MULTI-OBJECTIVE DESIGN OPTIMIZATION OF BEoL / *f*BEoL REGION DURING CHIP ATTACHMENT TO SUBSTRATE

by

HARDIK PAREKH

Presented to the Faculty of the Graduate School of

The University of Texas at Arlington in Partial Fulfillment

of the Requirements

for the Degree of

MASTER OF SCIENCE IN MECHANICAL ENGINEERING

THE UNIVERSITY OF TEXAS AT ARLINGTON

DECEMBER 2013

ACKNOWLEDGEMENTS

It is a great pleasure to acknowledge and extend my deepest gratitude to few people. First of all I would like to express my special gratitude to my advisor Prof. Dereje Agonafer. His encouragement and knowledge provided invaluable input in project completion.

I would also like to express my appreciation to Prof. Haji-Sheikh and Prof. Kent Lawrence to take time out to review my thesis and for serving on my committee.

Special appreciation goes to my mentor cum friend Fahad Mirza from EMNSPC. Work herein described wouldn't have been possible without his support and invaluable suggestion which helped me improve the quality of work.

I would like to appreciate John Fernandez for helping me on implementation front.

I would like to thank Ms. Sally Thompson for assisting me on each matter patiently. I take a great proud in being a part of EMNSPC lab where I got a chance to work with all brilliant colleagues.

I would also like to thank all my friends and well wishers who supported me throughout my graduate study and career.

Finally, I take immense pleasure to dedicate this work to my parents Bela and Mahesh, and my brother Chintan. Without their love, support and sacrifice, I would not have been in a position to complete this work. .

November 01, 2013

ABSTRACT

MULTI - OBJECTIVE DESIGN OPTIMIZATION OF BEoL / *f*BEoL REGION DURING CHIP ATTACHMENT TO SUBSTRATE

Hardik Parekh, M.S.
The University of Texas at Arlington, 2013

Supervising Professor: Dereje Agonafer

Semiconductor industry has recognized the need to replace traditional Al/SiO_2 interconnects with Cu/Low-*k* interconnects in the mainstream electronics devices following the latter's impact on power, RC delay, and cross-talk reduction. However due to lower elastic modulus, strength, and poor adhesion characteristic, reliability of the Cu/Low-*k* interconnects turns out to be a concern for its integration in the back-end-of-line (BEoL). Flip-chip attachment process (cooling from ~200C to room) can result in critical damage in nano-scale Cu/Low-*k* interconnects.

The objective of this study is to improve the reliability of Cu/Low-*k* interconnects during die attach reflow process for a specific die to substrate size ratio by varying a group of design parameters such as substrate thickness and solder bump footprint. Preliminary parametric study has shown that the variation in the concerned design variables has a significant effect on the solder bump (fBEoL) and low-*k* layer damage (BEoL) [1]. However, there is a trade-off between the solder bump and the dielectric damage with bump footprint, thereby arising a need to perform a multi-objective design optimization.

A simulation based multi-objective design optimization has been carried out to improve BEoL/*f*BEoL reliability under reflow loading by minimizing the following objective functions *1) strain energy in solder bump* and *2) peeling stress in dielectric (low-k layers)*. This work is of immense importance from process integration standpoint. It can provide a quantitative upstream

guideline to the process/electrical team on the BEoL/fBEoL damage.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

Figure Page

LIST OF TABLES

NOMENCLATURE

IC	Integrated Circuits
PCB	Printed Circuit Board
SMT	Surface Mount Technology
CSP	Chip Scale Package
SMD	Surface Mount Devices
C4	Controlled Collapse Chip Connection
FEoL	Front End of Line
BEoL	Back End of Line
UBM	Under Bump Metallization
CMP	Chemical-Mechanical Planarization
*f*BEoL	Far Back End of Line
RC	Resistive-Capacitive
CTE	Co-efficient of Thermal Expansion
CPI	Chip Package Interaction
ILD	Inter-Layer Dielectric
FE	Finite Element
BGA	Ball Grid Array
CCD	Central Composite Design
RDL	Redistribution Layer
FSG	Flourinated Silicate Glass
DOE	Design of Experiment

DNP	Distance from Neutral Point
OSF	Optimal Space Filling
LHS	Latin Hypercube Sampling Design
RS	Response Surface
GOF	Goodness of Fit
GDO	Goal Driven Optimization
MOGA	Multi-Objective Genetic Algorithm
NLPQL	Nonlinear Programming by Quadratic Lagrangian
MISQP	Mixed-Integer Sequential Quadratic Programming

CHAPTER 1

INTRODUCTION

1.1 Electronics Package

1.1.1 Packaging Functions

In contemporary age, almost every electrical device that is being used such as cell phone, laptops, gaming devices, cameras, contains different types of electronics packages. Electronic packaging is defined as the art of establishing interconnections between various levels of electronic devices, components, modules and systems. These packages include different electrical components like transistors, capacitors, resistors, diodes which in turn must be interconnected to perform a specific task. This interconnection of different components is known as circuit and as the level of integration increases, multiple circuit connections moves towards integrated circuits (IC) chips. These ICs and their interconnections must be mechanical supported and environmentally protected. These ICs are supplied with electrical power which being converted into heat. This temperature must be maintained within certain range for ICs to operate at their peak performance [1].

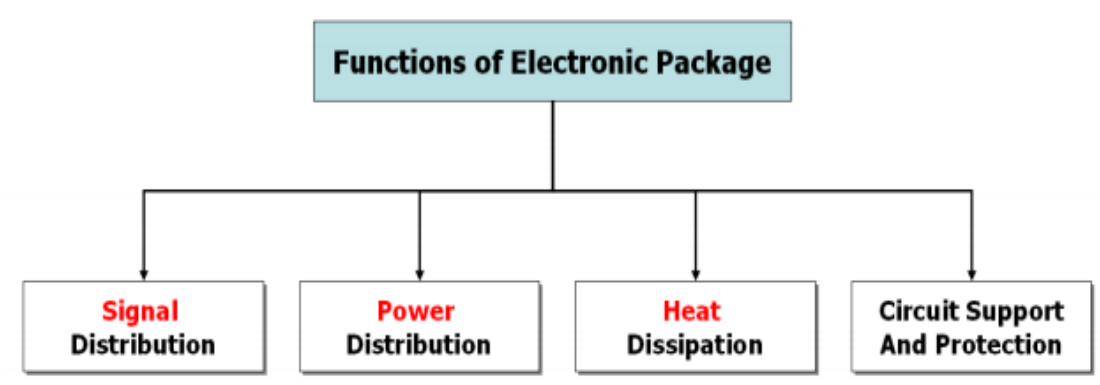


Figure 1.1 Basic Functions of Electronics Package [2]

Thus packaging serves four main functions:

- Signal Distribution
- Power Distribution
- Heat Dissipation
- Mechanical Support, chemical, electromagnetic and environmental protection

In addition to these requirements electronic package should function at pre-specified performance level of the product.

1.1.2. Package Hierarchy

Typical electronics systems have several levels of packaging and each level has its distinctive interconnection associated with it. This hierarchy of interconnects can be divided as follow.

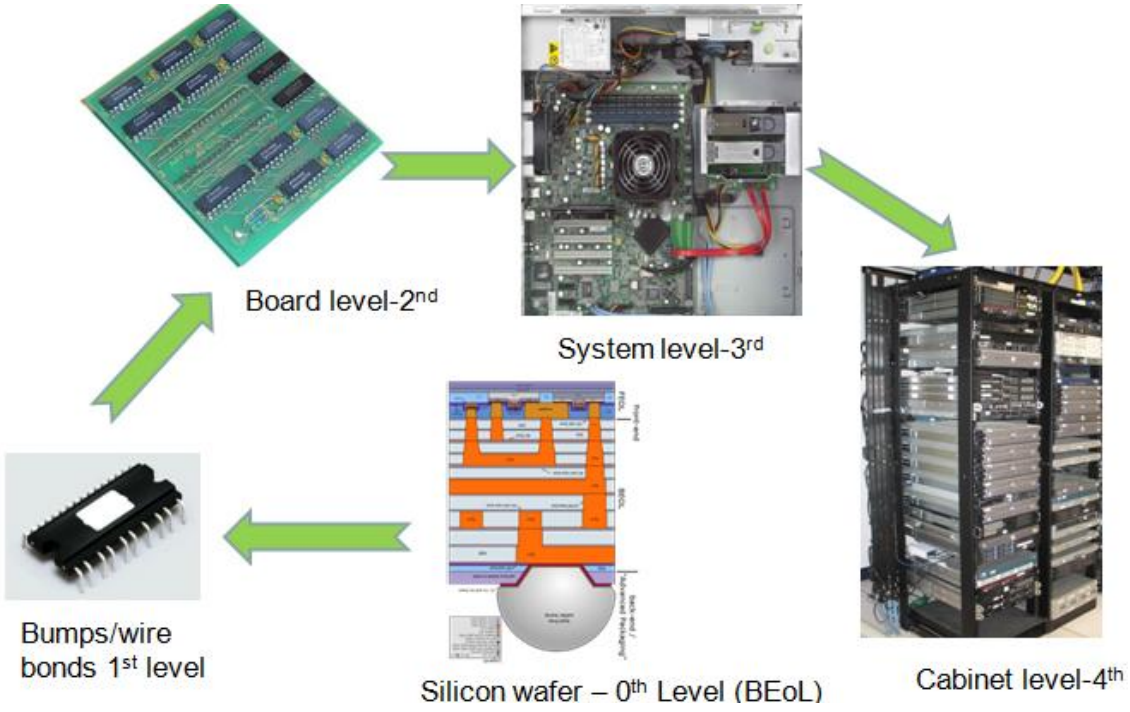


Figure 1.2 Package Hierarchy [3, 4]

0^{th} level: Gate to gate interconnection on Si wafer

1st level: Chip carrier, chip level interconnects that connects chip to substrate

2nd level: Chip-Substrate connected to Printed Circuit Board (PCB)

3rd level: PCB-to-PCB or Card-to-Motherboard connection

4th level: Rack holds several subassemblies (servers) which must be connected to each other to complete system.

In this thesis, study is limited to 0th and 1st level packaging interconnect.

1.1.3. Package Technology

Three breakthroughs in chip package technologies helped in miniaturization and convergence of different features in single electronic device.

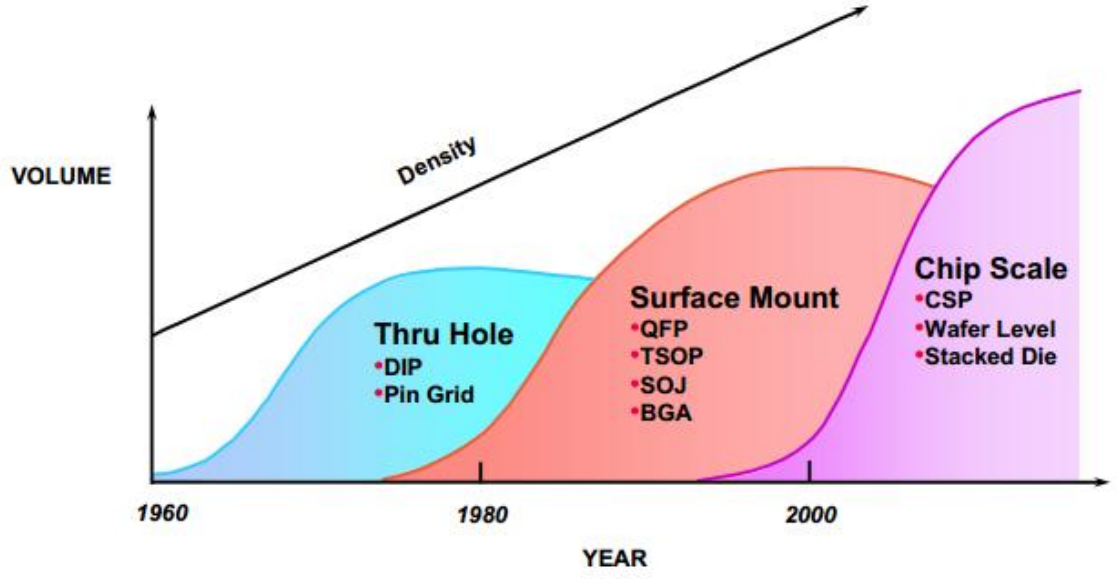


Figure 1.3 Electronics Packaging Technology [5]

Initially Semiconductor industry used Through hole technology to mount package on board in which pins of the component go through the previously drilled holes in PCB as shown in Figure 1.4. In this kind of packaging signal has to pass through all PCB layers and due to low density one sided mounting was available. To fill the gap Surface Mount Technology (SMT) was invented which allowed pins of the device to directly mount on the surface of PCB. Provided much higher density and can be mounted at both sides of PCB as shown in Figure 1.5.

