Conditions

Black's equation (2.41) predicts the acceleration factor at given metal temperature and current, as shown in Figure 5.2. The maximum allowable current at normal use conditions is a 'design rule' during product layout, and is shown on the 1x acceleration curve in Figure 5.2.

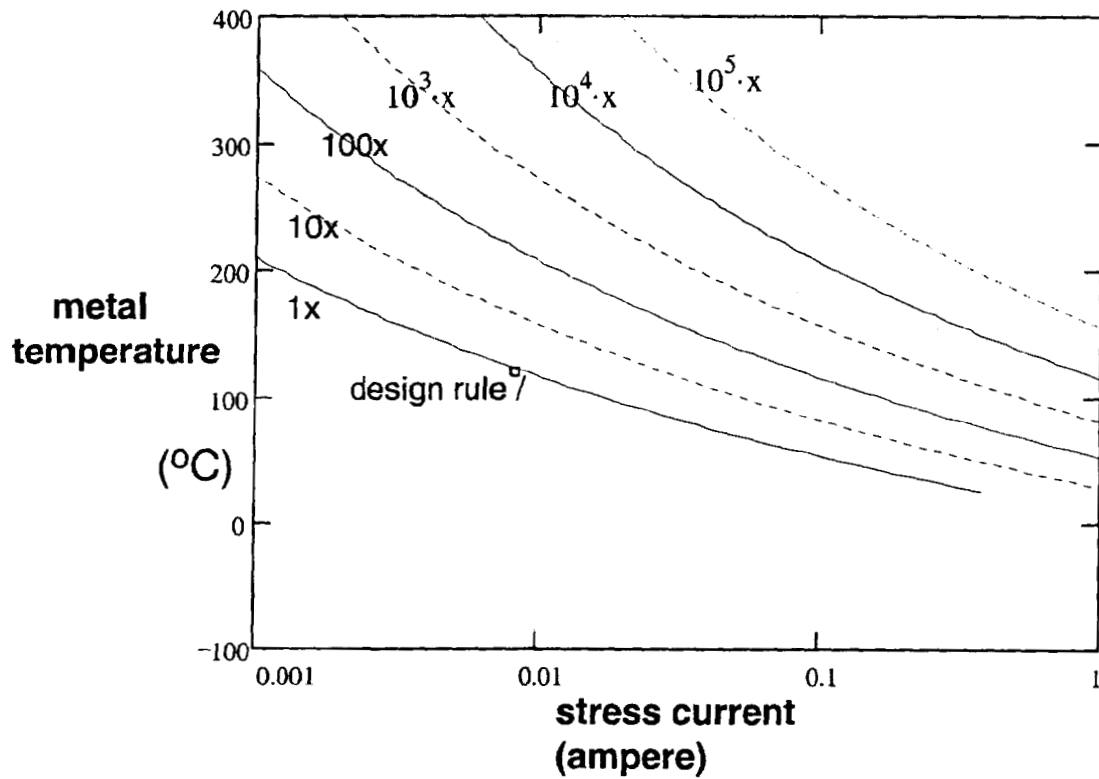


Figure 5.2: EM Iso-acceleration Curves for T_m

Metal temperature is the sum of ambient (wafer) temperature and selfheating from the stress current. Acceleration factor for given wafer temperatures are shown vs current in Figure 5.3. The curves are affected within 10x of the thermal runaway current, I_{run} . Selfheating has not been significant at design rule conditions, but will be for upper metal levels of future technologies.

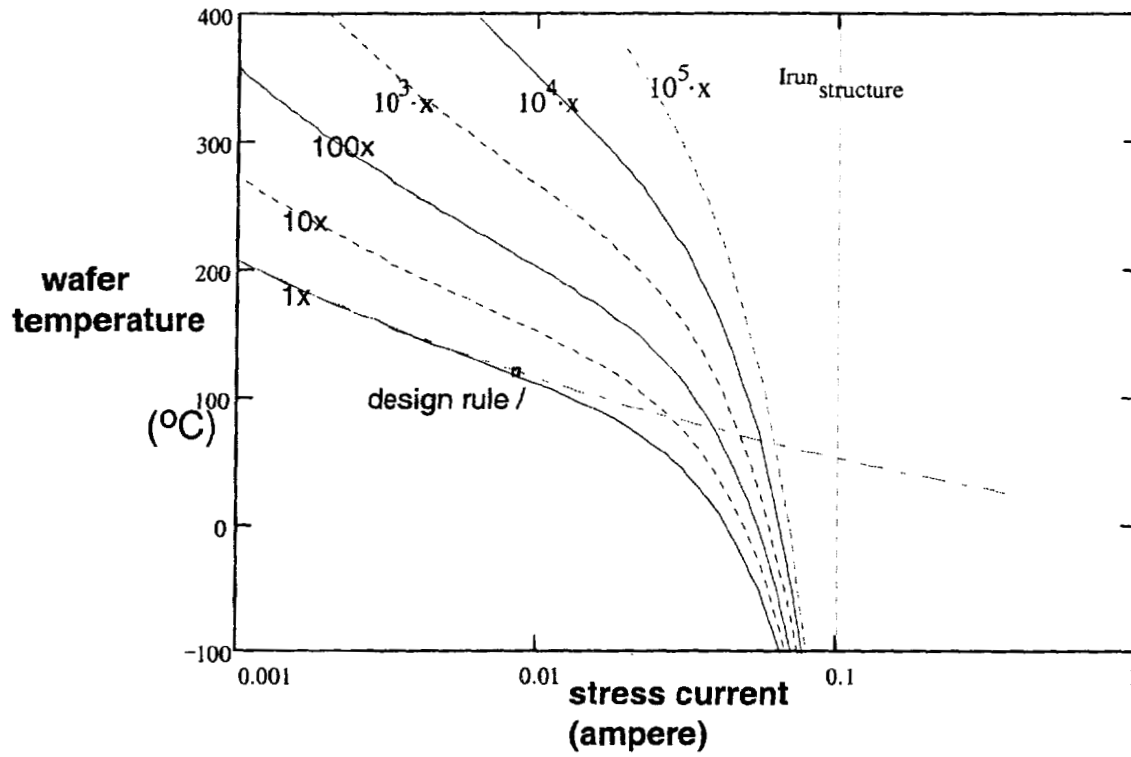


Figure 5.3: EM Iso-acceleration Curves for T_j

There is no useful acceleration at currents below the EM threshold, I_{min} . When the EM threshold for this structure is considered, the constant acceleration curves have a 'pole' at I_{min} as they have a 'zero' at I_{run} , as shown in Figure 5.4. Between I_{min} and I_{run} the region of current with uncorrected Black's dependence is vanishingly small. The maximum possible acceleration is limited by oxidation of the probe pad at high wafer temperature and catastrophic thermal instability at high current.

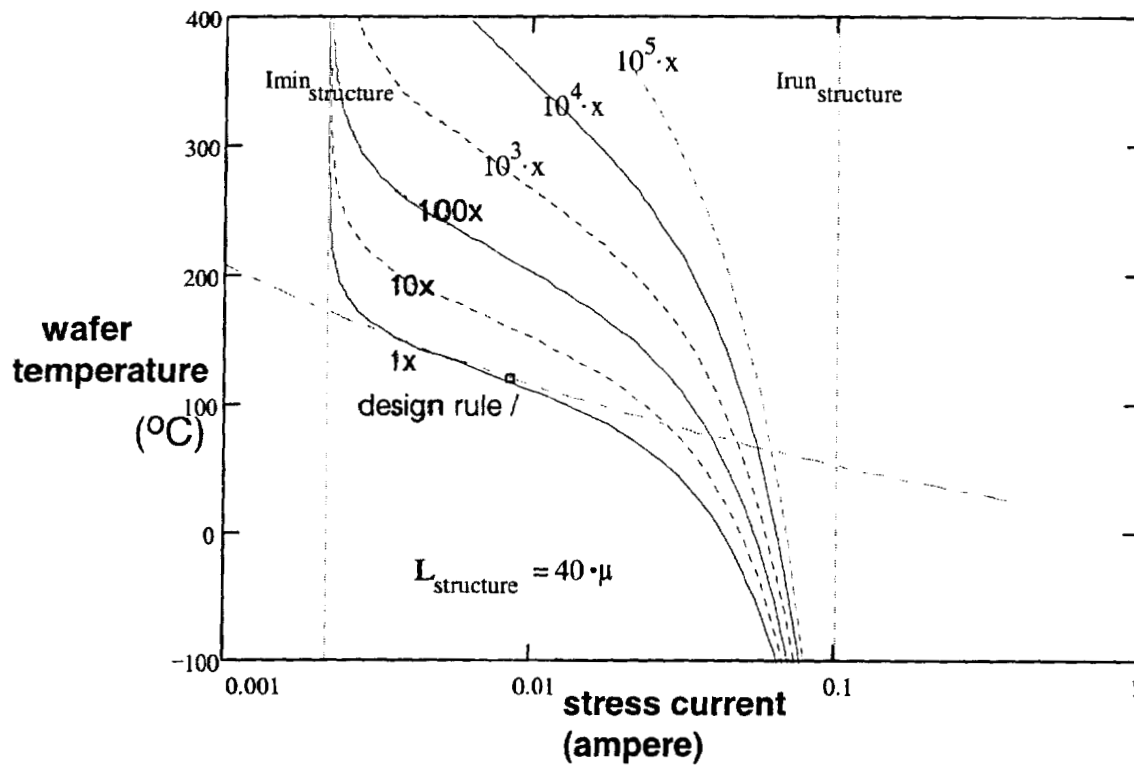


Figure 5.4: EM Iso-acceleration Curves Above Threshold

The useable temperature and current ranges define a 'window' or 'envelope' of appropriate stress conditions. To complete an 11 year useful life (10^5 hours) in a few days measurement (10^2 hours) requires 1000x acceleration. At reasonable wafer temperatures, stress conditions are outside the range of the uncorrected Black's dependence, and represent a rather narrow range of current.

The threshold and runaway currents vary with structure lengths, so by measuring a family of structures with varying lengths, the several small stress windows can be combined into one larger effective window.