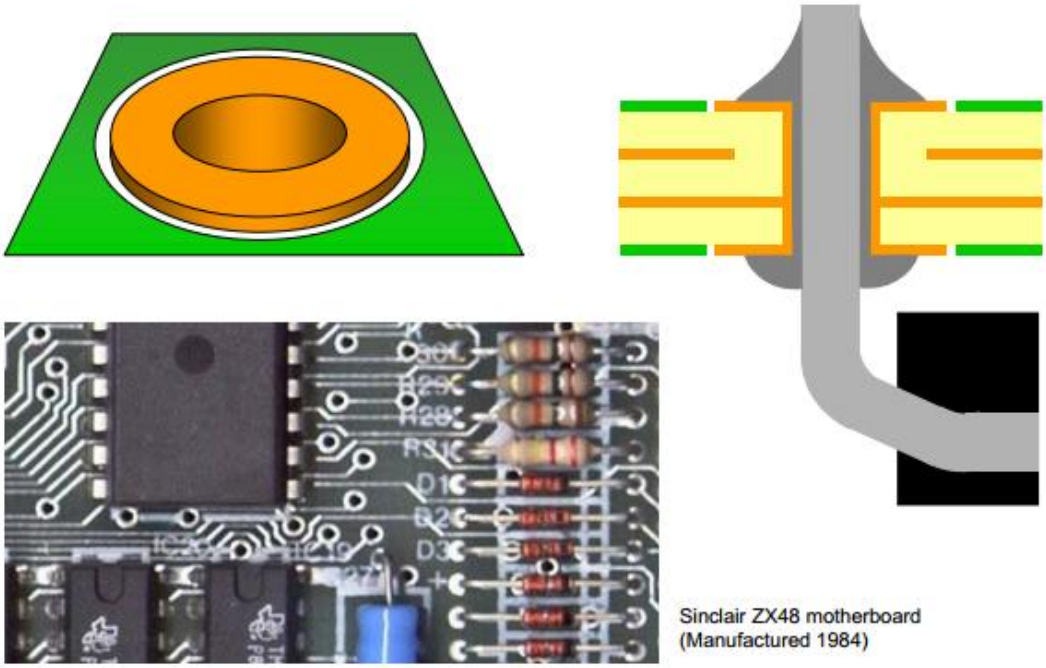


Figure 1.4 Through Hole Mounting [5]

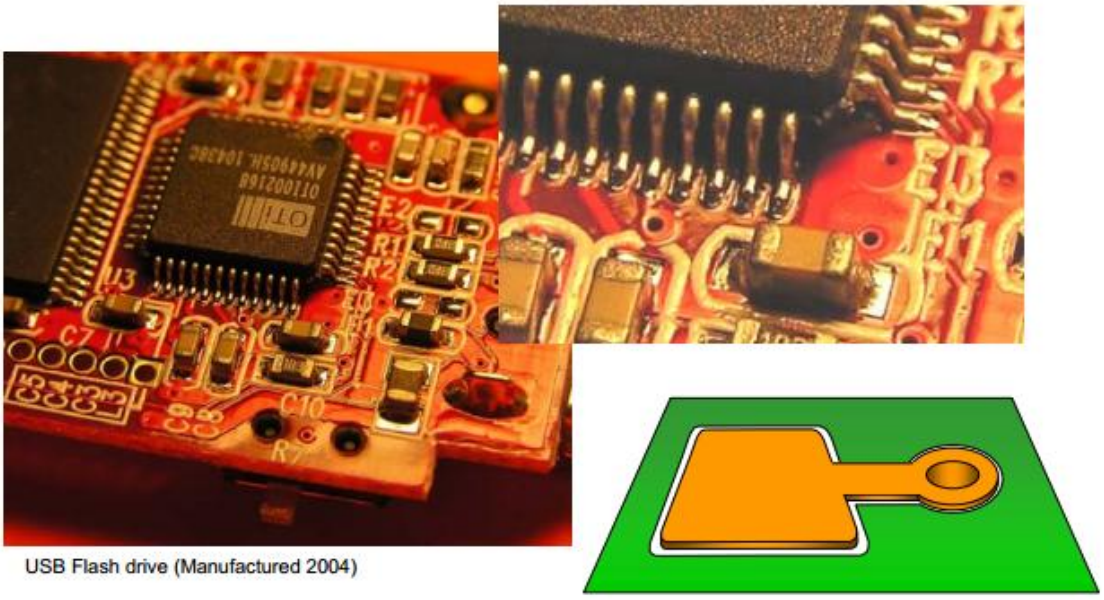


Figure 1.5 Surface Mount Technology [5]

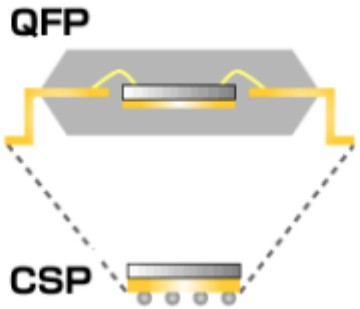


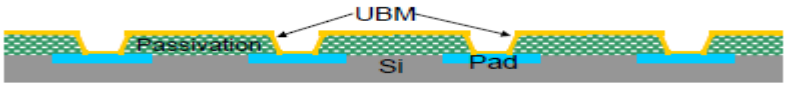
Figure 1.6 Chip Scale Packages [2]

Chip Scale Package, or CSP, based on IPC/JEDEC J-STD-012 definition, is a single-die, direct surface mountable package with an area of no more than 1.2 times the original die area. CSP is an evolution of SMT in which passive components such as resistors, capacitors must also be miniaturized to further improve the speed performance. These packages are hardly serviceable due to difficult fabrication process and there are issues related to long term reliability which need to be taken care of [5].

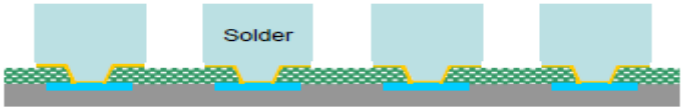
1.2 Flip Chip Package

1.2.1. Flip chip Package Introduction

Flip chip packages fall under Surface Mount Devices (SMD) and the whole study in this thesis is based on flip chip package. Flip chip is very common 1st level interconnect technology which is being used widely due to its benefits over wire bond type packages. In wire bond package number of interconnects are limited as they use the periphery of the silicon die and Interconnect are also longer than typical flip chip interconnects, which increases delay. In flip chip we can use the whole active side of the silicon die to create 1st level interconnects (bumps) and their height is also less compared to wire bond interconnects which reduces the signal path hence faster. This 1st level interconnects are also known as Controlled Collapse Chip Connection (C4) as introduced by IBM in 1964. The typical process flow of flip-chip packaging is described in Figure 1.7.

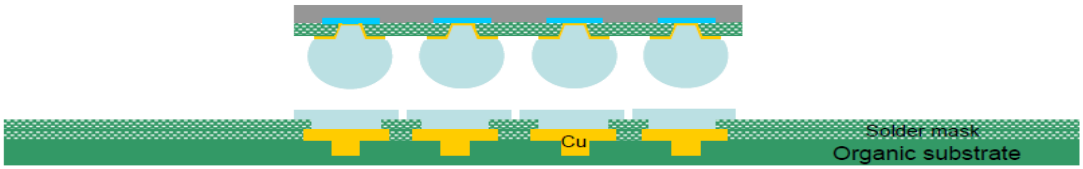


(a)

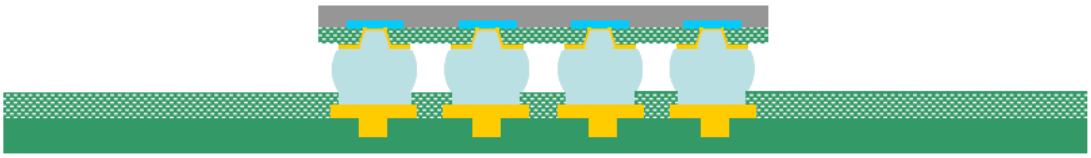


(b)

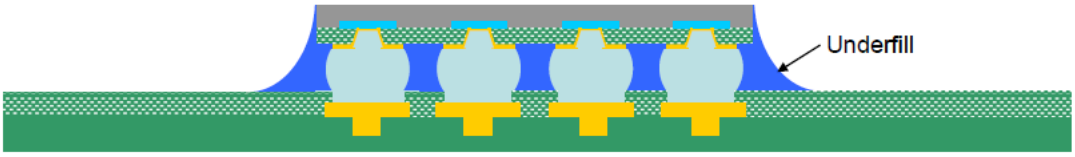
(c)



(d)



(e)



(f)

Figure 1.7 Flip Chip Process Flow [6]

Initially Silicon wafer goes through around 300-400 process steps including the creation of FEoL (Front End of Line) and BEOL (Back End of Line) which will be discussed later. After that completed Si wafer is patterned with passivation layer and then Under Bump Metallization (UBM) layers are electroplated on that pattern (Figure 1.7(a)). After electroplating, parts of UBM which are not required are etched away and solder alloy is deposited on top of the patterned UBM (Figure 1.7(b)). After Solder deposition whole wafer is heated to reflow the bumps to form spherical shape (Figure 1.7(c)). This wafer is then cut into small dies for further packaging steps. Prepared die is then flipped over and aligned with the substrate as shown in figure 1.7(d). All these C4 solder bumps (1^{st} level interconnects) are formed during reflow process and connected to organic substrate (Figure 1.7(e)). Finally the gap between die and substrate is filled up with underfill (polymer) material to reinforce the solder joints (Figure 1.7(f)) [6].

1.2.2. Moore's Law

It was first observed by Gordon Moore (Intel Corporation) in 1965 that number of transistor per unit area of the die is doubling every 18 months with proportionate decrease in cost. This observation is then known as "Moore's law", which predicts the transistor density on ICs for each advance technology nodes. This law has been driving force for the semiconductor industry worldwide. This law explains why technological progress in semiconductor industry has been able to continuously come up with the products that are more smaller, powerful and less expensive than the earlier products [7].

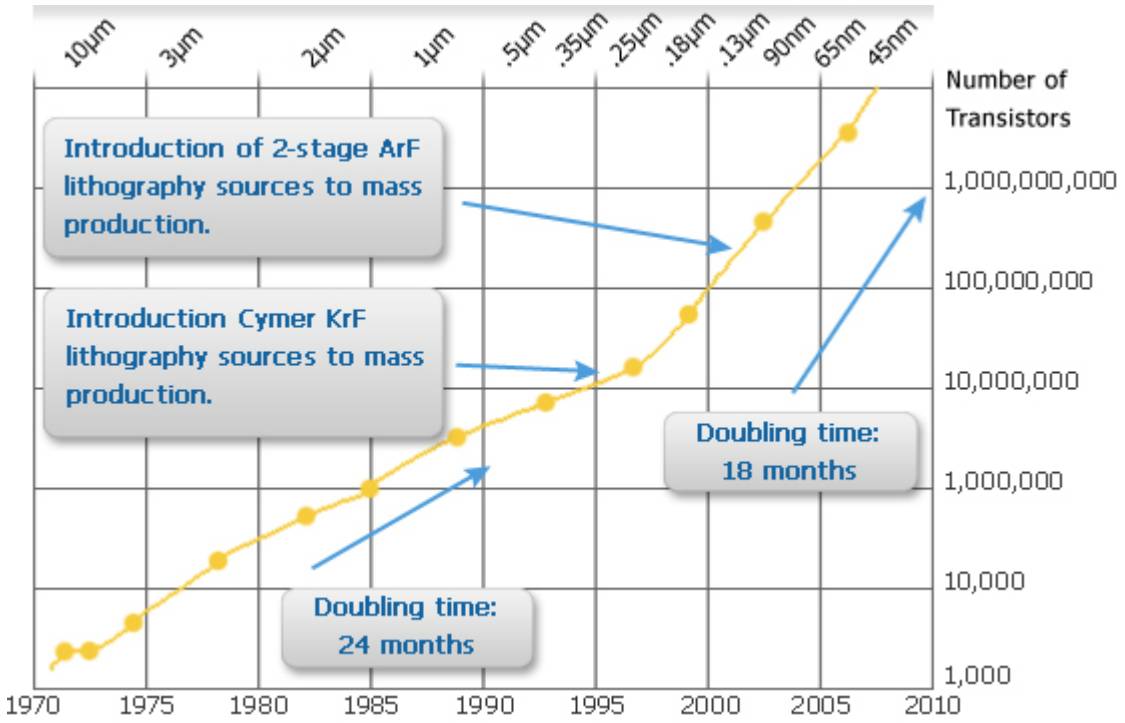


Figure 1.8 Moore's law [7]

1.3 Cu/low-k Interconnects

IC fabrication in itself is a very complex process which includes large number of steps which includes imaging, deposition and etching. This process is mainly divided into two portions. First portion is known as FEoL (Front End of Line). In this process individual components such as transistor, resistor and capacitors are patterned on the wafer. This process also includes the selection of type of wafer to be used, Chemical-Mechanical Planarization (CMP) and cleaning of wafer. This portion includes approximately 300 to 400 process steps containing cleaning, component formation (imaging, deposition and etching) [8].