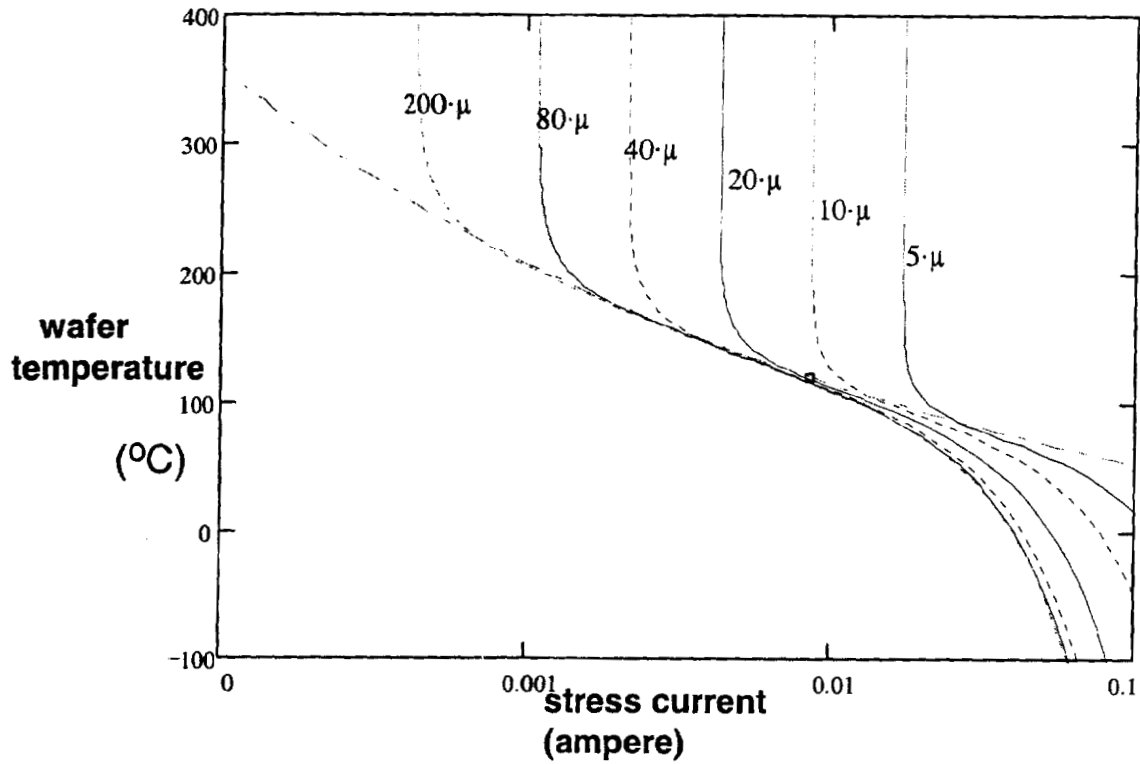


Figure 5.5: EM Iso-acceleration for Various Length Structures

5.3.1. Stress Envelope

Measurements of time-to-failure were performed at various constant current stress levels for various line lengths. Figure 5.6 shows the stress conditions to achieve 100x acceleration, and the actual temperature/current conditions used.

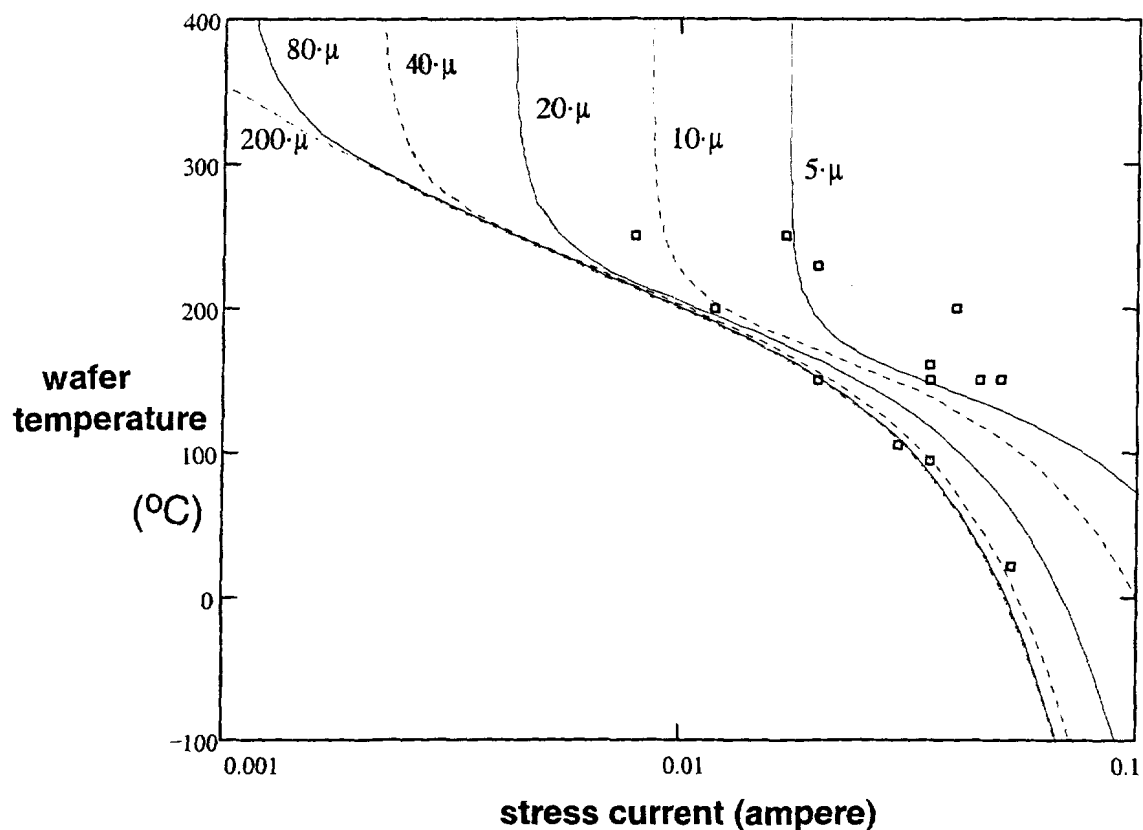


Figure 5.6: EM Stress Conditions and Stress Limits

Each condition was applied to the appropriate structures as not all stress conditions are useful for all structures. The selection of accelerating condition depends on the structure. Short segments change quickly, but require high current to exceed threshold. Long segments fail quickly without backpressure to slow void growth, and thermal gradients at each end affect a smaller fraction of the total segment.

5.4. Time-dependent Resistance Increase

The initial resistance increases with length, so the applied voltage should be proportional to current times length, as shown in Figure 5.7. As metal voids coalesce and grow during the accelerated stress, the stress current is forced into the higher resistance shunt layers, which increases the total resistance of the test structure. Depending on the void location, this process can be smooth or erratic. A void which forms directly at the cathode can only grow monotonically larger. Figure 5.8 shows resistance change typical

for a cathode void, along with an exponential curve fit. (The graph actually shows voltage at a constant current stress.)

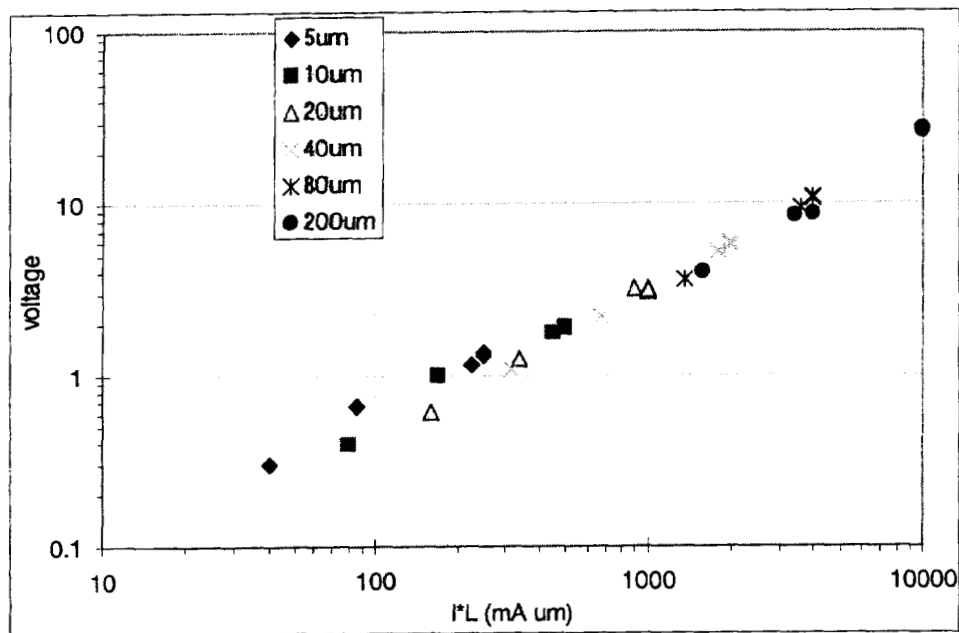


Figure 5.7: Initial Stress Voltage at Start of EM Stress

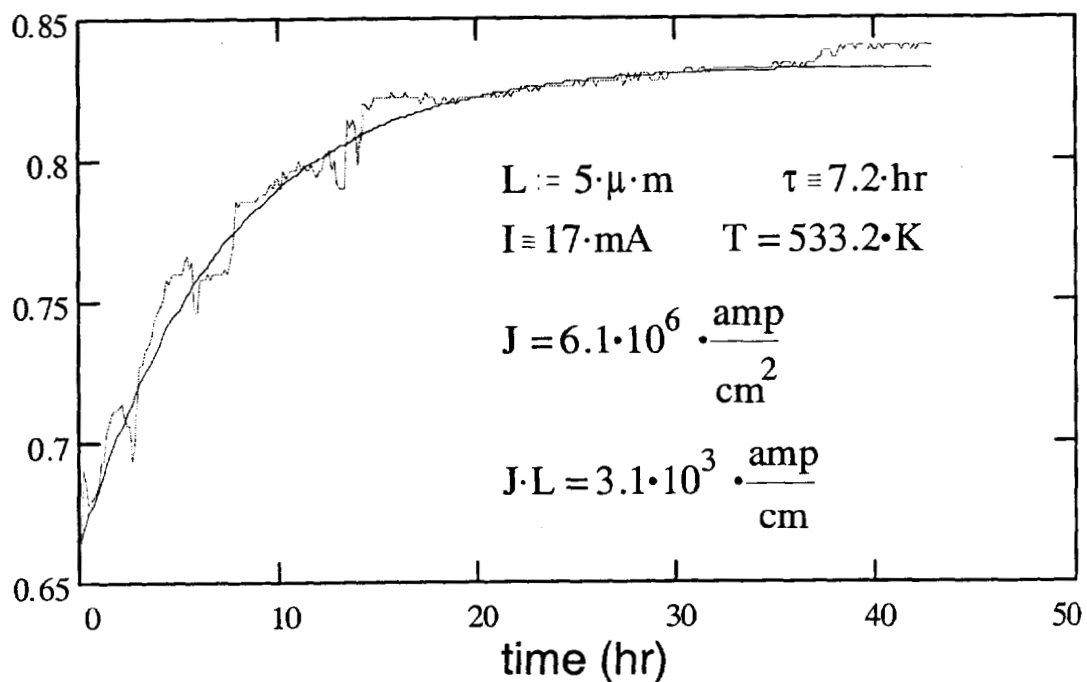


Figure 5.8: Resistance vs Time Near Threshold

Voids which form in the **ling** segment away from the ends grow and shrink erratically. The change in size is **due** to the net imbalance between arriving and departing vacancies. The resistance vs time is erratic, as shown in the examples in Figure 5.9.

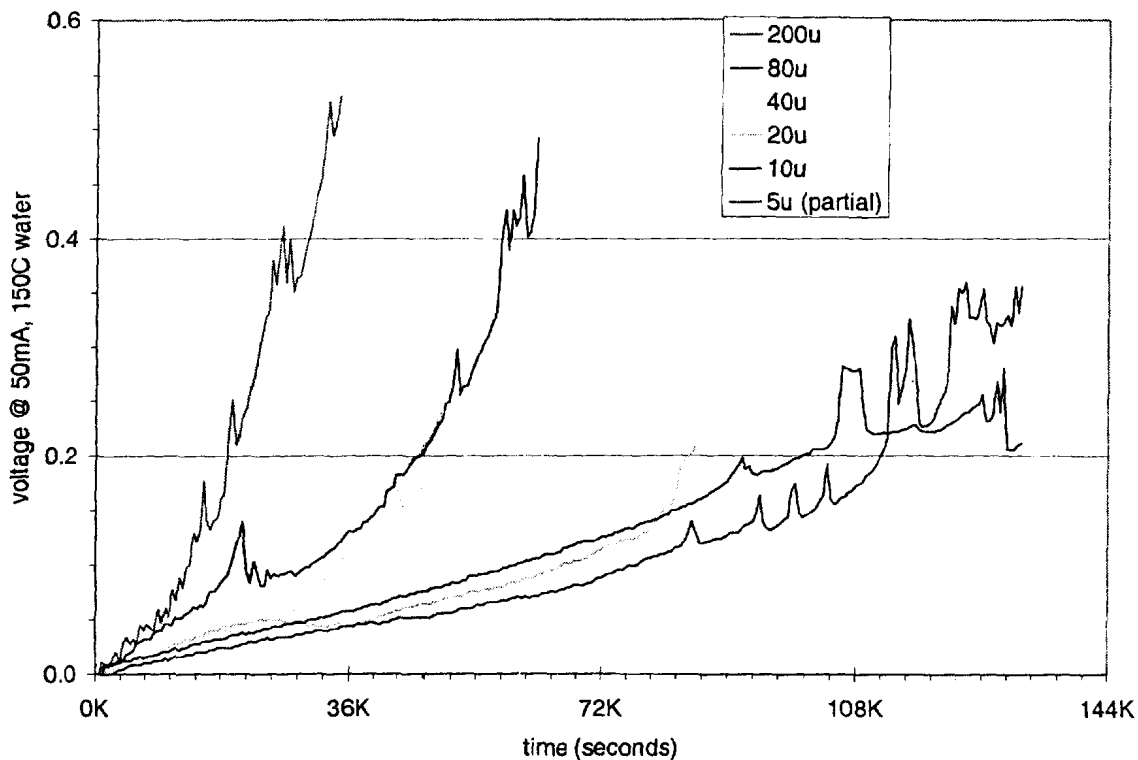


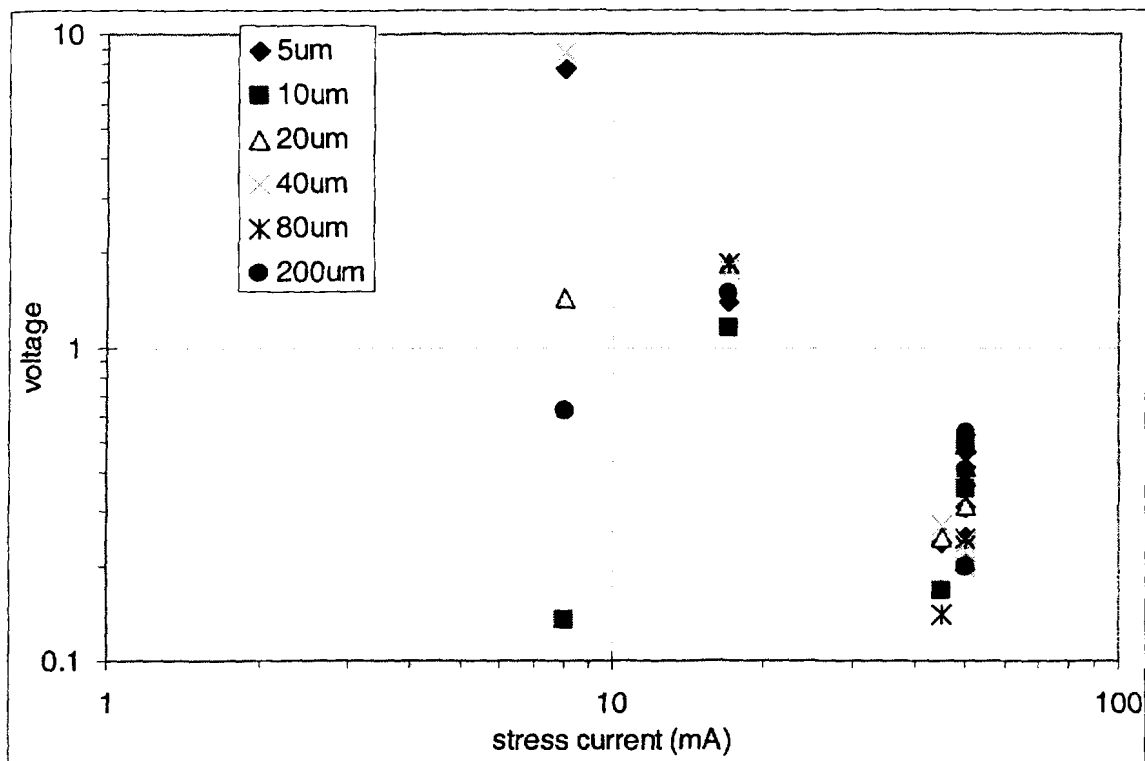
Figure 5.9: Resistance vs Time Near Runaway

Generally, voids occur at several locations along each segment. At near threshold conditions, the region near the cathode experiences the greatest tensile stress which favors growth of voids near the cathode. At well above threshold, the large instabilities in stress can favor the growth of one void for a time and then favor another. The few observations of saturation in resistance vs time occurred near threshold conditions, and at conditions well above threshold resistance was generally more erratic. For conditions below threshold, the resistance increase saturates as the void grows to its maximum size. [23,24,25]

High currents increase the temperature of the line, especially at voids. These local high-resistance regions become hot spots and reach thermal runaway long before the rest of the line. At lower currents, voids can grow larger with corresponding greater resistance

increase before reaching thermal runaway. Figure 5.10 shows that at high current, the stress voltage increases only slightly before a sudden opens failure, while at lower current some structures show very large resistance increase. To observe void growth, it is better to use lower stress currents with margin to runaway.

This contributes to the J^2 dependence in Black's equation. Void growth rate is proportional to current, while the void size at catastrophic open failure is inversely proportional to current, so the failure time = size/rate is inversely proportional to the square of current.



This value is about 7 orders of magnitude larger than single crystal bulk samples (Figure 2.4), which might be expected from the presence of grain boundaries and other surfaces. This effective D_v corresponds to diffusion of vacancies as if motion were uniform thru the metal cross section. Of many diffusion paths, reports suggest that the interface between copper and the overlying silicon nitride dominates over diffusion thru the lattice, grain boundaries, and other interfaces.[26,27] Assuming all diffusion occurs at the metal top surface, the time constant in (2.33) can be explained with a surface diffusion coefficient, requiring simulations to determine the effect of the characteristic length δ_{EM} in the correction factor τ_e .

5.6. Time to Failure Distributions

With a failure criteria of 30% resistance increase, the times-to-failure exhibit a lognormal distribution, as shown in Figure 5.11 (after eliminating various outliers). This distribution applies to conditions well above threshold, where samples are certain to fail at various times. At stress conditions near the threshold, some samples are immortal while less robust samples will fail – leading to a bi-modal distribution. Incorrectly fitting a bimodal distribution with a single Gaussian distribution will overestimate the dispersion. Such a large sigma would imply worse EM performance, but actual performance is much improved.

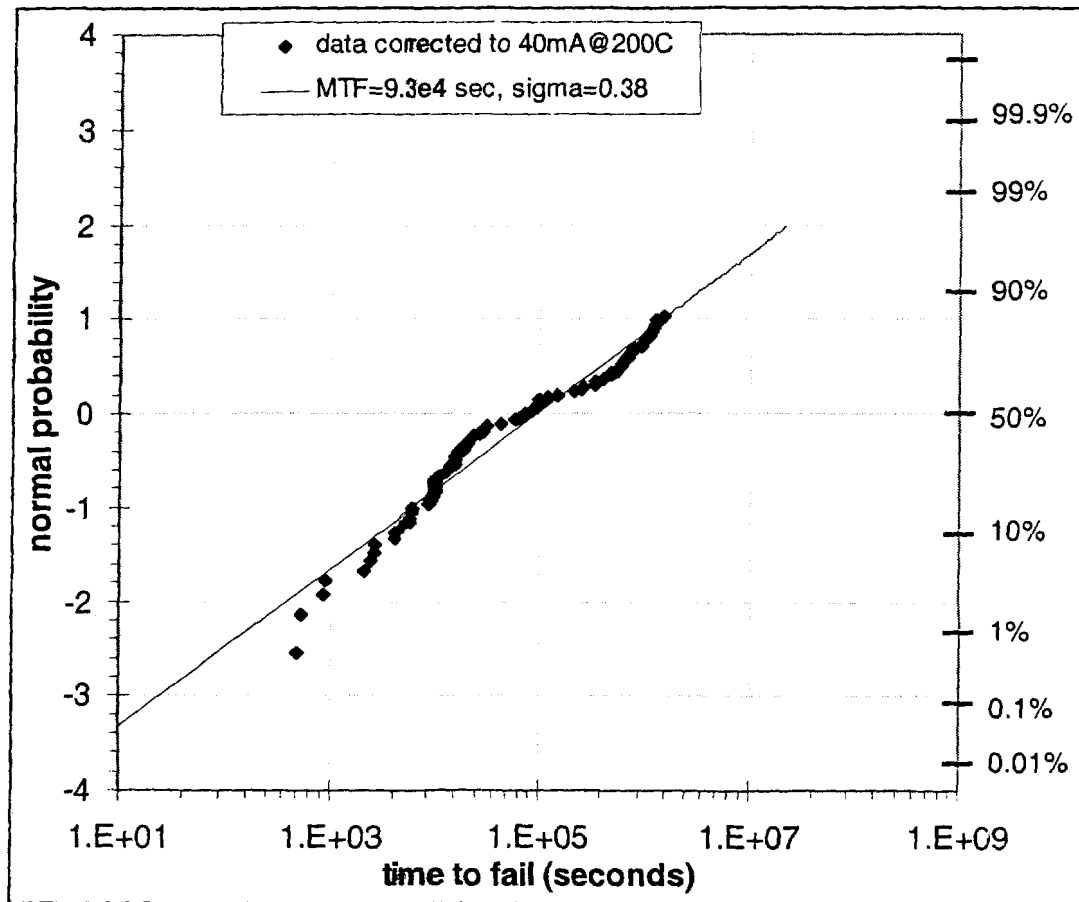


Figure 5.11: Probability Plot of Fail Times

5.7. Threshold Estimate

Just as there is a variation in the kinetics of time-to-failure, we expect variation in the thermodynamics of current-to-mortality (threshold). Figure 5.12 shows the I-L values at which structures were observed to fail or not, and samples were counted with increasing or decreasing the I-L respectively. As shown, the weakest sample failed at I-L=40 mA· μ m, while the strongest sample showed no signs of failing after much time at I-L=640 mA· μ m. Although the 'weak' sample may have been damaged by some electrical overstress (lightning strike) and the 'strong' sample may have failed given enough time, we will assume that this variability is due to mechanical constraint of the hydrostatic forces induced by electromigration. At each I-L value, the 'fraction mortal' is