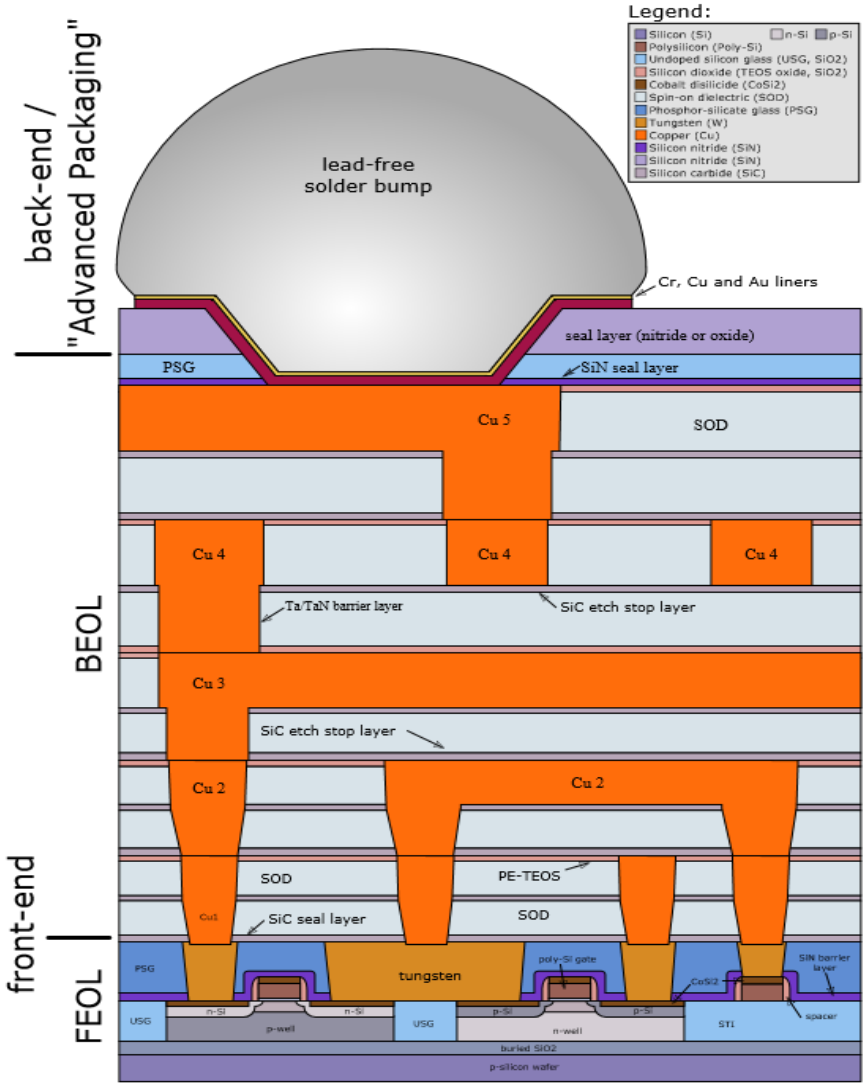


Figure 1.9 FEoL and BEoL/*f*BEoL in typical ICs [9]

As a result of FEoL we have all the components on the Si Chip, but they are all isolated. To function they need some kind of interconnection which starts with the deposition of first metal layer. This Portion of IC fabrication is known as Back End of Line. Thus BEoL is the second portion of IC fabrication, where all components like transistor, resistor and capacitor get connected through different metal layers. Apart from metal layers there are insulating layers (dielectric) deposited in between metal layers. Two metal layers get connected through vias to communicate as per electrical routing [10] (Figure 1.9).

In continuation of Moore's law, Semiconductor industry has increased the power of ICs by reducing its feature size. Along with that industry started altering the material to manufacture high speed devices. To meet up the miniaturization and convergence requirement features like gate length, dielectric thickness and interconnect line width should be scaled down. This process would affect the RC (resistive-capacitive) delay. Apart from that cross talk and leakage current between two metal lines are the critical factors which need to be taken care of during the design [11, 12].

As a result of scaling down feature size is reducing means increased density. This makes the proximity of circuit interconnects closer and big line–to–line capacitance eventually a greater signal delay. This interconnects delay increases with square of reduction in feature size whereas gate delays generally decreases linearly with the same reduction in feature size. Which displays that in contemporary time, speed performance of device is no longer feature size dependent but is dependent on interconnects distance [11].

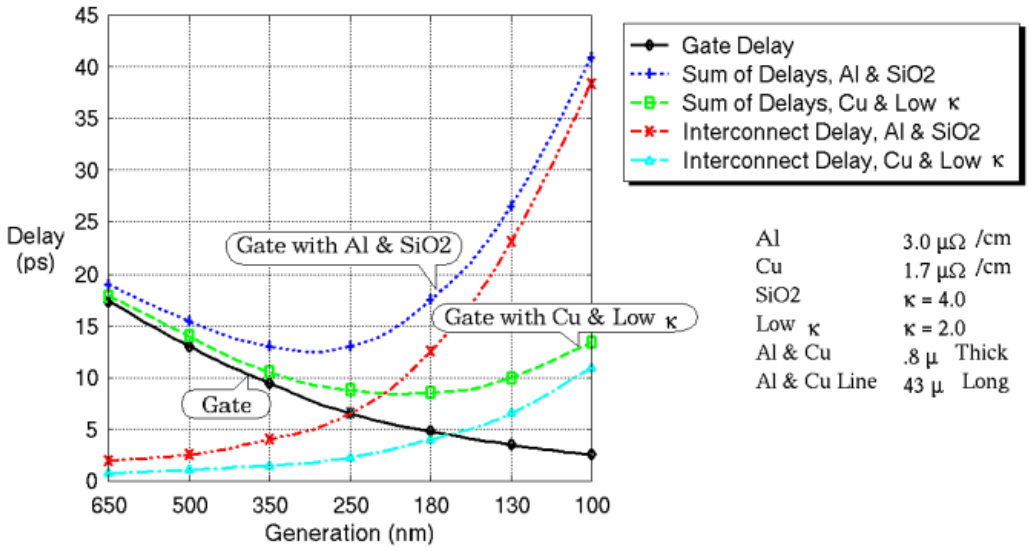


Figure 1.10 Interconnect delay and gate delay for technology generation transition [13]

This increased delay conventionally addressed by adding more number of metal layers, but it increases production cost as well as heat generation in devices which affects reliability and performance of device. To beat this problem and still allow continuation of Moore's law semiconductor industry migrated from conventional Al to Cu which has higher conductivity. Transition from Al to Cu alone improves speed performance approximately around 30%. However if Cu is integrated with low-k / ultra low-k (K<3.9) material in place of SiO_2, speed performance increases significantly around 266%. This significant results made it mandatory for semiconductor industry to replace Al/SiO_2 interconnect technology with Cu/low-k interconnect technology [11]. Figure 1.10 explains how semiconductor industry recognized need to migrate from Al/SiO_2 to Cu/low-k interconnect technology in mainstream electronics devices following latter's impact on RC delay reduction [13].

1.3.1. Reliability Issues with Cu/low-k Interconnects

As technology advances BEoL structure in package continues to evolve with decrease in dimensions and increase in number of layers as well as complexity. In contemporary era, industry is more driven towards integrating low-k (k<3.9) / ultra low-k (k<2.5) with Cu in BEoL to further reduce RC delay. However with reduction in dielectric constant k, mechanical properties of the material deteriorate as more porosity is introduced in material.

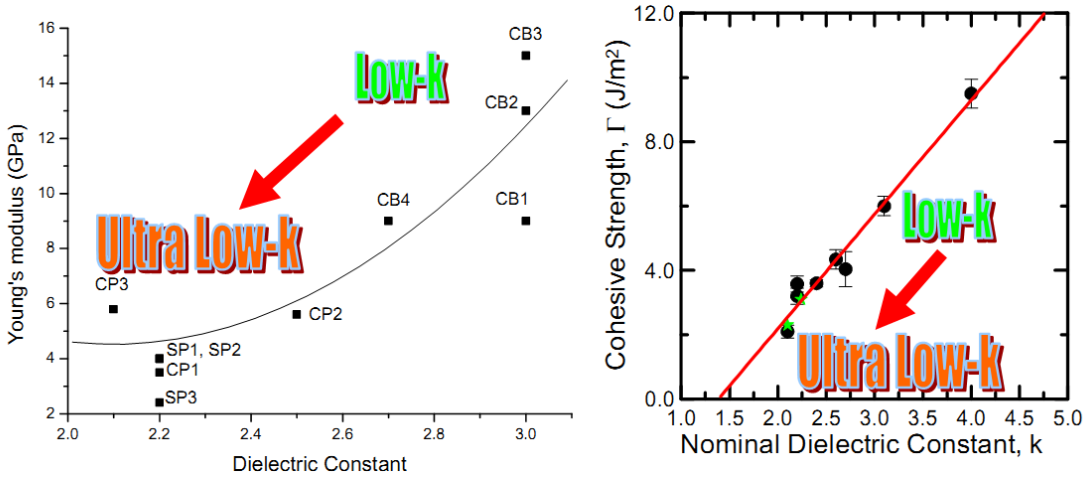


Figure 1.11 Low-k / Ultra Low-k material properties [14]

Figure 1.11 explains that with reduction in dielectric constant k value, material starts becoming mechanically weak. Hardness, toughness and even cohesive strength of material decrease which brings more issues in integrating this materials with Cu in interconnects though they helps in electrical performance (RC delay reduction) of the device. Brittleness and poor adhesion characteristics of low-k dielectric material may brings two kinds of failures during wafer fabrication (sawing dies from wafer) and subsequent assembly process (die attach process). First is cohesive fracture of dielectric and second is interfacial delamination. Due to its brittleness chances are higher that crack is introduced in low-k material during sawing process itself which might propagate in subsequent assembly process due to higher thermo-mechanical stress. Even poor adhesion might lead layers to separate from each other at higher temperature [6].

The main factor which causes reliability issue is CTE (Co-efficient of Thermal Expansion) mismatch between Si die and Organic substrate. Typically most of the parts of flip chip package have high modulus and they are cured at higher temperature such as 100°C to 200°C. So when assembly is cooled down to room temperature, whole structure starts bending because of the CTE mismatch which is called warpage. Due to this warpage bending stresses induces in low-k layers in BEoL, which may cause failure of brittle low-k material. This similar problem may arise during subsequent assembly process as different components (substrate, PCB) get attached to each other at higher temperature [12].

1.4 Motivation of Work

Reliability issues mentioned in section *1.1.7* regarding Cu/low-k interconnect technology needs to be addressed. Structural integrity is the major reliability concern during the fabrication process. To mitigate this CPI (Chip Package Interaction) risk and to reduce crack growth and interfacial delamination there is a need to reduce peeling stress in low-k layers during fabrication process itself. There are certain ways to mitigate this risk of which one possible way is explored in this study.

1.5 Objective and Scope

In this study thermo-mechanical reliability of inter-layer-dielectric (ILD) is examined during die attach – reflow process. Design for reliability explains that failures in microelectronics systems are because of overstress in mechanism or due to wearout (gradual failure even at lower stress-fatigue). As in this study we are focusing on reflow process it's obvious that failure is due to overstress in component rather than fatigue as it is one time loading condition. There are two ways we can design device for long term reliability against failures.

- By reducing the stress that cause failure
- By increasing the strength of material

This can be achieved by selecting/altering material or by changing the package dimensions (skeleton). There are thin chances to change the material of dielectric as it has its electrical preference because of RC delay reduction. So when focusing on design of specific form factor of package, reliability risk can be mitigate by changing the dimensions of other package parts such as bump foot print and substrate thickness.

In this study second possible approach is explored to meet the thermo-mechanical reliability issues during reflow process by varying the group of variables.

1) Solder bump foot print
2) Substrate thickness

A 3-D Multi-Level FE (sub modeling) technique has been leveraged to carry out whole study in ANSYS Workbench v14.5. Initially parametric study has been carried out focusing on the damage parameters in BEoL (0^{th} level interconnect - peeling stress in low-k layers) and *f*BEOL (1^{st} level interconnect – plastic work in solder bump) with above mentioned parameters. Results of this study exhibit qualitatively opposite behavior in both damage parameters. This explains it is difficult to come up with one specific dimension for solder bump and substrate thickness for specific form factor of package. Answer to this issue is to optimize both design

variables keeping the objective to minimize both damage parameters. This multi objective design optimization technique has been implemented with Design Exploration tool box in ANSYS Workbench v14.5.

CHAPTER 2

LITERATURE REVIEW

Hoofman et al [16] exhibit that with continuous decrease in inter-metal spaces along with the introduction of new porous low-k materials and thinning down the metallic barrier thickness, dielectric reliability might become a serious issue in the interconnect as well. The downscaling of barrier thickness has also a dramatic effect on the copper reliability. It was explained that, some of these challenges can be dealt by continuous innovation of materials, processes and integration approaches. It is however also possible that different solution, such as air gaps and low-k backfill, might also be considered to remove speed-sensitive paths from minimum-pitch portions of circuits, whenever possible, in order to profit better from the advantages of copper and low-k materials, without running into their associated critical areas.

Michel et al [17] explains that with reduction in dimensions and increment in complexity of interconnects, process development alone will bring a limited benefit in performances, yield and reproducibility of the final product. At the same time the design of complex circuits is confronted to issues related to the materials and processes introduced in manufacturing. A concurrent and synergetic development of process-aware design techniques and of product-driven processes is mandatory for achieving the expected progress in scaling dimensions further.

Fu et al [18] discussed CPI reliability issues, impact of crack stop design, mechanical properties of low-k /ultra low-k material, underfill material selection, lead-free manufacturing and low-k/ultra low-k layers on CPI reliability. In addition it has been explained that by optimizing the various parameters such as design stronger crack stop, improve the adhesion and cohesive strength of low-k/ultra low-k materials, select right underfill material , reliable low-k/ultra low-k devices can be manufactured.

Uchibori et al [19] investigated the effect of CPI on Cu interconnects reliability for a number of low-k materials including two developed technology nodes using 3D multi-level sub-

modeling method. It was found that the mechanical reliability of Cu interconnects deteriorates possibly by six times as generation of technology nodes advances from 90 nm to 65nm.