calculated, and the resulting lognormal plot is shown in

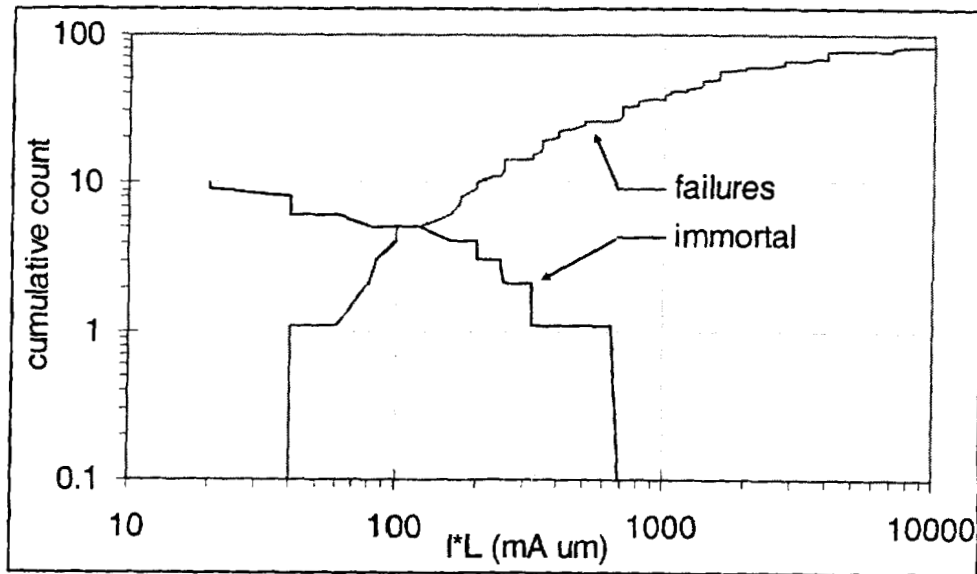


Figure 5.12: Variation in EM Thresholds

Reports in the literature have shown threshold values for Cu damascene with various barriers and low-k dielectrics. The best values range from 2850 to 7000 A/cm, which for the $0.28 \mu\text{m}^2$ cross-section (see page 50) is equivalent to 79 to 194 mA· μm .

$$\left(\frac{2850}{7000} \right) \cdot \frac{\text{amp}}{\text{cm}} \cdot \text{area}_{\text{segment}} = \left(\frac{78.945}{193.9} \right) \cdot \text{mA} \cdot (\mu\text{m})$$

. If this extrapolation to small sample sizes is warranted, designers using $I \cdot L < 11.2$ mA· μm can be assured that less than 0.1% of structures will ever fail.

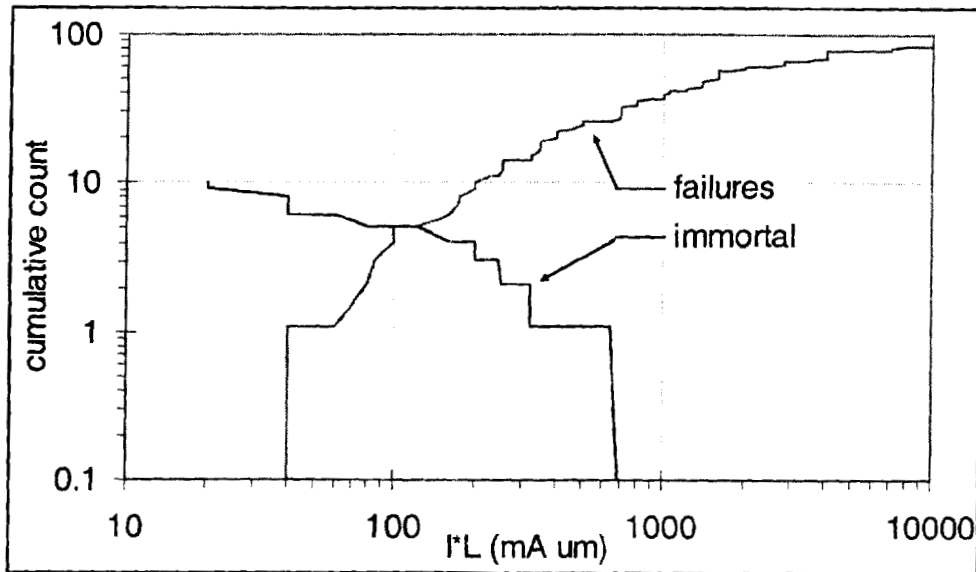


Figure 5.12: Variation in EM Thresholds

Reports in the literature have shown threshold values for Cu damascene with various barriers and low-k dielectrics. The best values range from 2850 to 7000 A/cm, which for the $0.28 \mu\text{m}^2$ cross-section (see page 50) is equivalent to 79 to 194 mA· μm .

$$\left(\frac{2850}{7000} \right) \frac{\text{amp}}{\text{cm}} \cdot \text{area}_{\text{segment}} = \left(\frac{78.945}{193.9} \right) \cdot \text{mA} \cdot (\mu \cdot \text{m})$$

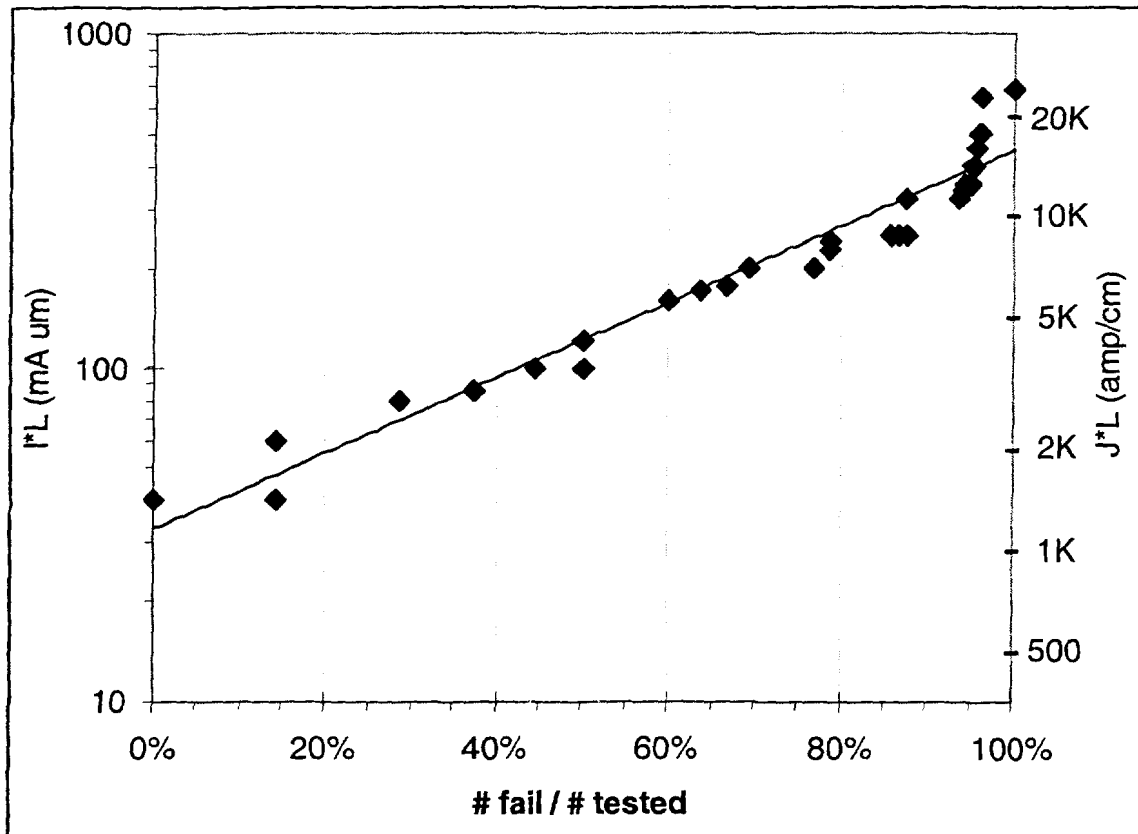


Figure 5.13: Statistical Distribution of EM Thresholds

5.8. Multiple Activation Energy Model

Black's model (2.41) and the time constants (2.33) depend on temperature, and the variation of temperature complicates applying the models. The literature suggests various temperature estimates have been used, including maximum, average, weighted average, and via temperature.

Re-writing Black's equation as a logarithm (5.1), and treating the activation energy as the sum of components (5.2) is useful for regression. Obviously, if all T_i were the same, then activation energy is sum of the components ($E_a = \sum E_i$) but by using different estimates of temperature as independent degrees of freedom, regression can compare goodness of fit to each estimate. This is not an attempt to find physical mechanisms corresponding to these activations, but only a construct to identify which temperature best fits Black's equation. The choices of ambient (wafer), via (minimum),

center (maximum), and average temperatures were applied to a solution of linear equations with measured time-to-failure. Once again, matrix methods were used to find a least squares solution for the overdetermined system. Structure-to-structure variation is characterized by sigma of the log-normal distribution, which corresponds to the residual of the regression fit. The choice of few samples at many conditions provides better response to parameters than more numerous samples at fewer conditions.

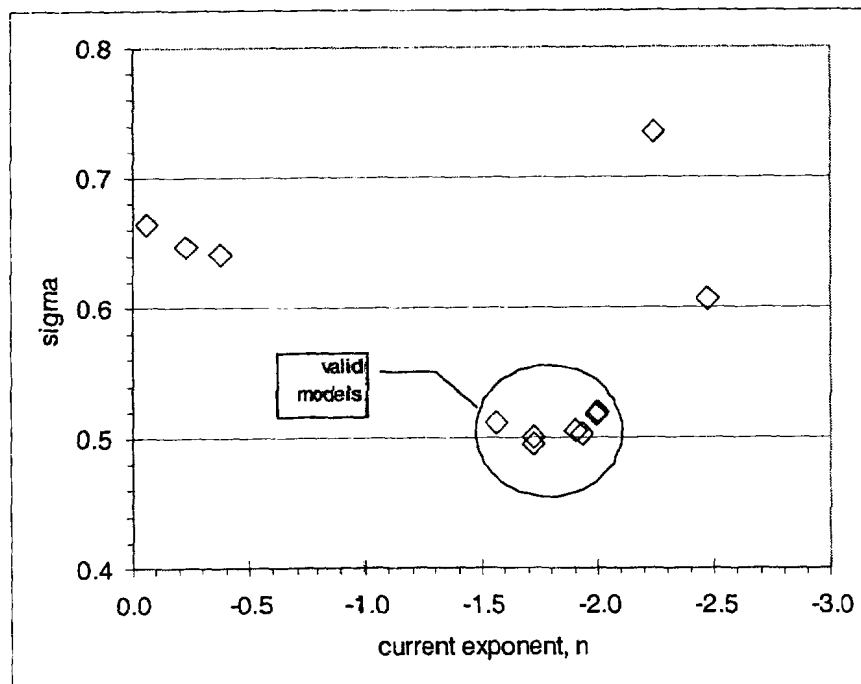
$$\ln(tfail) = offset + n \cdot \ln(I) + \frac{E_a}{k_B \cdot T} \quad (5.1)$$

$$\ln(tfail) = offset + n \cdot \ln(I) + \frac{E_0}{k_B \cdot T_0} + \frac{E_1}{k_B \cdot T_1} + \frac{E_2}{k_B \cdot T_2} + \dots \quad (5.2)$$

Table 5.1 and Figure 5.14 show the fitting parameters for several multiple-activation models for the data of Figure 5.11. Black's equation with a single temperature estimate (models 1 thru 4) has a large lognormal sigma (residual), and the current exponent, n , departs from the expected value of -2 for models 3 and 4. (While sources in the literature propose a current exponent $n=1$, this applies only to unconstrained drift velocity which for this data is a negligible portion of the total time-to-failure.) Regression can produce negative E_a estimates (models 12, 34, 123, & 124), which is an artifact of linear algebra using degrees of freedom to best correlate with data. Models 23 and 24 are among the best fit, and are strongly weighted to the segment end (via) temperature but with an additional hotter component. The use of 3 temperatures does not improve the fit and overcomplicates Black's equation with excess parameters.