Wang et al [20] studied CPI for Cu interconnects using 3D finite element analysis based multi-level sub-modeling technique. Packaging induced crack driving forces for relevant interfaces in Cu/low-k structures were deduced and compared with corresponding interfaces in Cu/TEOS to assess the effect of ILD on packaging reliability. It was investigated that the packaging effect due to die attach process where a high thermal load occurs during solder reflow before underfilling. In order to study the influence of low-k properties on packaging reliability both SiLK and MSQ dielectrics were investigated. This effect found to be intense with lead-free solder than the eutectic lead solder and the high lead solders. Packaging effect is generally lower for the Cu/MSQ structure than the Cu/SiLK structure and the difference can be attributed to the higher Young's modulus of the MSQ material.

Zhai et al [21] demonstrated the experimental results which supports that Chip Package Interaction related risk increases as we move towards low-k material which are lower modulus. Analysis indicated that package level solution such as underfill material selection, impact specific failure mode alleviate the CPI related risks. Risk of corner delamination can be reduced by selecting lower modulus underfill material. Other side thermal expansion of underfill leads to Cu/low-k film delamination. This study can be uased a s guidance to material selection as well as BEoL stack up configuration to reduce CPI related risks.

Kuo et al [22] investigated thermo-mechanical and mechanical properties of polyimide material which is widely used in microelectronics packaging. Young's modulus and stress-strain curve of this material presents strong visco-plastic behavior which is highly temperature and srain rate depended. It was seen that power law breakdown phenomenon occurs at higher stresses as the temperature increases. It was seen that this polyimide material exhibit softening and less creep under cyclic load. There was11% difference in uniaxial tensile test and nano indentation test.

Wang et al [23] performed sub-modeling analysis for path dependent thermo-mechanical issue and he benchmarked this method on bimetallic strip of two commonly used electronic packaging materials such as silicon and eutectic solder alloy. Comparison of global model with simplified geometry and sub- model with detailed geometry was carried out under thermal fatigue life assessment. Results found more close to experimental data in case of sub-modeling. In addition to that it has been explained that accuracy of sub-model depends on the scale and local feature of concern should be far away from sub-model boundaries.

Che et al [24] performed reliability assessment on low-k structure using 2D finite element method and its results were compared with 3D analysis to verify accuracy of 2D approach. Group of parameters such as solder bump material, low-k stacks, UBM thickness, barrier layer thickness and shear heights were used to conduct parametric study. Shear force and interfacial peel stress were compared between Sn–Ag based lead free solder and high Pb solder bumps. It was found that 15 layers low-k stack is safer than 13 layers low-k stack. It was also observed that thinner UBM can results as interfacial stress reduction, induced by solder deformation. Shear arm height was taken higher than UBM thickness in both simulation and experiment.

Che et al [25] conducted calibration study to develop global-local modeling method containing sub-modeling and global-local beam (GLB) technique. Sub-model with only solder bump found to be less accurate than sub-model with hybrid solid, solder bump, PCB and component. It was seen that, solid-beam-solid GLB model is more accurate than shell-beam-shell GLB model and modified shell model comes with less error than non-modified shell elements. It explains that structure where only elastic behavior of solder is considered such as small deformation issue or modal analysis, GLB model works fine. Real ball shape of solder joint was replaced by effective cubic shape joint considering that they have similar stiffness based on GLB theory. Effective cubic shape joint provides more accurate results with solder joint plastic deformation.

Lai et al [26] verified sub-modeling technique in flip chip BGA through plane two-dimensional analysis for the thermo-mechanical reliability assessment. Sub-model was constructed to enclose outer most solder bump to focus on its thermo-mechanical response. It was noticed that dimensions of sub-model in this study was applied appropriately. Moreover, when boundary of the sub-model is more away from the enclosed bump, good prediction can be done for the fatigue indices through sub-modeling. Due to incapability of computers to perform 3D analysis with detailed solder bump implementation, sub-modeling technique was applied in particular path dependent thermo-mechanical problem.

Lofrano et al [27] studied temperature dependence of stress induces voiding with both 2D and 3D finite element built for three different Back End of Line interconnect structures. Structural dependency was first studied in 2D axisymmetry model but limitation was that, stress gradient in asymmetrical plane were underestimated. 3D model showed higher gradient hence larger driving force for void nucleation. To identify the effect of dielectric on stress gradient in via, additional analysis was performed. The analysis exhibited that, when material with lower elastic modulus (low-k) is used than SiO_2, higher stress gradient is seen in via.

Auersperg et al [28] explains that more optimization based on parametric study and finite element analysis is being used at early stage of product development process to design for thermo-mechanical reliability. Apart from this, reduction in feature size and introduction of high-tech porous material in micro-electronics packages brings new reliability challenges. He focuses up to multi scale modeling approaches, damage and fracture mechanics approaches on the basis of continuum mechanics and measurement techniques of material properties in the miniaturized range addressed. Author explains that multi material micro-electronic system not only requires variety of environmental loading and combined multiple failure criteria, but intrinsic stress situation from previous technological step. Author proposed new technique based on stress release FIB milling and high-resolution displacement measurement.

Lau et al [29] studied optimization of the reflow soldering process with multiple quality characteristics such as thermal stress, peak temperature, reflow time and temperature uniformity of board level BGA packages using the grey relational analysis based on Taguchi method. In this study internal flow of reflow was modeled in CFD and structural heating BGA was created using FEM code. To compute the weight of each characteristic entropy measurement method was implemented along with grey based Taguchi method. It was figured out that Taguchi method is not suitable for solving the problem of multiple characteristics soldering process. A simple analysis corresponding the cooling time of 40 s, preheating slope of 1.5 K/s, soaking slope of 0.3 K/s, ramp slope of 1.0 K/s, cooling slope of 1.0 K/s, soaking temperature of 413 K, and peak temperature of 518 K was carried out to reach the optimal setting of the infrared-convection reflow oven. It was found that cooling slope is the most influential factor that can reach minimum soldering defects followed by soaking temperature and peak temperature. The grey relation grade was significantly improved by 46.6 % for optimal condition and implementation of current effective approach helped in reducing soldering defects during reflow process.

Xu et al [30] developed multi-objective design optimization method for electronics package reliability and it is based on simulation and capable of improving the design in provided design space itself. Use of RSM makes it even simple for complex and highly non-linear FE models. As results of this optimization is computational based, need arise to verify optimal design point by experimental tests. However this simulation based optimization tool guides us to reliability improvement from product design phase itself. This method could be useful to potentially eliminate sever reliability issues during design phase and helps in reducing product cycle and cost.

Hossain et al [31] performed optimization of generic PWB-level package using inbuilt FEM tool in ANSYS. Life of solder joints were estimated using Darveaux's method for package reliability analysis. Three types of optimization process were carried out to study influence of

each design variable on fatigue life of solder joint under thermal cycling. Factorial method gave 29.1% improvement in fatigue life. The factorial study in this tool employs two level techniques to sample two extreme points of each design variable. To find optimum design point in design space, author suggest central composite design as it uses response surface approximation and reduce the simulation time significantly for approximating a huge number of simulations. This study supports that CCD based response surface methodology satisfies all goodness of fit measures done using various statistical tool for given response output parameters. It was observed that coefficient predicted for each RS polynomial provides a good fit for the given output response parameter. Through all methods it was observed that solder ball standoff height and CTE of PWB have a significant influence on the fatigue life of a solder joint.

CHAPTER 3

FINITE ELEMENT MODELING

3.1 Introduction to Finite Element Method (FEM)

The Finite Element Method is the computational technique used to obtain approximate solution to boundary value problem in engineering. FEM is virtually used in almost every industry that can be imagined. Common application of FEA applications are mentioned here.

- Aerospace/Mechanical/Civil/Automobile Engineering
- Structural Analysis (Static/Dynamic/Linear/Non-Linear)
- Thermal/Fluid Flow
- Nuclear Engineering
- Electromagnetic
- Biomechanics
- Geomechanics
- Biomedical Engineering
- Hydraulics
- Smart Structures

"The Finite Element Method is one of the most powerful numerical techniques ever devised for solving differential equations of initial and boundary value problems in geometrically complicated regions" [32]. Sometimes it is hard to find analytical solution of important problems as they come with complicated geometry, loading condition, and material properties. So FEA is the computational technique which helps in reaching the satisfactory results with all the complex conditions that can't be solved through analytical procedure. There are wide range of sophisticated commercial code available which helps in reaching the approximately close solution in 1D, 2D and 3D. In this FEA method, the whole continuum is divided into a finite numbers of small elements of geometrically simple shape. These elements are made up of numbers of nodes. Displacement of these nodes is unknown and to find it, polynomial

interpolation function is used. External force is replaced by an equivalent system of forces applied at each node. By assembling the mentioned governing equation, results for the entire structure can be obtained.

$$\{F\} = [K]\{u\}$$

Where, {F} = Nodal load/force vector

[K] = Global stiffness matrix

{u} = Nodal displacement

Structure's stiffness (K) depends on its geometry and material properties. Load (F) value has to be provided by user. The only unknown is displacement (u). The way in general FEA works is, it creates the number of small elements with each containing few nodes. There are equations known as Shape function in software, which tells software how to vary displacement (u) across the element and average value of displacement is determined at nodes. Those stress and/or displacement values are accessible at nodes which explains that finer the mesh elements, more accurate the nodal values would be. So there are certain steps that we need to follow during the modeling and simulation in any commercial code to reach approximately true solution, which would be explained hereby [12]. In this study commercially available FEA tool, ANSYS Workbench v14.5 has been leveraged.

3.2 FEA Problem Solving Steps

These five steps need to be carefully followed to reach satisfactory solution to FEA problem.

1) Geometry and Material definition
2) Defining Connection between bodies
3) Meshing the model
4) Defining load and boundary condition
5) Understanding and verifying the results

The ANSYS is the general purpose FEA tool which is commercially available and can be used for wide range of engineering application. Before we start using ANSYS for FEA modeling and simulation, there are certain set of questions which need to be answered based on observation and engineering judgment. Questions may be like what is the objective of analysis? How to model entire physical system? How much details should be incorporated in system? How refine mesh should be in entire system or part of the whole system?

To answer such questions computational expense must be compared to the level of accuracy of the results that needed. After that ANSYS can be employed to work in an efficient way after considering the following.

- Type of problem
- Time dependence
- Nonlinearity
- Modeling simplification

From observation and engineering judgment, analysis type has to be decided. In this study the analysis type is structural; to be specific out of different other structural problem focus in this study is on Static analysis. Non-linear material and geometrical properties such as plasticity, contact, and creep are available.

3.2.1 Geometry and Material Definition

Geometrical nonlinearity needs to be considered before analysis. This nonlinearity is mainly of two kinds.

1) Large deflection and rotation: If total deformation of the structure is large compared to the smallest dimension of structure or rotate to such en extent that dimensions, position, loading direction, change significantly, then large deflection and rotation analysis becomes necessary. Fishing rod explains the large deflection and rotation.

2) Stress Stiffening: Stress stiffening occurs when stress in one direction affects the stiffness in other direction. Cables, membranes and other spinning structures exhibit stress stiffening.

Material nonlinearity is also the critical factor of FE analysis, which reflects in the accuracy of the solution. If material exhibit linear stress-strain curve up to proportional limit or loading in a manner is such that it doesn't create stress higher than yield values anywhere in body then linear material is a good approximation. If the material deformation is not within the loading condition range is not linear or it is time/temperature dependent then non linear properties need to be assigned to particular parts in system. In that case plasticity, creep, viscoelasticity need to be considered. Apart from that if structure exhibit symmetry in geometry, then it needs to be considered when creating model of physical structure which is advantageous in saving the computational time and expense [33]. Once the geometry and material properties are taken care of contacts between different bodies needs to be considered such as rigid, friction, bonded etc.

3.2.2 Meshing Model

As discussed in section 3.1, large number of mesh counts (elements) provides better approximation of solution. There are chances in some case that excessive number of elements increases the round off error. It is important that mesh is fine or coarse in appropriate region and answer to that question is completely dependent on the physical system being considered. In some cases mesh sensitivity analysis is also considered to balance computational time with accuracy in solution. Analysis is first performed with certain number of elements and then with twice the elements. Then both the solutions are compared, if solutions are close enough then initial mesh configuration is considered to be adequate. If solutions are different then each other then more mesh refinement and subsequent comparison is done until the convergence is achieved [34]. There are different types of mesh elements for 2D and 3D analysis in ANSYS which are mentioned below.

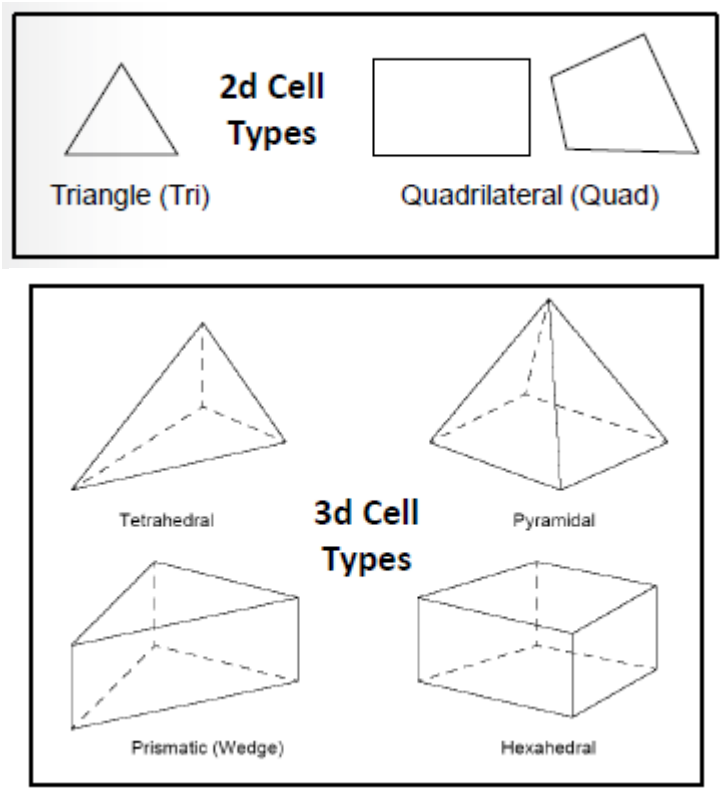


Figure 3.1 2D and 3D elements type [34]

There is a need to look at the aspect ratio of the elements when defining mesh for 2D and 3D geometry. Aspect ratio of the element is the ratio of largest and smallest dimension. In case of elongated and skinny elements aspects ratio becomes higher. Elements with higher aspect ratio not necessarily produce bad results however that depend on the type of loading and boundary condition but they introduce potential trouble in convergence of solution [35].

3.2.3 Boundary Conditions

Limitations set to the problem are known as boundary condition. In order to solve the problem these limitations are mandatory, without that the system would be assume as a rigid body. These boundary conditions (limitations) helps software to understand where element is likely to move and where it is restricted. In absence of boundary conditions, system would float in space as a rigid body without deformation under acting load. So when assuming boundary conditions, it needs to be making sure that system is constrained enough to have deformation to

prevent rigid body motion. In addition to that system should not be under constrained or over constrained to get convergence to the solution [31].

3.3 Submodelling

To get the most accurate results in region of interest out of your whole system sub modeling technique is used. In FE analysis it may occur that mesh is too coarse to provide satisfactory results in the area of interest where stress in higher. Sub modeling is sometimes known as global-local analysis or cut boundary displacement method. Cut boundary is the boundary of the sub model where it has been cut through global model. Displacement calculated on the boundary of the cut from global model is applied as a boundary condition for the sub model at cut boundary planes. Figure 3.2 explains how area of interest (high stress) from global model of pulley hub and spokes is differentiated in sub model.

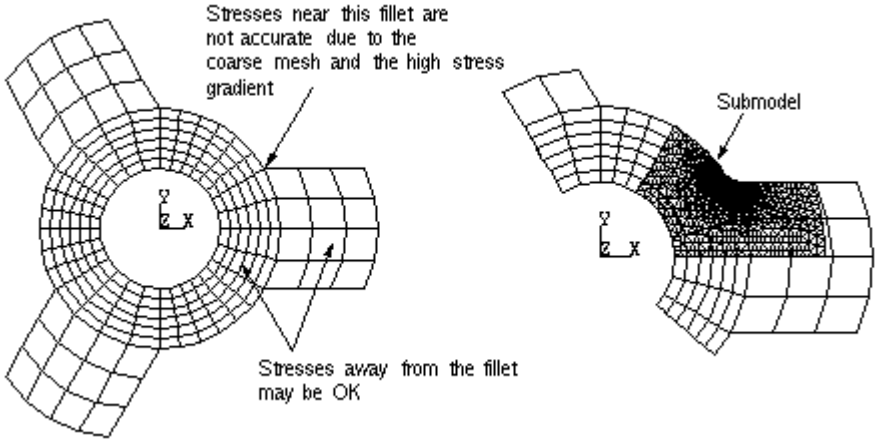


Figure 3.2 Sub modeling of pulley hub and spokes [36]

St. Venant's principle supports sub modeling technique. It states that if actual force distribution is replaced by statically equivalent system, the distribution of stress and strain is altered only near the region of load application. This explains that stress concentration effects are localized around the concentration, so if the boundary of the sub model is far enough away from the area of interest, reasonably accurate results can be calculated in the sub model.

Apart from just the accuracy there are other benefits of sub modeling, which are stated here.

- ✓ The need for transition region in solid FE models is reduced or eliminated.
- ✓ It allows you to experiment on different design and area of interest.
- ✓ It helps you in getting adequate mesh refinement.
- ✓ You can independently tackle sub model, even geometry modification or improvement can also be done.

There are some limitations when implementing sub modeling is that, it can be used for only solid and shell elements and assumption is taken that cut boundaries are far away from area of interest [36].

CHAPTER 4

FINITE ELEMENT MODELING AND METHODOLOGY

4.1 Modeling of Flip Chip Package

In this chapter Finite Element modeling and methodology is discussed which has been carried out to analyze the influence CPI on 0^{th} (low-k layers stacked in BEoL) level and 1^{st} (solder bumps in *f*BEoL) level interconnects. Whole study is done after the reflow process which is shown in figure 1.7(d). In reflow Silicon die is being attached to the substrate at higher temperature (approx 200°C), so sometimes it is known as die attach process. After die is attached to the substrate, whole assembly is brought down to the room temperature. While bringing down the whole assembly to the room temperature, structure starts bending due to the CTE (Coefficient of Thermal Expansion) mismatch of Si die and organic substrate. In microelectronics industry this phenomena of structure bending is known as warpage. This warpage induces higher thermo-mechanical stresses in the structure at all the packaging level. To analyze this thermo-mechanical stresses and their effect on package reliability at different level of interconnects (0^{th} & 1^{st}) is studied using commercially available code ANSYS Workbench v14.5. Since our area of interest in flip chip package is cu/low-k layers (nm) and solder bumps (µm) which are couple of magnitude lower than the rest of the package components such as die and substrate, which makes the Finite Element Modeling difficult. To bridge this large dimensional difference multi-level sub modeling (global-local) simulation approach has been leveraged. In this study 3D finite element model was created and analyzed at two different levels using multi-level sub modeling technique. A flip chip package with 7 mm x 7 mm die size, 150 µm bump pitch, 16 mm x 16 mm substrate was built and analyzed during reflow process to determine thermo-mechanical behavior. There are certain assumptions which have been made to carry out this finite element analysis.

- All the parts in 3D package is assumed to be bonded to each other

- Temperature change in package during reflow process is assumed to be same throughout the package
- Except solder bump, all other materials are assumed to behave as linear elastic [12].

4.2 Package components

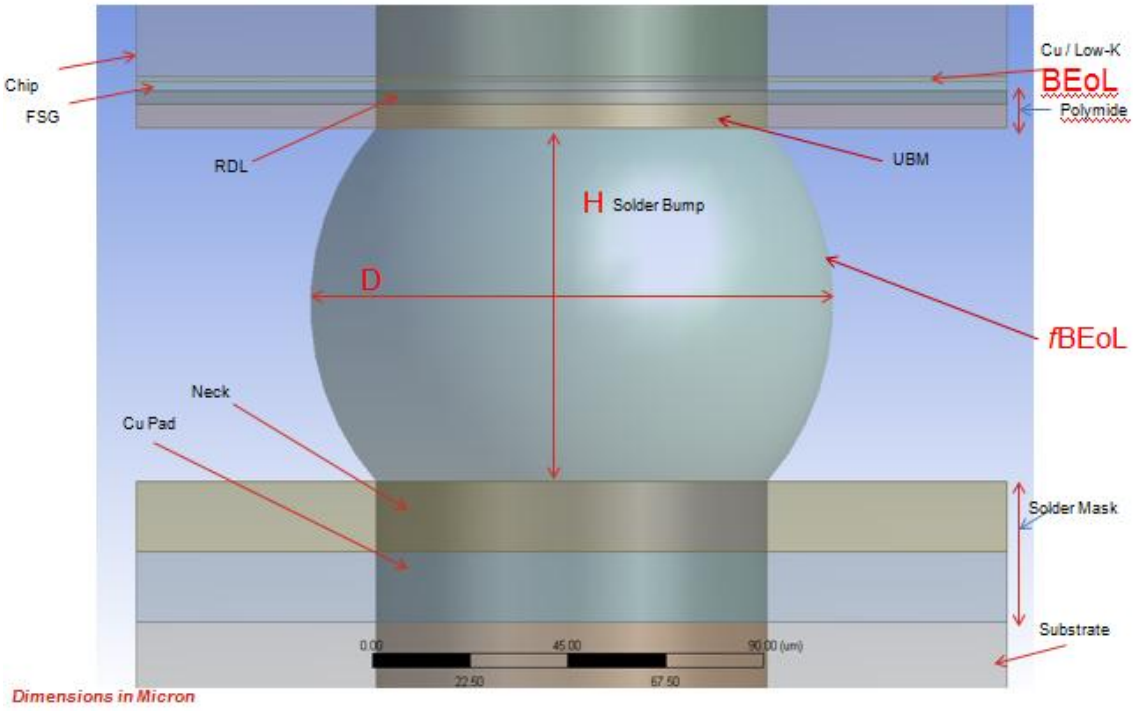


Figure 4.1 Package Components (Global)

Package component details in global model have been considered as shown in Figure 4.1. Importance of some of the package components is explained here. Conventionally Aluminum (Al) was used as a final metal layer in most IC bond pad which provides satisfactory surface for wire bonding but in flip chip packages they doesn't work that well for solder bump. Al creates oxide insulator immediately upon exposure of air which jeopardize bumps electrical connection. To cope up with these issues higher numbers of electronics packages are fabricated with Under Bump Metallization (UBM). This UBM can be deposited in conventional wafer fabrication facilities. UBM adheres well with Al as well surrounding passivation layers. It

provides good wettable surface for solder bump to reflow. To meet this requirements multiple layers of different metal is deposited to create UBM. This UBM includes adhesion layer, diffusion barrier layer, solderable layer, and an oxidation barrier layer, which are being made of Titanium (Ti), Cooper (Cu) and Nickel (Ni) [37].

Initially Redistribution Layers (RDL) helped solder bumping on the die which was originally made for wire bonding. RDL helps in converting peripheral bond pads of wire bond to area array bond pads. In some cases RDL provides a way to distributed power and ground contacts. It also helps in converting off-chip connection from chip scale to board scale. Wafer-level-chip-scale-packages usually redistributed to Ball Grid Array (BGA) pads [38].

Flourinated Silicate Glass (FSG) is the last layer of dielectric material which plays critical role in package reliability and environmental condition. It is a dielectric layers which covers the bond pad and protects the semiconductor beneath the bond pad. It also helps prevent moisture penetration to underlying semiconductor and improves the reliability.

4.3 Sub Modeling

Initially global package was modeled and solved to determine the magnitude and behavior of package warpage. To save the significant amount of computational time, octant symmetry (Figure 4.4) of package has been considered to represent whole package. This geometry has been considered as large deflection geometry as chances are there that overall deflection of the structure is more than 10% of structure thickness. Since BEoL region of Cu/low-k layers is too small in thickness compared to other components in global model, effective Cu/low-k block has been employed with effective material properties.

Most critical bump is diagonally corner bump, which was indentified after the global model (Figure 4.4) analysis. As the stress concentration in corner bump is high, that particular part must be analyzed independently to assess CPI risk. This means individual sub model need to be analyzed in that area to determine accurate response. To create sub model, solder bump

pitch has been used as width and breadth of sub model [12]. Cut boundary displacement has been applied as a boundary condition of cut boundary planes of sub model (Figure 4.5).

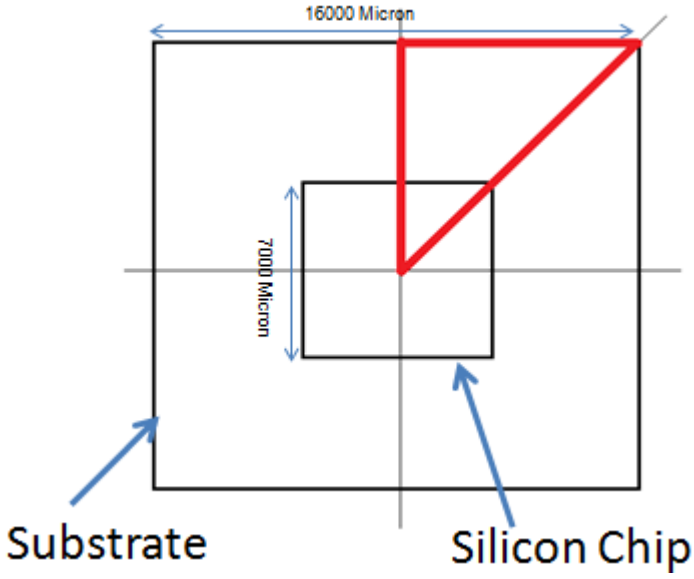


Figure 4.2 Octant symmetry of global package.