Table 5.1: List of Acceleration Models

model	offset	n	Activation Energy				sigma
			Twafer	Tvia	Tavg	Tmax	
1	-0.33	-2.24	0.66				0.73
2	-0.55	-2.48		1.02			0.61
3	0.08	-0.23			0.33		0.65
4	0.12	-0.06				0.29	0.66
12	-0.53	-1.56	-1.27	2.43			0.51
13	-0.39	-2.00	0.52		0.29		0.52
14	-0.39	-1.99	0.56			0.26	0.52
23	-0.45	-1.90		0.76		0.18	0.50
24	-0.45	-1.93		0.73	0.21		0.50
34	0.04	-0.38			0.64	-0.28	0.64
123	-0.48	-1.72	-0.56	1.45	0.14		0.50
124	-0.48	-1.72	-0.54	1.45		0.12	0.50
134	-0.39	-2.00	0.55		0.11	0.16	0.52
234	-0.45	-1.93		0.73	0.19	0.02	0.50

**Figure 5.14: Scatterplot of Acceleration Models**

5.9. Summary

The chapter presented the method of deriving useful temperature/current stress conditions, and measurements at those conditions. Resistance vs time at near-threshold conditions were used to derive EM threshold (Blech Length), while statistical analysis of above threshold times-to-failure determined the coefficients of Black's equation. The long and short-line EM performance is translated into 'do-not-exceed' limits for current in metal lines. The implications of these design rules for product layout are discussed next.

Chapter 6

VLSI Design Rules

EM product failure is preventable, if designers use appropriate limits during IC layout. A more robust metallization technology allows narrower lines, reduced die size, and improved product performance, but perhaps more significant is the benefit of improved designer productivity. In the past, layout was performed with automated routing tools which connected features with lines of arbitrary width. The completed layout was simulated, currents estimated, and lines flagged as errors (violations) when current limits were exceeded. It was then necessary to manually redraw many lines, costing much delay and many man-hours. Since simulations of the revised layout would flag other errors, this method became iterative and required re-layout after re-layout. The industry is working toward a 'do it right the first time' methodology. Early estimates of required line width would greatly reduce the numbers of iterations to complete an error-free design. This chapter derives a method which suggests a safe width from the design netlist (prior to layout).

6.1. Supply Line Duty Factor

The repeater circuit shown in Figure 6.1 can represent much more complicated circuits. The figure shows 8 repeater gates between latches, and the signal must transit all 8 gates during a single clock cycle. Other logic circuits have greater fan-out and fan-in, but in similar fashion they must complete their task during a clock cycle. The number of gates per clock cycle is known as the 'logic depth' (which is 8 for this example).¹

¹ The debate of 'RISC' (Reduced Instruction Set Computer) versus 'CISC' (Complete Instruction Set Computer) focused on the optimum logic depth, with RISC favoring fewer (perhaps 4 to 6) gates in series and CISC allowing more (8 to 12) gates. The premise is that RISC allows higher clock frequency and increases utilization of each transistor.

IC interconnect lines can be categorized as either signals or supplies, with very different current waveforms. Signals generally charge and discharge a capacitor, with zero average current. The V_{CC} supply carries a pulse of charge as the pull-up transistor turns on, and the V_{SS} supply carries a charge pulse when the pull-down transistor first turns on. Lower transistor + interconnect resistance means a quicker pulse with higher peak current, but the average current given by (6.1) is the same for both V_{CC} and V_{SS} supplies.

$$I_{avg} = C \cdot F \cdot V_{cc} \quad (6.1)$$

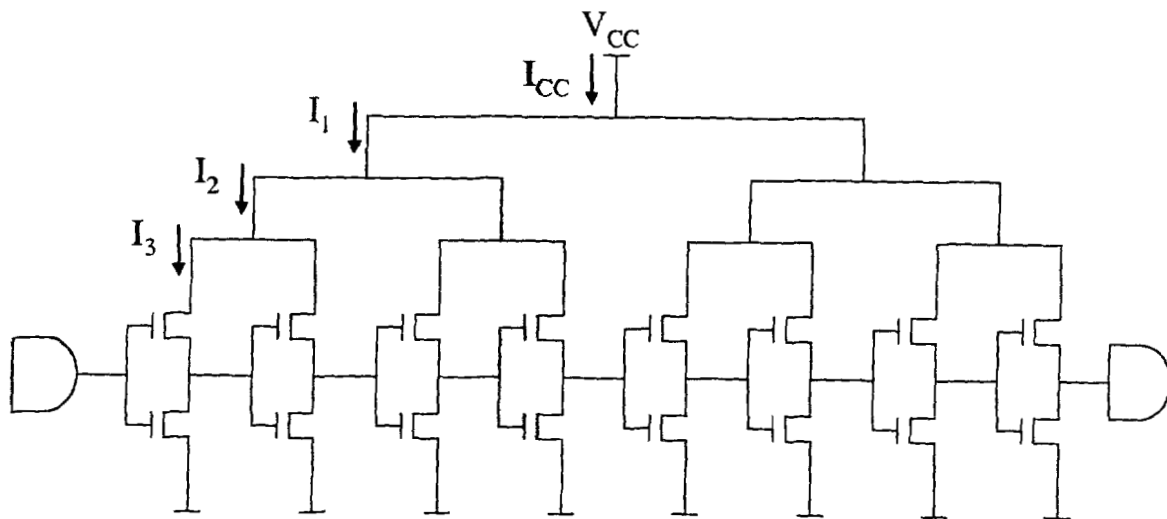


Figure 6.1: Simple Inverter Chain Circuit

Of course, the clock frequency is optimized to match the logic depth, so the signal just barely completes the transit in time. In this best case scenario, current I_{CC} in Figure 6.1 is nearly DC, while the pulse widths of I_1 , I_2 , and I_3 successively decrease. The ratio of pulse width to clock period is known as duty factor, D . Duty factor is about 100% for I_{CC} , about 50% for I_1 , about 25% for I_2 , and about 12% for I_3 . Approximating each supply current waveform with square pulses, the following relationships hold:

$$I_{peak} = \frac{I_{avg}}{D} \quad (6.2)$$

$$I_{rms} = \sqrt{(I_{peak})^2 \cdot D} = \frac{I_{avg}}{\sqrt{D}} \quad (6.3)$$

6.2. Design limits

Designers must account for several limits during layout of lines. In addition to EM current limits, selfheating temperature and voltage ($I \cdot R$) drop can be significant. RMS current affects temperature and EM reliability limits the average current which depends on the RMS current thru the temperature dependence of EM.[28] Given the electrical resistance (6.4) and thermal resistance (6.5), temperature rise is given in transcendental form (6.6) and explicit form (6.7), which allows calculating the greatest load capacitance (6.8) for a maximum temperature rise, T_{max} . Figure 6.2 shows that an insignificant temperature rise near DC current becomes unacceptable below 1% duty factor (for the example of 0.2μ metal over 5μ oxide).

$$R = \frac{\rho_0 \cdot (1 + \alpha \cdot \Delta T) \cdot L}{w \cdot t_m} \quad (6.4)$$

$$\theta = \frac{t_{ox}}{L \cdot w \cdot c_{ox}} \quad (6.5)$$

for w = line width, L = segment length, t_m = metal thickness, t_{ox} = underlying dielectric thickness, and F = clock frequency.

$$\Delta T = (I_{rms})^2 \cdot R \cdot \theta = \frac{C^2}{D} \cdot \frac{1 + \alpha \cdot \Delta T}{bx}, \quad bx := \left(\frac{w}{F \cdot V_{cc}} \right)^2 \cdot \frac{t_m \cdot c_{ox}}{\rho_0 \cdot t_{ox}} \quad (6.6)$$

$$\Delta T = \frac{1}{\frac{D}{C^2} \cdot bx - \alpha} \quad (6.7)$$

$$C_{\Delta T(D)} := \sqrt{\frac{D \cdot bx}{\frac{1}{T_{max}} + \alpha}} \quad (6.8)$$

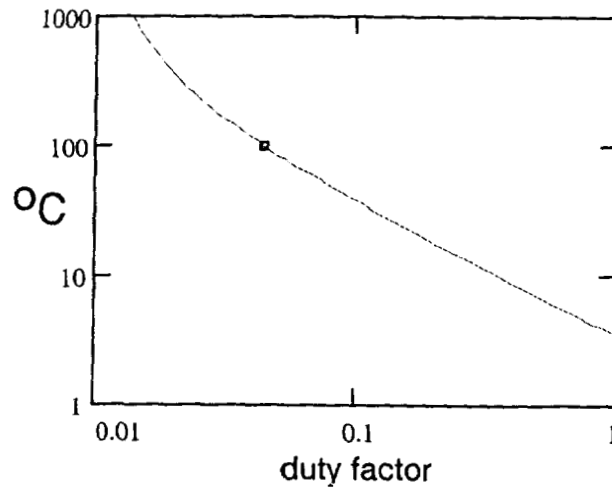


Figure 6.2: Temperature Rise

The voltage drop given by (6.9) also becomes unacceptable for low duty factor, as shown in Figure 6.2 (for an example of 20μ segment length). A maximum voltage drop, V_{max} , limits maximum allowed load capacitance (6.10).

$$\Delta V = I_{peak} \cdot R = \frac{C \cdot F \cdot V_{CC} \cdot \rho_0 \cdot (1 + \alpha \cdot \Delta T) \cdot L}{D \cdot w \cdot t_m} \quad (6.9)$$

$$C_{\Delta V} = \frac{V_{max} \cdot D \cdot w \cdot t_m}{F \cdot V_{CC} \cdot \rho_0 \cdot (1 + \alpha \cdot T_{max}) \cdot L} \quad (6.10)$$