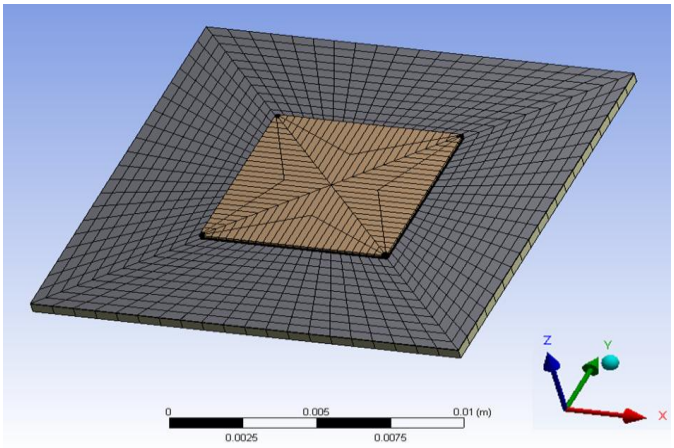


Figure 4.3 Global Model

To further focus on Cu/low-k layers in BEoL region, sub model 2 was modeled as shown in Figure 4.6. To simplify the geometry of Cu and low-k layers, vias were not considered

in modeling but effective properties were given to each Cu and low-k layers. Such total 10 layers (5 Cu and 5 low-k alternate layers) were considered in BEoL region.

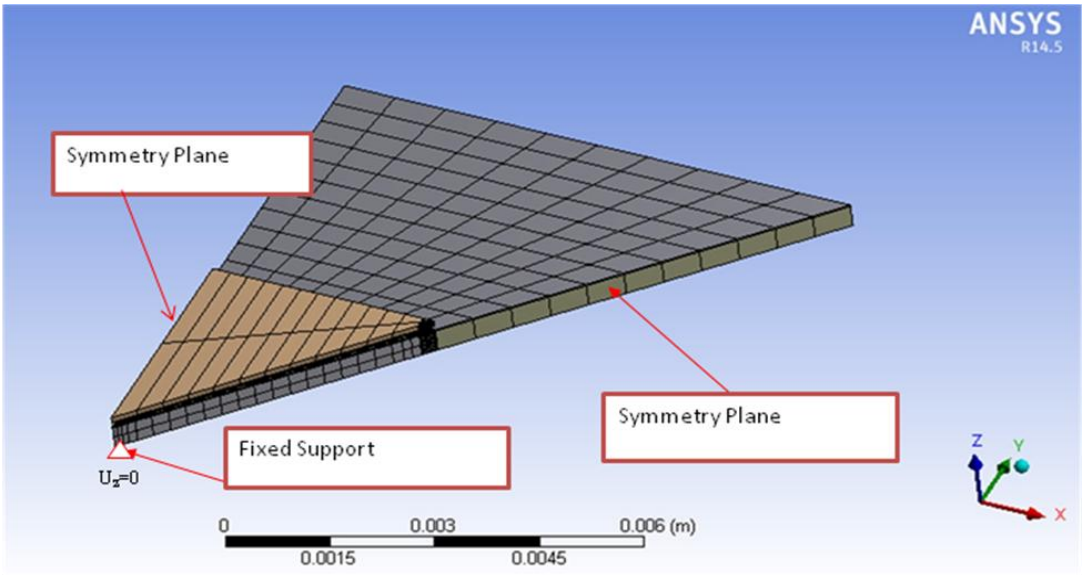


Figure 4.4 Octant Model with Boundary Conditions

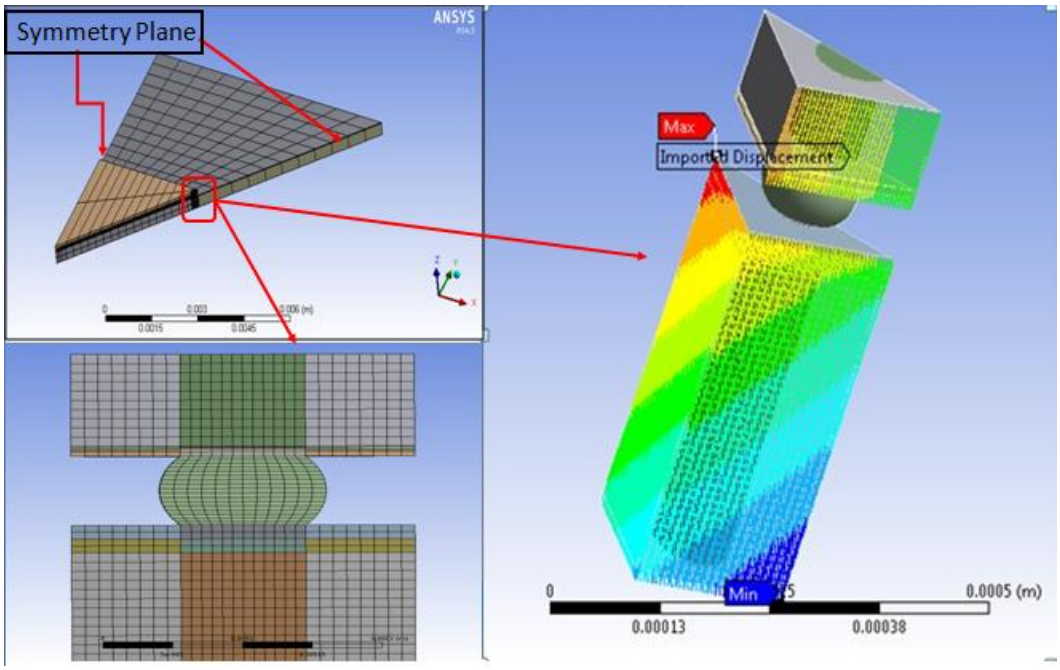


Figure 4.5 Sub model 1 with cut boundary displacements

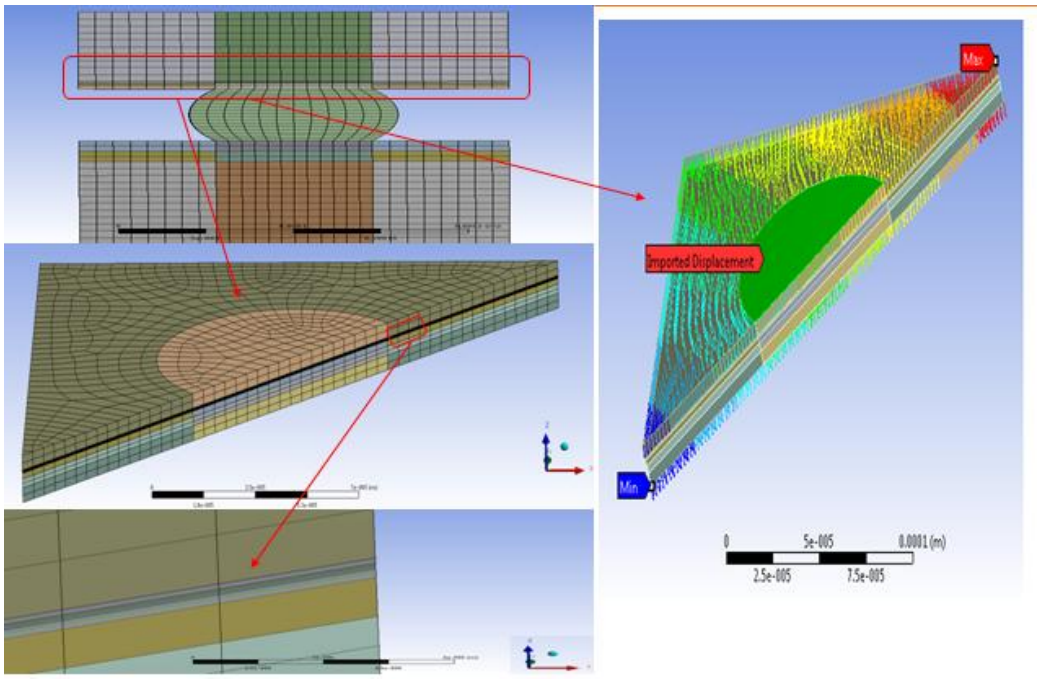


Figure 4.6 Sub model 2 (BEoL) with cut boundary displacements

4.4 Meshing, Boundary Condition and Loading Condition

The global model was discretely meshed using different meshing option in ANSYS Workbench v14.5. Special meshing operation was carried out to make sure that model has mesh continuity throughout the thickness of package in the very far corner unit cell as it is the region of interest. Rest part of the packages where meshed in such a way that, reasonable mesh continuity is achieved in less time so that reasonable response can be captured through nodal averaging in ANSYS. Mesh refinement was carried out in sub model 1 and sub model 2 as our response parameter (damage) belongs to these two sub models. Mesh sensitivity analysis has been implemented in this two sub models to reach the maximum accuracy.

Boundary conditions have been imposed as shown in figure 4.4. Two symmetry boundary conditions is applied to two symmetry boundary planes of the octant symmetry model and in order to prevent rigid body motion further constrain was applied at very center bottom of the package by restricting the vertical movement [39]. The model has been simulated to study the

damage in BEoL/*f*BEoL regions during die attach reflow process of 200°C to room at 30°C per minute

4.5 Package Dimensions and Material Properties

Detailed geometry, package dimensions and material properties has been referred from the literature [40]. Detailed package dimensions are as shown in Table 1. Materials for all the parts of structure have been considered as linear elastic from literature [40] and shown in Table 2 except the solder bumps. To capture the creep and plasticity in solder bump, temperature and time dependent nonlinear properties has been leveraged by Anand's constants for SAC 405 alloy [6] as shown in Table 3. To save significant amount of computational time octant symmetry of the package has been solved [39].

Table 4.1 Detailed package dimensions (mm) [40]

Substrate	16 x 16
Solder Mask	16 x 16 x 0.03
Silicon	7 x 7 x 0.1
Polyimide	7 x 7 x 0.008
FSG	7 x 7 x 0.002
Cu/low-k block	7 x 7 x 0.001
Copper Pad Thickness	0.015
Neck (Height)	0.015
UBM (Thickness)	0.005
RDL (Thickness)	0.003
Micro Bump pitch	0.175

Table 4.2 Material properties [40, 41, 42]

Material	**E (GPa)**	**Poisson's Ratio (υ)**	**CTE (/°C)**
Substrate	24.5	0.37	17e-6 (In Plane)
			52e-6 (Out of Plane)
Solder Mask	4	0.4	52e-6
Die	131	0.28	3e-6
Effective Cu/low-k block	21.3	0.34	21e-6
FSG	70	0.16	0.25e-6
Polymide	4	0.35	57e-6
Cu Pad	110	0.34	17e-6
UBM	120	0.32	14e-6
RDL	110	0.34	17e-6
Low-k Layer	13.2	0.3	25e-6
Cu Layer	56.2	0.3	18e-6

Table 4.3 Anand's constants for SAC 405 solder alloy [6]

Variable	**Sn4Ag0.5Cu**
S_o (MPa)	20
Q/R (1/K)	10561
A (1/s)	325
A	10
M	0.32
h_o (MPa)	8.0E5
S^(MPa)	42.1
N	0.02
A	2.57

4.6 Design of Experiment

This study has been carried out to mitigate the CPI risk and minimize damage in BEoL and *f*BEoL by optimizing the group of critical parameters like substrate thickness and solder bump foot prints. Initially parametric study has been carried out to see if optimization is really

necessary. Design space for both the input parameters for this parametric study has been selected considering the most recent 28 nm technology node and it is within the fabrication capabilities of current semiconductor industry equipments.

1) Substrate thickness varied from 200 µm to 400 µm
2) Solder height has been varied from 50 µm to 75 µm , keeping the diameter by height (D/H) ratio constant as after reflow solder paste takes uniform bump shape where D/H ratio remains almost same.

Design of experiment for the parametric study is shown in table 4.

Table 4.4 Design of experiment

Case	Substrate Thickness (um)	Solder Bump	
		Solder Diameter (um)	Height (um)
1	400	120	75
		100	62
		80	50
2	300	120	75
		100	62
		80	50
3	200	120	75
		100	62
		80	52

Results of this parametric study has been analyzed and explained how it leads to the optimization of this design variables in next chapter.

CHAPTER 5

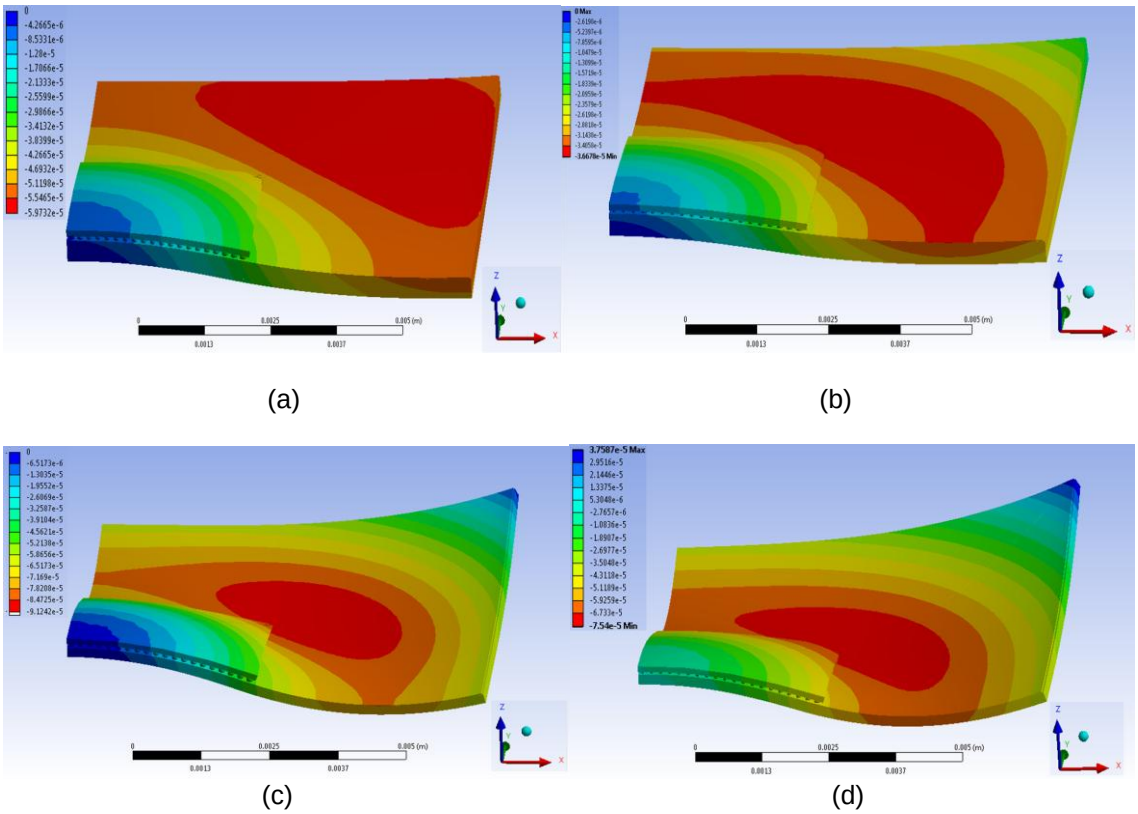
RESULTS AND DISCUSSION

5.1 Background and Methodology

Reflow process (die attach) has been simulated using ANSYS Workbench v14.5 to analyze the thermo-mechanical response of BEoL stack and *f*BEoL bumps. DOE is shown in section 4.6. The aim of this study was to focus on BEoL and *f*BEoL damage when varying the substrate thickness and solder bump foot print. To come up with adequate relationship, initially warpage at package level model has been compared. Further, sub model-1 with width and length of 175 µm (solder bump pitch) has been analyzed for stress/strain distribution at critical diagonally far corner bump. To focus on Cu/low-k region sub model-2 has been carried out with discretized Cu/low-k layers and analyzed for probable damage.

5.2 Package Warpage

Simulation at global level has been carried out on the octant symmetry of the package but for better understanding of the warpage profile, quarter symmetry is demonstrated in the results.



(a) (b)

(c) (d)

Figure 5.1 Warpage, where (a) Sub(t): 400, D-H: 120-75, (b) Sub(t): 400, D-H: 80-50, (c) Sub(t): 200, D-H: 120-75, (d) Sub(t): 200, D-H: 80-50 (all dimensions are µm)

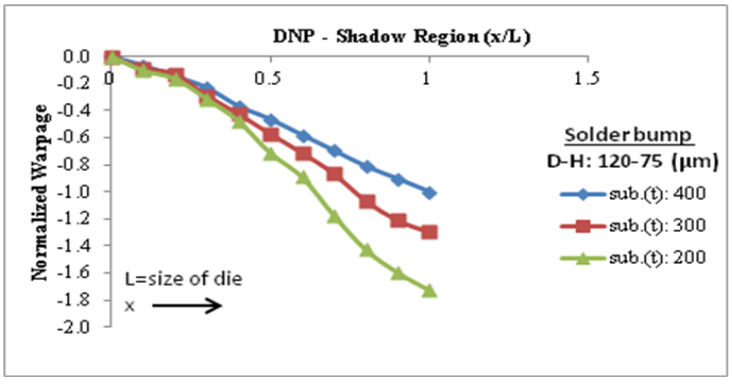


Figure 5.2 Normalized warpage (shadow region) with substrate thinning

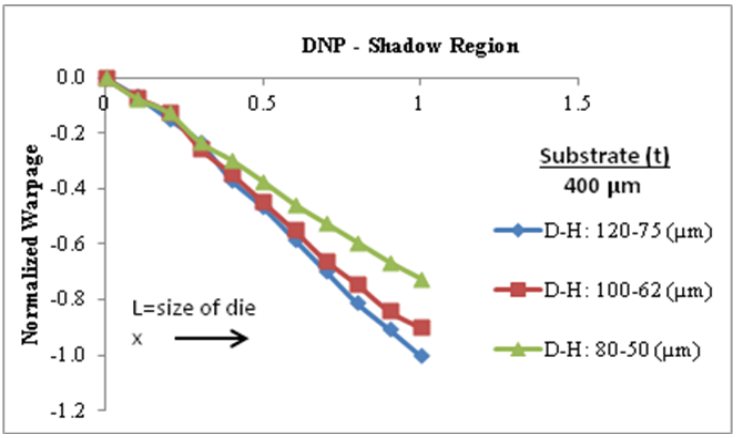
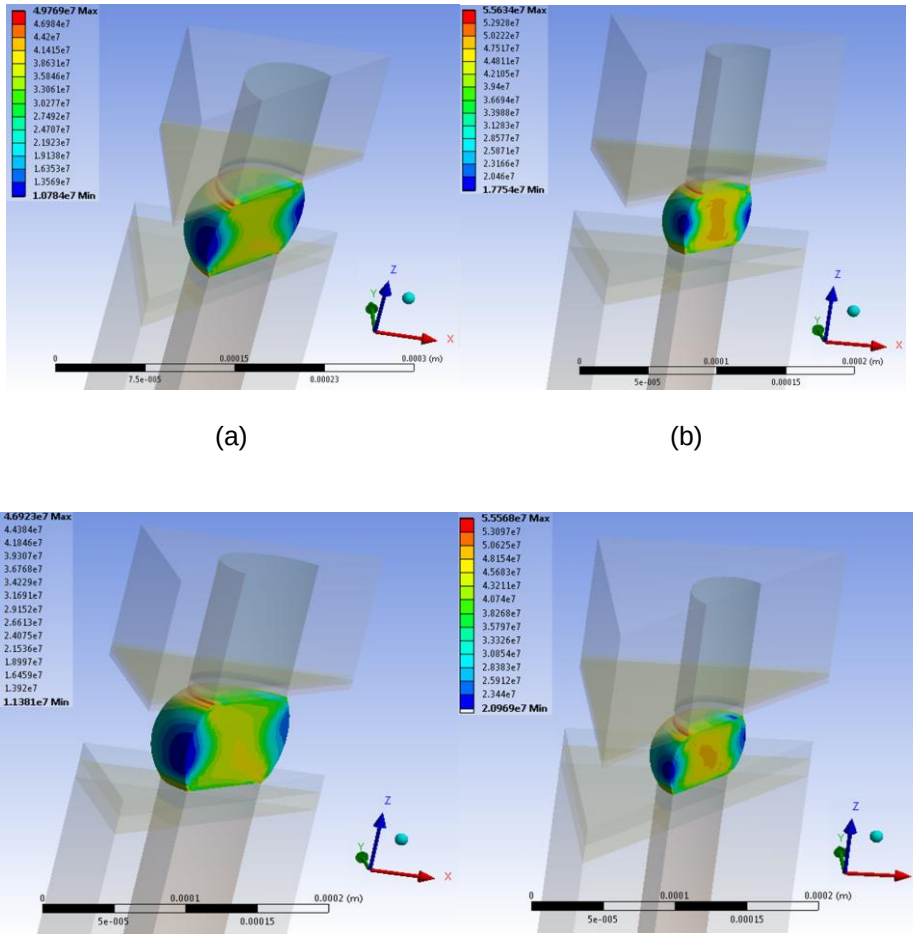


Figure 5.3 Normalized warpge (shadow region) with solder bump foor print reduction

Figure 5.1 shows the warpage profile of the quarter symmetry of the package while figure 5.2 and 5.3 shows the normalized warpage of the shadow region of the package. Sub(t) is considered as substrate thickness and D-H denotes diameter-height of the solder bump. It is evident from the Figure 5.2 that as the substrate thickness decreases, warpage increases. This relation holds true for different solder bump footprints. It is also apparent from figure 5.2, that the warpage behavior of the package is convex (crying) in shape in the die shadow region. From figures 5.1 (a) through (d), it is clear that warpage behavior of shadow region is governed by the CTE mismatch between the die and the substrate, but as the distance from neutral point (DNP) increases beyond the shadow region; substrate thickness and CTE mismatch between solder mask and substrate significantly contributes towards the out of plane bending. As the substrate thickness decreases, the convex shape of the warpage straightens (starts moving towards the smiling) and if the substrate thins down further, the package starts deforming in opposite (concave) direction towards the corners.

From figure 5.3, it is apparent that the warpage reduces as you wane the solder bump dimension. The bending moment in the package decreases with smaller bumps and results in less severe warpage. This relation holds true for each thickness of substrate.

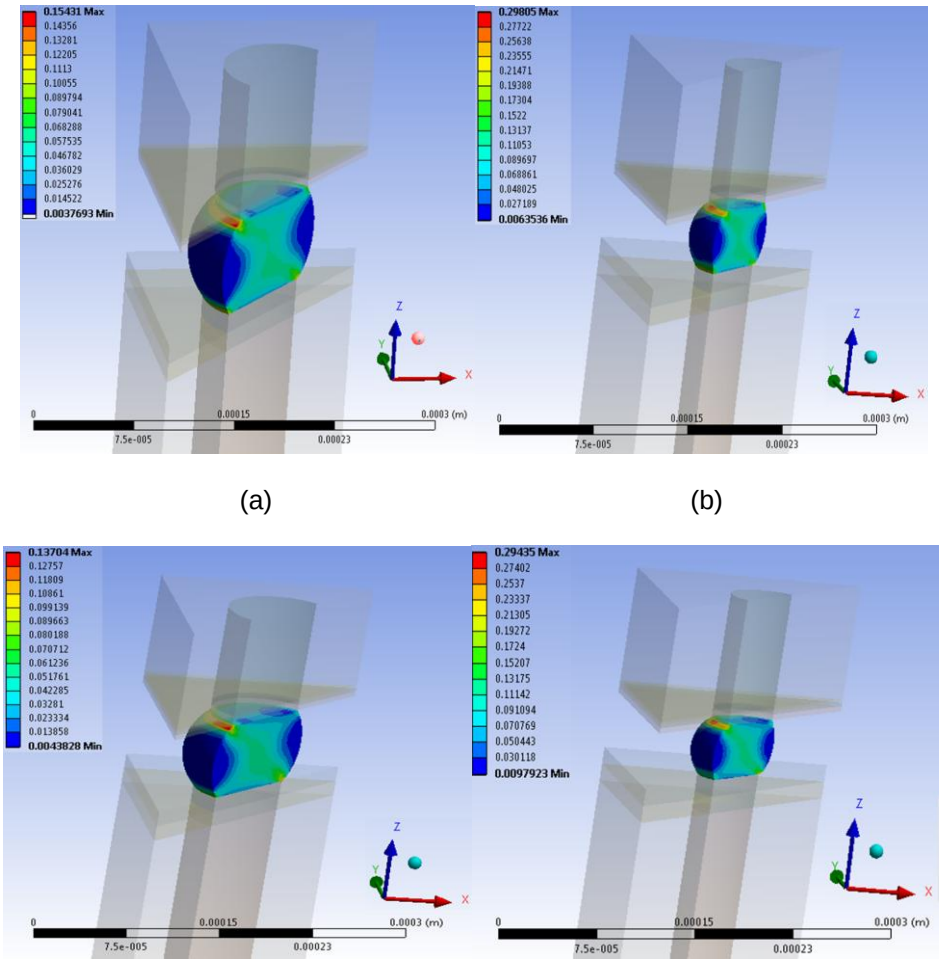
5.3 *f*BEoL Damage (Solder Bump)



(a) (b)

(c) (d)

Figure 5.4 Equivalent stress in solder bump: sub model-1, where (a) Sub(t): 400, D-H: 120-75, (b) Sub(t): 400, D-H: 80-50, (c) Sub(t): 200, D-H: 120-75, (d) Sub(t): 200, D-H: 80-50 (al dimensions are in µm)



(a) (b)

(c) (d)

Figure 5.5 Equivalent total strain in solder bump: sub model-1, where (a) Sub(t): 400, D-H: 120-75, (b) Sub(t): 400, D-H: 80-50, (c) Sub(t): 200, D-H: 120-75, (d) Sub(t): 200, D-H: 80-50 (all dimensions are in µm)