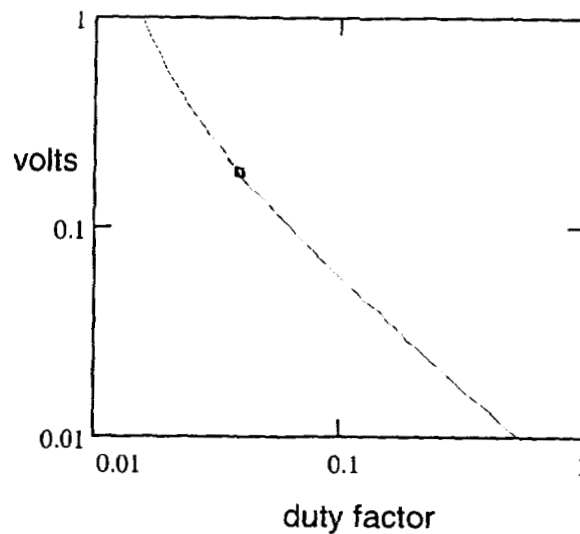


Figure 6.3: I*R Drop

The current density allowed in long-lines (6.11) is derived from EM Black's equation (6.36). The long and short line EM limits are combined in (6.12). Figure 6.3 compares the various capacitance limits, showing that I*R drop limits the shortest pulses, while long-line EM is important near DC (duty factor = 1). The EM curve in Figure 6.3 includes the short-line region (constant with temperature) and the long-line region (where allowed capacitance increases with larger duty factor because of lower temperature). Temperature certainly affects the long-EM limit, but the explicit temperature rise limit may also be important for intermediate duty factors.

$$J_{\text{limit}}(\Delta T) := J_{\text{dr}} \exp \left[\frac{E_a}{n \cdot k_B} \cdot \left(\frac{1}{T_{\text{dr}} + \Delta T} - \frac{1}{T_{\text{dr}}} \right) \right] \quad (6.11)$$

$$C_{\text{EM}} = \frac{w \cdot t_m}{F \cdot V_{\text{CC}}} \cdot \text{greater} \left(J_{\text{limit}}(\Delta T), \frac{J_{\text{L}_{\text{max}}}}{L} \right) \quad (6.12)$$

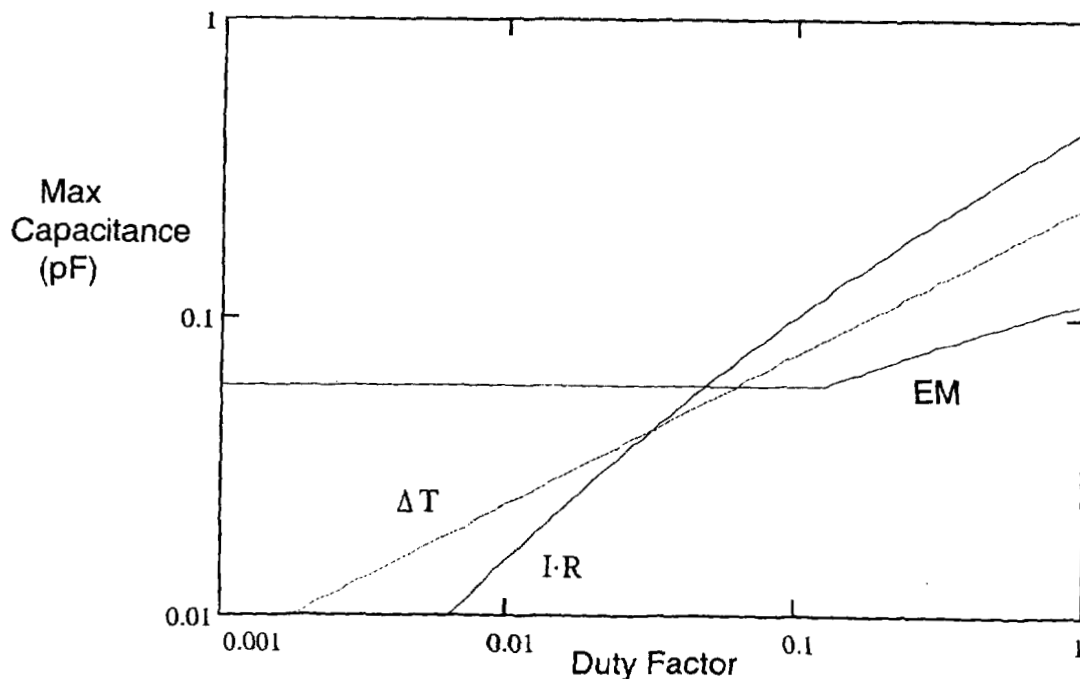


Figure 6.4: Comparison of Capacitance Limits

6.3. Consequences

The simultaneous solution of pairs of equations (6.8) and (6.10) defines curve (6.13) separating regions in L-D space. Voltage is critical for longer lines, while temperature is more critical for shorter lines, as shown in Figure 6.5. Information about allowed current or load capacitance has been lost, but this 'map' indicates which mechanism limits capacitance. If a supply line complies with the critical limit, then the other mechanism is also in compliance. Designer effort is reduced by layout and checking a single limit.

$$D = \left(\frac{L}{V_{\max}} \right)^2 \cdot T_{\max} (1 + \alpha \cdot T_{\max}) \cdot \frac{\rho_0 \cdot c_{\text{ox}}}{t_m \cdot t_{\text{ox}}} \quad (6.13)$$

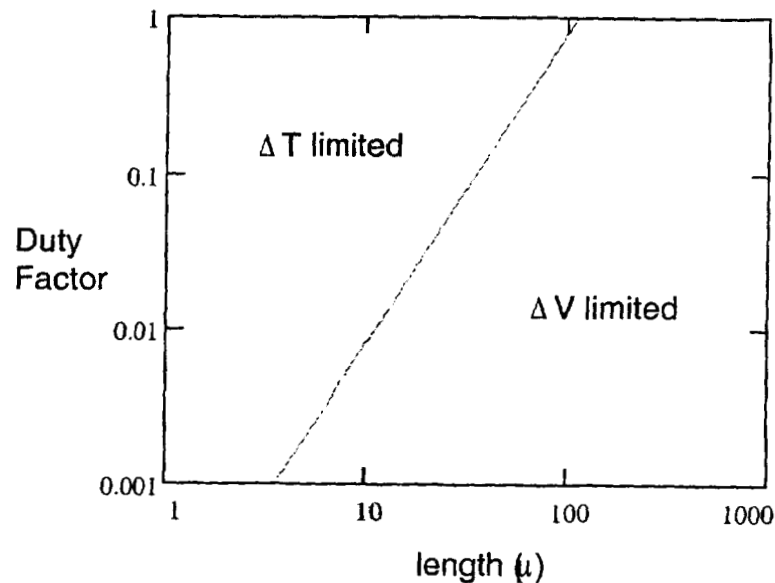


Figure 6.5: Temperature/Voltage Boundary

$$D = \left(\frac{JL_{\max}}{L} \right)^2 \cdot \rho_0 \cdot \frac{t_{\text{ox}} \cdot t_m}{c_{\text{ox}}} \cdot \left(\frac{1}{T_{\text{rise}} \left(\frac{JL_{\max}}{J_{\text{dr}} \cdot L} \right) + \alpha} \right) \quad (6.14)$$

The similar boundary between long and short line EM from (6.12) is given in (6.14) and shown in Figure 6.6.

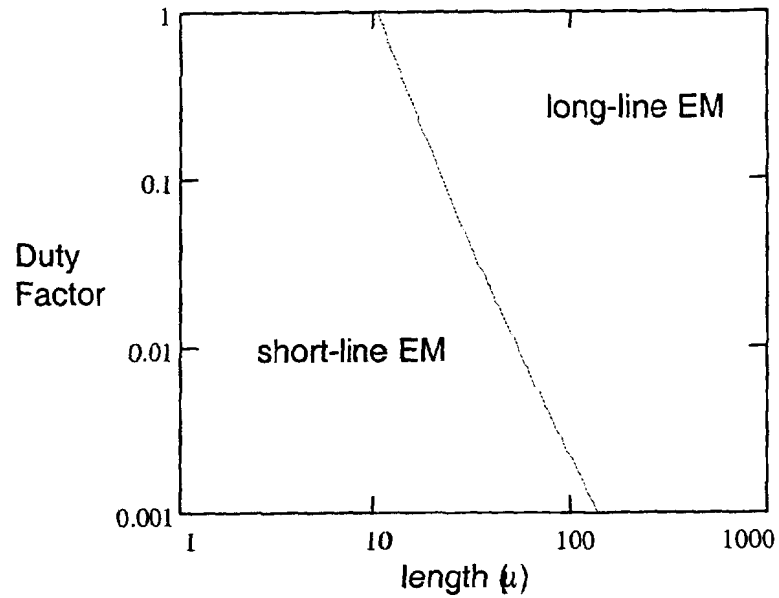


Figure 6.6: Long/Short EM Boundary

Boundaries between temperature rise and long line EM (6.15), between temperature rise and short line EM (6.16), and between voltage drop and short line EM (6.17) are explicit, while the boundary between voltage drop and long line EM requires a numerical method solving for D and L at which (6.10) equals the long line term of (6.12): $C_{\Delta V} \approx C_{EM}$.

$$D = J_{lim} (T_{max})^2 \cdot \rho_0 \cdot \left(\frac{1}{T_{max}} + \alpha \right) \cdot \frac{t_{ox} \cdot t_m}{c_{ox}} \quad (6.15)$$

$$D = \left(\frac{JL_{max}}{L} \right)^2 \cdot \rho_0 \cdot \left(\frac{1}{T_{max}} + \alpha \right) \cdot \frac{t_{ox} \cdot t_m}{c_{ox}} \quad (6.16)$$

$$D = \rho_0 \cdot \left[\left(\frac{JL_{max}}{L} \right)^2 \cdot \frac{t_m \cdot t_{ox} \cdot \alpha}{c_{ox}} + \frac{JL_{max}}{V_{max}} \right] \quad (6.17)$$

The various boundaries are combined on Figure 6.7, resulting in a map indicating which mechanism will limit a particular geometry.

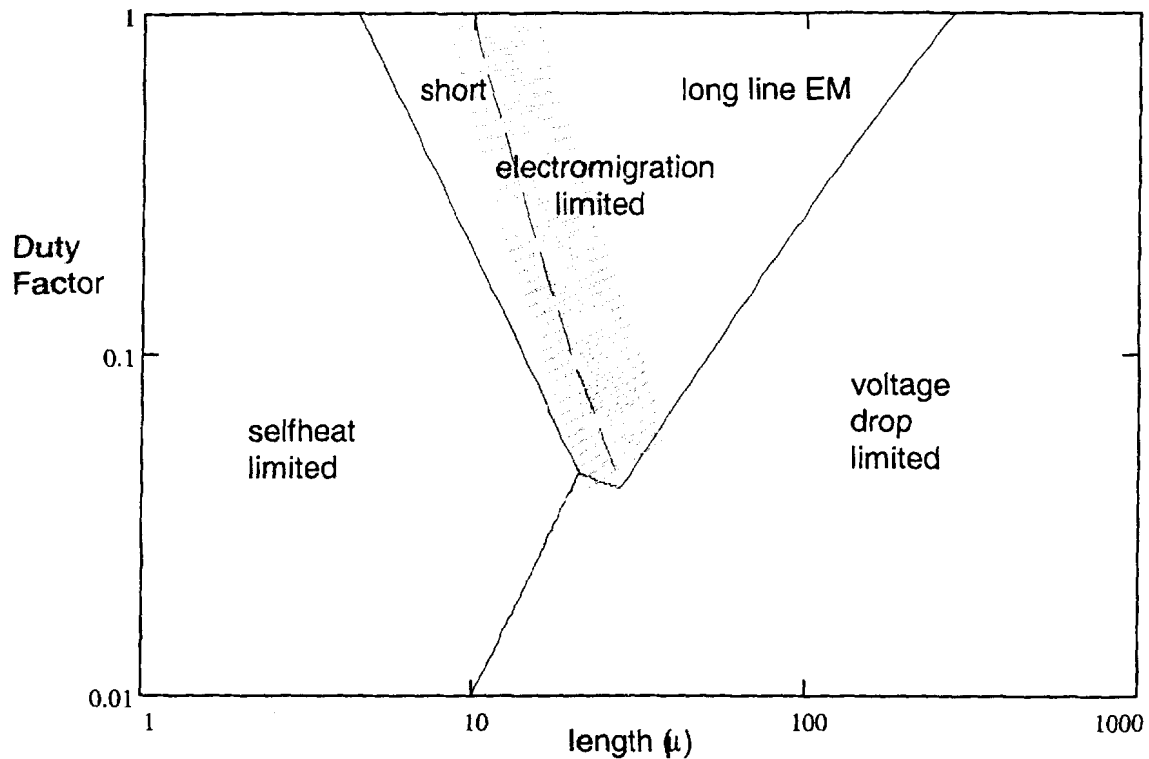


Figure 6.7: Map of Design Limits. (Hashed area is near-threshold region.)

This map shows that EM is not an issue for lines longer than a millimeter. Since printed circuit boards and wires have larger thermal resistance, the selfheat region expands to longer lengths, further reducing the area of the EM region. Solder bumps may show EM degradation, or temperature may be more critical. As PC boards shrink to sub-millimeter dimensions with improved thermal capability, look for EM to become an issue.

This map uses a simple thermal model (6.5) which assumes only dielectric underlying the metal line. For better accuracy, future work might include length dependence from (6.19) and/or fringing from (6.9). Figure 6.7 represents a worst case analysis for a line with unknown surroundings, and is a conservative design tool to safely define metal width during layout.

6.4. Supply Tree vs Grid.

The local supply connections divide to individual transistors (branch like a tree) as in Figure 6.1, but global supplies are usually organized as a mesh as shown in Figure

6.8. (Even if the mesh is not evident on-chip, it becomes de-facto for packaged products with many V_{CC} and V_{SS} pin connections.) 'Manhattan' routing is the practice of implementing north/south lines and east/west lines in alternating metal levels – for example, lines in metal levels 1,3,5 may run perpendicular to lines in levels 2,4,6.

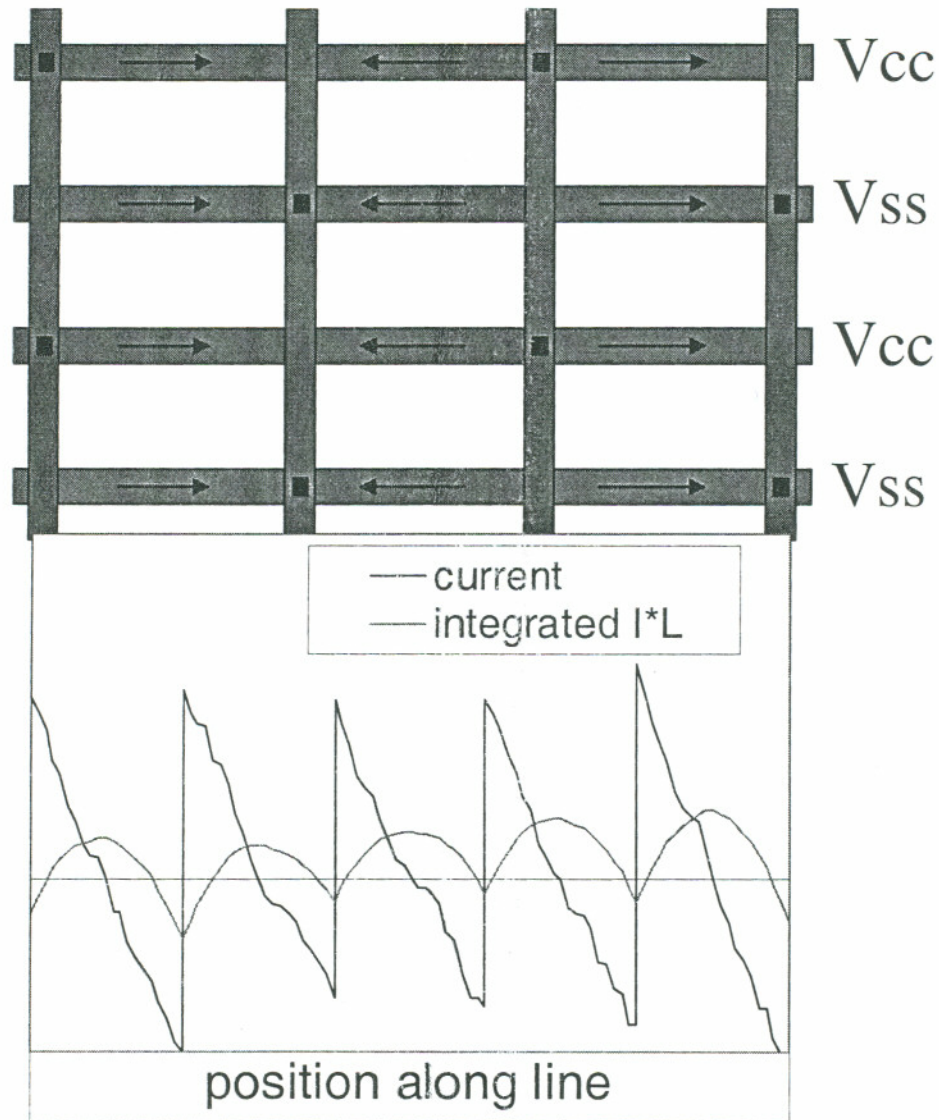


Figure 6.8: Supply Mesh (arrows show direction of current flow)

With Manhattan routing, a given supply line receives current from the next higher metal level thru vias periodically along its length, and distributes current to the next lower metal level thru more numerous vias. This configuration causes the current in each line to reverse at each upper via and the current drops to zero at some point between the

upper vias. The length between current reversals can be considered a partition for Blech effect purposes. Grid lines may be very long, but the Blech length applies separately to each partition bounded by numerous anodes and cathodes along the long metal line. Designers can place much higher currents in the supply mesh if layout design rules take advantage of the greatly improved reliability due to the Blech effect. Most supply meshes are self-heat or voltage drop limited, and should not be EM limited.

6.5. Summary

Given the current limits derived in Chapter Chapter 5, consequences for practical supply network layout were discussed. Only a limited range of segment lengths and duty factors will be susceptible to EM damage because other structures and conditions will be limited by either selfheating or voltage drop. The last chapter summarizes the major findings of this investigation and suggests areas to explore in the future.

Chapter 7

Conclusion

7.1. Findings

Reports by other investigators use alternative techniques to quantify EM threshold. Resistance saturation presumes that voids reach a maximum size at equilibrium, and elastic containment implies that a greater stress creates a larger void. This model is supported by data from several sources which show that saturation resistance increases with I-L stress. EM threshold is the I-L stress at which the maximum (saturated) resistance is below the resistance increase failure criteria. This investigation found few examples of resistance saturation - not sufficient to show a trend with current stress. An alternative to saturation resistance is resistance slope technique which gives equivalent EM threshold values.[29]

Of course time-to-failure increases at low current, but does the current dependence deviate from Black's dependence as current decreases to threshold? Are the kinetics of failure far from equilibrium different from kinetics near threshold? Other investigators have presented linear graphs of the inverse of time-to-failure ($1/MTF$) as a linear function of current, similar to the straight line in Figure 7.1.[30] This implies $n=1$ with an offset, so that extrapolations to the x-axis crossing ($1/MTF=0$) determine a finite current at infinite time-to-failure. The expectation is that the degradation process slows down and eventually stops below equilibrium. The dashed curve in Figure 7.1 shows Black's power-law current dependence for comparison. With typically large statistical variance in measurements, large sample sizes are required to determine which dependence better fits the data. A better test is to measure a wider current range and plot on a log-log graph, as shown in Figure 7.2. While the curves are similar in the range of 1.5 to 4 times the threshold current, Figure 7.1 disguises the large relative difference between the models near threshold, which shows clearly in Figure 7.2.

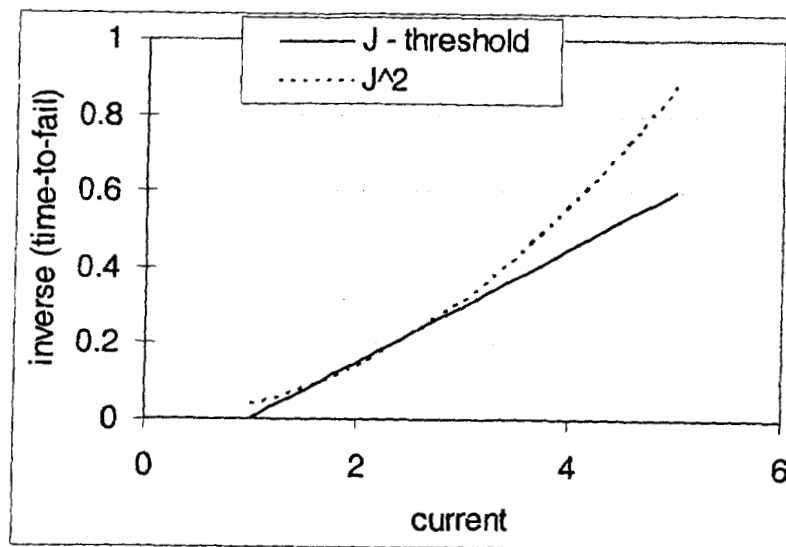


Figure 7.1: Linear Plot of Inverse MTF (Threshold Dependence)

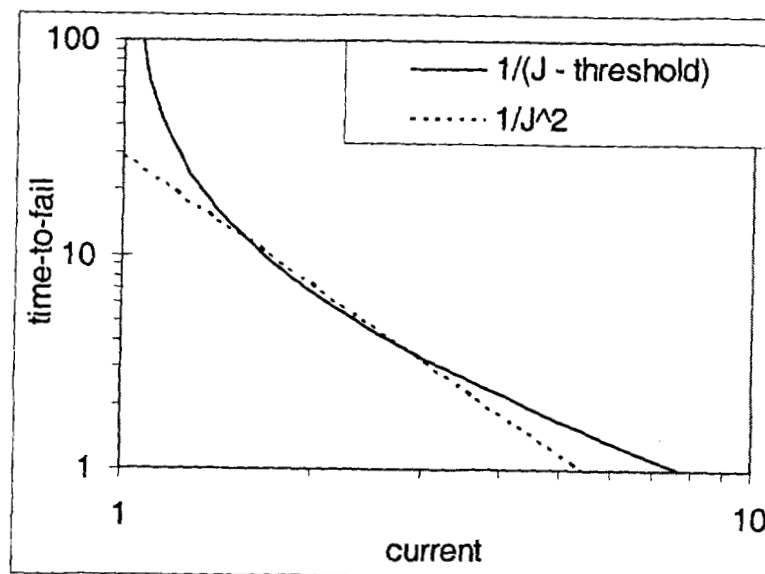


Figure 7.2: Log Plot of MTF (Black Dependence)

Data in this dissertation follow the power-law dependence over a 7x current range. At high current, structures fail much quicker than a linear model predict. While the mechanism may transition to a linear model at low current, no statistical test was found to validate that. Instead, examination of the residuals of the regression shows no sign of increased lifetime as threshold is approached. In exploring near-threshold stress

conditions, the threshold determines whether a structure fails or not, and if so, the power-law dependence determines the time-to-fail. Both threshold and time-to-failure exhibit significant variation. The high-current stresses have thermal gradients which may have reduced the threshold (see section 2.2.1). Black's equation remains valid to segment lengths on the order of the Blech length - and even below what the EM threshold would have been at lower current stress.

7.2. Application to Process Monitors

The fears of overstress may be a result of inadequate understanding of thermal effects on EM performance. These data show that meaningful extrapolations can be made on measurements from low currents (below mean threshold) to high currents (exceeding half the runaway current). With a suite of carefully characterized test structures, both long-line and short-line EM performance can be evaluated. Resistance measurements give clues of the dynamics of void growth and quantify some of the material parameters that determine EM threshold.

With a throughput exceeding 1 sample per probe per day a coordinated wafer-level sample has cost and process feedback advantages over package-level oven stress. Reduced time-to-failure (bad EM performance) will greatly increase wafer stress throughput, providing early warning of problems. Assembly overhead dominates oven stress throughput, slowing response. For production monitors on product wafers, wafer-level measurements allow flexible sample plans. Only probed structures are damaged with wafer-level tests so product is unaffected, while packaging requires dicing the wafer which may destroy product and restrict availability of samples. The across-wafer variation has decreased to less than the wafer-to-wafer and lot-to-lot variations. A sampling is best if it corresponds to the source of variation, so more wafers and more lots could be sampled by using fewer samples per wafer.

7.3. Future Work

The use of wafer-level EM as process monitors allows examination of various process alternatives. The cladding/barrier underlayer affects the deposited Cu microstructure and resulting EM performance.[31] Alternative underlayers were

considered for this investigation, but PVD TiN was the only process option that produced functional test structures. Recent reports propose that EM performance strongly depends on surface diffusion at the overlying SiN etch-stop. Resistance saturation measurements can compare diffusivity of various overlayers. Long-line EM kinetics include possible thermal feedback and waves, but the electrical signal by itself is too noisy to demonstrate the expected instabilities. Detection of spatial stress variation with AFM or two-beam interferometry would confirm EM waves, and in-situ measurements during stress should detect wave motion. Measuring wavelength, phase velocity, and group velocity would quantify vacancy generation/recombination and diffusion kinetics.

While this investigation showed that a simple short-loop process meets the present-day needs of the International Semiconductor Technology Roadmap (ISTR) shown in Table 1.2, significant challenges remain for future technologies. As TEOS is replaced by alternative ILDs with reduced dielectric constant (low-k dielectrics), containment and adhesion will be degraded.[32] The industry has serious challenges to maintain scaling of future technologies.

References

- [1] The International Technology Roadmap for Semiconductors, 2002 update, can be found at "<http://public.itrs.net/Files/2002Update/2002Update.pdf>", p. 75 (downloaded August 5, 2003)
- [2] John Emsley, "The Elements", 2nd ed., Clarendon Press, Oxford, ISBN 0198555687 (1991)
- [3] N.L.Peterson, "Self-Diffusion in Pure Metals" *Journal of Nuclear Materials*, vol. 69&70, pp. 3-37 (1978)
- [4] Neil W. Ashcroft & N. David Mermin, "Solid-State Physics", Holt, Rinehart and Winston, ISBN 0030839939, p. 617 (1976)
- [5] Milton Ohring, "Reliability and Failure of Electronic Materials and Devices", Academic Press, ISBN 0125249853, p. 23 (1998)
- [6] H.Weaver in "Electro- and Thermo- Transport in Metals and Alloys", ed. R.E.Hummel, H.B.Huntington, TMS-AIME, New York, 1977 p.37-52.
- [7] Milton Ohring, "Reliability and Failure of Electronic Materials and Devices", Academic Press, ISBN 0125249853, p. 480 (1998)
- [8] A.D. LeClaire, "Some Predicted Effects of Temperature Gradients on Diffusion in Crystals" *Physical Review*, vol. 93 (2), p. 344 (1954)
- [9] I. Blech, "Electromigration in thin aluminum films on titanium nitride", *Journal of Applied Physics*, vol. 47, pp.1203-1208 (1976)
- [10] R. Kircheim, "Modelling Electromigration and Induced Stresses in Aluminum Lines", *Materials Reliability Issues in Microelectronics III*, Material Research Society proceedings 309, ISBN 1558992057, pp. 101-110 (1993)
- [11] Neil W. Ashcroft & N. David Mermin "Solid-State Physics" Holt, Rinehart and Winston, ISBN 0030839939, pp.7-8 (1976)
- [12] Ru, C.Q.; "Intrinsic Instability of Electromigration Induced Mass Transport in a Two-Dimensional Conductor"; *Acta Materialis*, vol. 47 (13), p.3571-3578 (1999)

- [13] Ru, C.Q.; "Thermomigration as a driving force for instability of electromigration induced mass transport in interconnect lines"; *Journal of Materials Science*, vol. 35, p.5575-9 (2000)
- [14] I.A. Blech, E.S.Meieran, "Direct Transmission Electron Microscope Observation of Electrotransport in Aluminum Thin Films", *Applied Physics Letters* vol. 11 (8), pp. 263-6 (1967)
- [15] Ruth DeJule, "Dual-Damascene: Overcoming Process Issues", *Semiconductor International*; vol. 23 (6), pp. 94-102 (June 2000)
- [16] Frank Incropera, David DeWitt "Fundamentals of Heat and Mass Transfer" Fourth Edition, John Wiley & Sons, ISBN 0471304603, p.55 (1996)
- [17] C. G. Shirley, "Steady-state temperature profiles in narrow thin-film conductors", *Journal of Applied Physics*, vol. 57 (3), pp.777-784 (1985)
- [18] C. P. Yuan, T. N. Trick, "A Simple Formula for the Estimation of the Capacitance of Two-Dimensional Interconnects in VLSI Circuits", *IEEE Electron Device Letters*, vol. EDL-3 (12), pp. 391-393 (1982)
- [19] "CRC Handbook of Chemistry and Physics" 84th edition, CRC Press, pp. 12-45,46 (2003)
- [20] Dieter Schroder, "Semiconductor Material and Device Characterization" 2nd edition, John Wiley and Sons, New York, ISBN 0471241393, p.140 (1998)
- [21] "MATLAB® The Language of Technical Computing" Version 5, The Math Works Inc., Natick MA, pp. 4-13 to 4-17 (1999)
- [22] Anke Zitzelsberger, et al. "On the Use of Highly Accelerated Electromigration Tests (SWEAT) on Copper" Proceedings 41st Annual International Reliability Physics Symposium pp.161-165 (2003)
- [23] M. A. Korhonen, P. Borgesen, D. D. Brown, Che-Yu Li, "Microstructure based statistical model of electromigration damage in confined line metallizations in the presence of thermally induced stresses", *Journal of Applied Physics*, vol. 74 (8), pp. 4995-5004 (1993)
- [24] Z. Suo, "Stable state of interconnect under temperature change and electric current", *Acta Materiala*, vol. 46 (11), pp. 3725-32 (1998)

- [25] J.J. Clement, S.P.Riege, R.Cvijetic, C.V.Thompson, "Methodology for electromigration critical threshold design rule evaluation", *IEEE Transactions on Computer Aided Design of Integrated Circuits and Systems*, vol. 18 (5), pp. 576-81 (1999)
- [26] C.V.Thompson "What's Different About the Reliability of Al-based and Cu-based Interconnects?" Materials Research Society 2003 Spring Meeting, paper E1.1 (2003)
- [27] A. von Glasow, et al., "The Influence of the SiN Cap Process on the Electromigration and Stressvoiding Performance of Dual Damascene Cu Interconnects" Proceedings 41st International Reliability Physics Symposium, pp. 146-150 (2003)
- [28] William R. Hunter, "Self-Consistent Solutions for Allowed Interconnect Current Density - Part I: Implications for Technology Evolution" *IEEE Transactions on Electron Devices*, vol. 44 (2), pp. 304-309 (1997)
- [29] Lucile Arnoud, "Electromigration Threshold Length Effect in Dual Damascene Copper-Oxide Interconnects", Proceedings 40th Annual International Reliability Physics Symposium, pp. 433-434 (2002)
- [30] S. Thrasher, C. Capasso, L. Zhao, R. Hernandez, P. Mulski, S. Rose, T. Nguyen, H. Kawasaki, "Blech Effect in Single-Inlaid Cu Interconnects" Proceedings 39th Annual International Reliability Physics Symposium, pp.177-179 (2001)
- [31] G. B. Alers, A. Vijayendran, P. Gillespie, L. Chen, H. Cox, K. Lam, R. Augur, K. Shannon, K. Pfeifer, M. Danek, "Electromigration Improvement with PDL TiN(Si) Barrier in Copper Dual Damascene Structures" Proceedings 41st Annual International Reliability Physics Symposium, pp.151-155 (2003)
- [32] C. S. Hau-Riege, A. P. Marathe, V. Pham, "The Effect of Low-K ILD on the Electromigration Reliability of Cu Interconnects with Different Line Lengths" Proceedings 41st Annual International Reliability Physics Symposium, pp.173-177 (2003)

Biography

Bill Meyer was born in Red Lake Falls, Minnesota, on February 22, 1954. He received his bachelor's degree in Electrical Engineering from the University of Minnesota, Minneapolis, in June 1977. After briefly attending graduate school, he was employed by Intel Corporation as a process reliability engineer. After 20 years of work on dielectrics, statistics, and aluminum interconnect, he returned to graduate school and received his master's degree in Material Science from the Oregon Graduate Institute in June 2000. Publications include:

W.K.Meyer and G.K.Weohner "Abstract: Sputter Yields Of Bimetal Mixtures" J. Vac. Sci. Technol. 16 (2) p.808 (1979)

Crook, Dwight L.; Meyer, William K. "Redundancy Reliability", Annual Proceedings - International Reliability Physics Symposium, IEEE; p.1-10 (1981)

Meyer, William K.; Crook, Dwight L. "Model For Oxide Wearout Due To Charge Trapping" Annual Proceedings - International Reliability Physics Symposium, IEEE; p.242-247 (1983)

Meyer, William K.; Crook, Dwight L. "Non-Aging Screen To Prevent Wearout Of Ultra-Thin Dielectrics", Annual Proceedings - International Reliability Physics Symposium, IEEE; p 6-10 (1985)

Varner, Elizabeth B.; Marieb, Thomas; Mack, Anne Sauter; Lee, Jin; Meyer, William K.; Goodson, Kenneth E., "Thermal conductivity measurements of interlevel dielectrics", Materials Research Society Symposium Proceedings MRS v 473 p 279-284 (1997)

Gleixner, R.J.; Meyer, W.K.; Nix, W.D.; "Electromigration modeling for design rule development in microelectronic interconnects" AIP Conference Proceedings, n 491, p 51-61 (1999)

Suo, Z.; Ma, Q.; Meyer, W.K. "Immortal interconnects - prevent cracking and limit void size" Materials Research Society Symposium - Proceedings, MRS v 563, p 71-76 (1999)