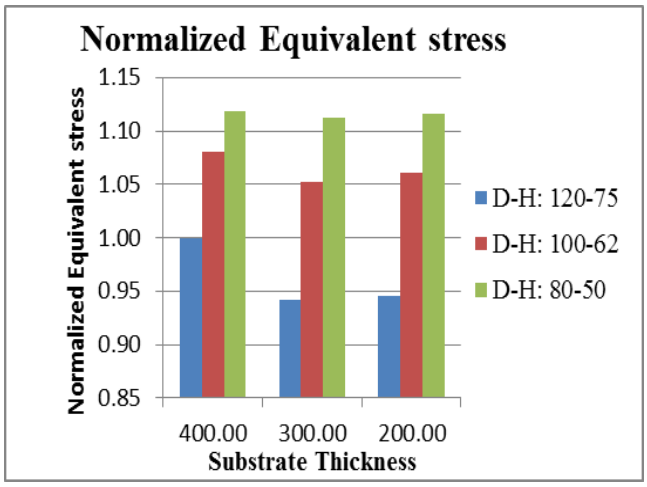


Figure 5.6 Normalized equivalent stress in solder bump: sub model-1

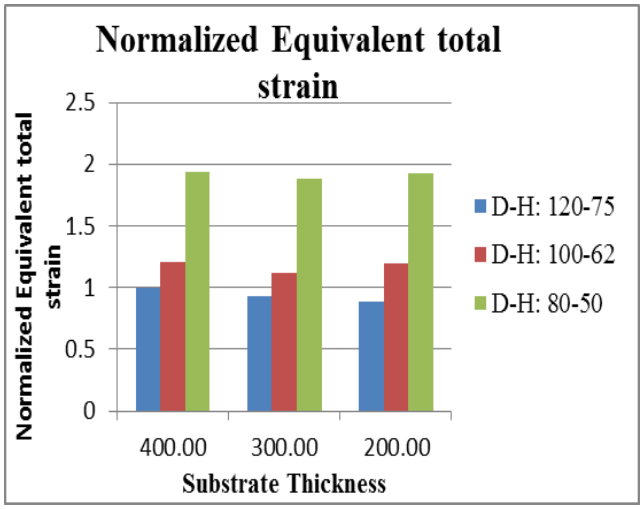


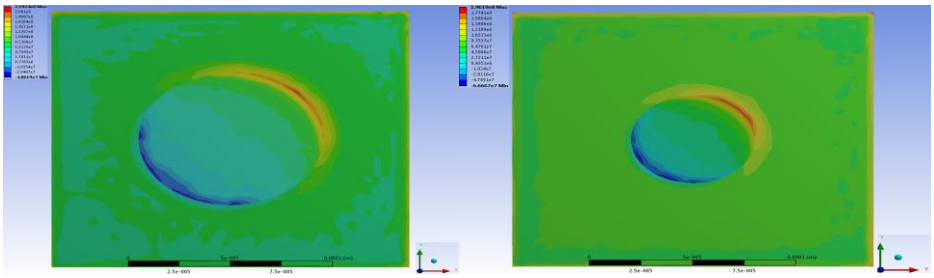
Figure 5.7 Normalized total equivalent strain in solder bump: sub model-1

Figure 5.4 and 5.5 shows the equivalent (von-Mises) stress and total equivalent strain in solder bump (*f*BEoL) in sub model-1.

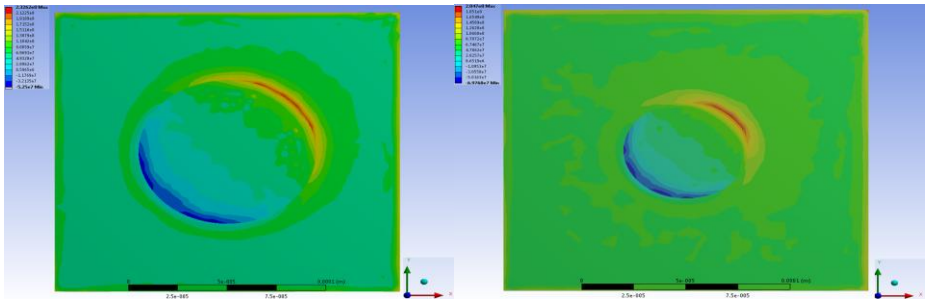
From figure 5.6 and 5.7, as expected, the equivalent stress and equivalent total strain in the solder bump increases as the solder footprint decreases (diameter and height decrease). The effective area is reduced for the same load; however, both these parameters (solder stress/strain) reduce within ~5% with substrate thinning. This result also explains that with

solder bump foot print reduction stress/strain increases so is the strain energy and plastic work. That means foot print reduction brings more damage to the solder bumps (*f*BEoL).

5.4 BEoL Damage (Low-k Layers)



(a) (b)



(c) (d)

Figure 5.8 Max. Principal stress in low-k layers: sub model-2, where (a) Sub(t): 400, D-H: 120-75, (b) Sub(t): 400, D-H: 80-50, (c) Sub(t): 200, D-H: 120-75, (d) Sub(t): 200, D-H: 8-50(all dimensions are in µm)

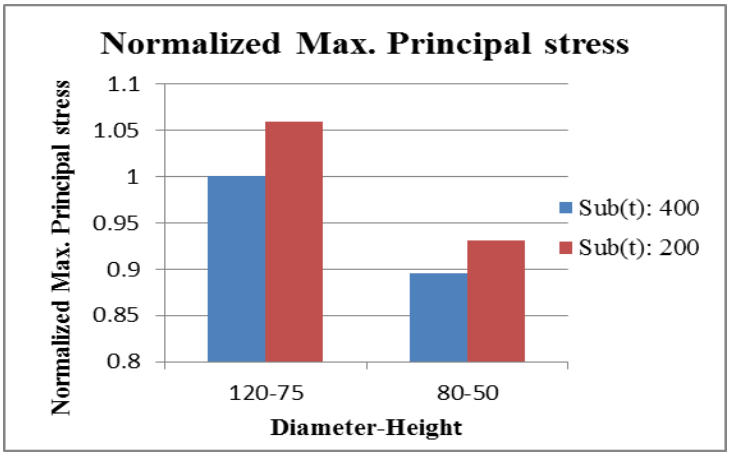


Figure 5.9 Normalized max. principal stress in low-k layers (BEoL)

From figures 5.8 it seems that, location of the maximum principal stress in low-k layers is around the centre region (UBM footprint on the dielectric) of the extreme corner (diagonal) unitcell. Figure 5.9 shows the variation in normalized maximum principal stress in the low-k layers for the extreme cases considered in this study. It suggests that as the substrate thickness decreases the maximum principal stress in low-k dielectric layers increases by about ~6%. Figure 13 also demonstrates that the maximum principal stress decreases with the reduction in solder diameter/height. As low-k layers are porous and brittle in nature, max. principal stress is considered to be a failure criteria. This max. principal stress in low-k represents peeling stress which can be reason for cohesive fracture or interfacial delamination in the BEoL region.

5.5 Summary

A parametric numerical study has been completed using multi level FE technique for the flip chip package under die attach reflow process to study the CPI. The thermo-mechanical behavior of the *f*BEoL (bump region) and the BEoL Cu/low-k regoin has been demonstrated through this effort. The input variables that were considered are solder bump (diameter and height) and substrate thickness.

Damage Location	Substrate Thinning	Solder Footprint Reduction
Warpage (Global)	↑	↓
Strain Energy (Solder Bump_SM1 - *f*BEoL)	↓	↑
Peeling Stress (Low-k layers_SM2-BEoL)	↑	↓

Figure 5.10 Summary

Figure 5.10 explains the whole summary of parametric study results through which certain things can be derived. It seems that substrate thinning significantly affects the package warpage so there should be check on it as it could lead to the issues in subsequent assembly process. Substrate thinning slightly affects the BEoL and *f*BEoL. Though this effect is lower but

they are qualitatively in opposite direction. Solder bump stress/strain (plastic work) decreases with increase in foot print at same time peeling stress in low-k layers and package warpage increases. Thus parametric study demonstrates a trade-off between the Cu/low-k (BEoL) and C4 bump (*f*BEoL) damage when the design parameters are varied. To minimize the damage to the BEoL and *f*BEoL components there is need to perform multi-objective design optimization to mitigate CPI risk during chip attachment.

CHAPTER 6

MULTI-OBJECTIVE DESIGN OPTIMIZATION

6.1 Background

Semiconductor industry has recognized the need to replace the traditional Al/SiO_2 interconnects with Cu/Low-*k* interconnects in the mainstream electronic devices following the latter's impact on both the power and RC delay reduction. However, due to lower elastic modulus and poor adhesion characteristic, reliability of the Cu/Low-*k* interconnects is a concern. Flip-chip packaging process (die attachment to substrate) can result in critical damage in Cu/Low-*k* interconnects. The objective of the study is to improve the reliability of Cu/Low-*k* interconnects during die attach reflow process by varying a set of design parameters including substrate thickness and solder bump footprint (diameter and stand-off height). Initial chip-package-interaction (CPI) parametric study demonstrated that there is a trade-off between the Cu/low-k (BEoL) and C4 bump (*f*BEoL) damage when the design parameters are varied. This trade-off provides a premise to perform a multi-objective design optimization (MODO). In this chapter, a simulation-based multi-objective design optimization has been exhibited to minimize the damage to the BEoL and *f*BEoL components, thereby mitigating CPI risk during chip attachment. The objective of this optimization analysis is to minimize two output variables - plastic work in solder bump (C4) and the peeling stress in low-*k* layers. ANSYS code with a built-in optimization tool has been leveraged to carry out the multi-level (global-local approach), multi-objective optimization study to determine optimal design parameters that enhance the package reliability. The proposed simulation-based optimization approach will prove to be critical in the product development phase and would lead to a robust and a reliable design

6.2 Introduction to Optimization

As it is known that a good design point is often the trade-off between various objectives, exploration cannot be done by using an optimization algorithm which leads to one design point.

It is important to gather information within the applicable design space to answer all kind of "what if" question. To gather all this information and efficiently reach the final design point, ANSYS Workbench has "Design Exploration" optimization tool. Design exploration describes the relationship between the design variables and the performance of the product by using Design of Experiments (DOE), combined with response surfaces. DOE and response surfaces provide all of the information required to achieve Simulation Driven Product Development [43].

6.2.1 Design of Experiments (DOE)

Design of Experiment is the technique to scientifically determine location of sampling points within the available design space to cover as much as possible space. There is wide range of algorithm available in engineering literature which tries to locate the sampling points such that design space for input parameters is explored in most efficient way. In ANSYS within the design exploration tool there are 7 different algorithms available which helps in generating DOE within specified design space (specified by upper and lower bound) for each input parameters, which are mentioned below [43].

1. Central Composite Design (CCD)
2. Optimal Space-filling Design (OSF)
3. Box-Behnken Design
4. Custom
5. Custom + Sampling
6. Sparse Grid Initialization
7. Latin Hypercube Sampling Design (LHS) [43]

In this study Custom + Sampling DOE was considered. Sampling design points were leveraged from Optimal Space-filling Design algorithm and extreme combination of design space were added as a custom points.

6.2.2 Response Surface (RS)

The Response Surfaces are functions where the output parameters are described in terms of the input parameters. This input and output values are taken from solved DOE. They are built from the Design of Experiments in order to provide quickly the approximated values of the output parameters, everywhere in the analyzed design space, without to perform a complete solution. The accuracy of a response surface depends on factors such as complexity of the variations in the solution, number of design points and type of response surface selection. To create response surface through DOE there are certain meta-models available in design exploration tool which are mentioned below [43].

1. Standard Response Surface – Full 2nd Order Polynomial
2. Kriging
3. Non-Parametric Regression
4. Neural Network
5. Sparse Grid [43]

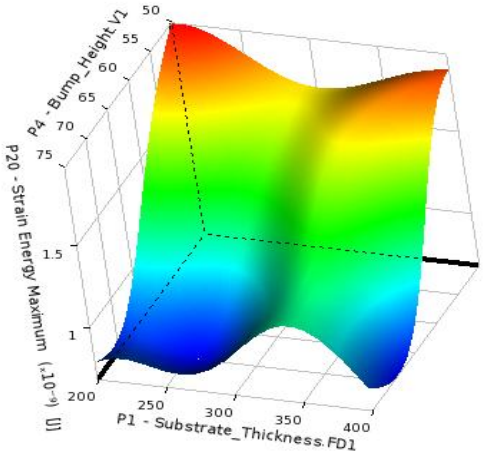


Figure 6.1 Strain Energy Response Surface

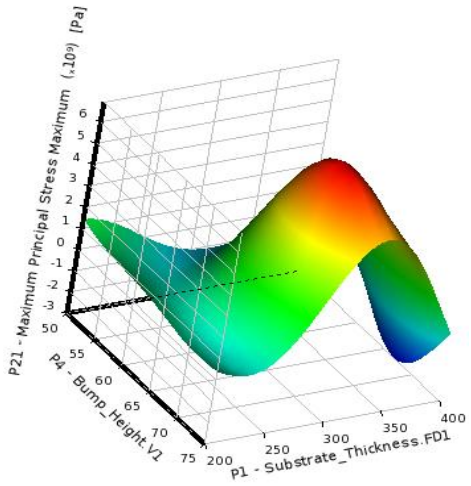


Figure 6.2 Peeling Stress Response Surface

In this study Non-Parametric Regression meta-model has been used. This algorithm covers the predictably high nonlinear behavior of output with respects to its input. Response surface for both output parameters considered in this optimization study is shown in figure 6.1, 6.2.

6.2.3 Goodness of Fit (GOF)

Goodness of Fit is an evaluation parameter for response surface which helps you determine how well your response surface is fitting all the design points in design space. To determine the accuracy of current response surface GOF is used. To check the acceptability of GOF various parameters are there in GOF matrix which will be discussed here. If GOF is not acceptable that means current response surface is not accurately representing a parametric model. In That case there is need to refine your response surface [43].

Criteria for GOF are mentioned in matrix form in response surface project schematic for each individual output parameters. These criteria are as mentioned below.

1. Coefficient of Determination: This is the ratio of explained variation to the total variation. The best value it is 1.
2. Maximum Relative Residual: The maximum distance (relatively speaking) out of all of the generated points from the calculated response surface to each generated

point. The best value is **0%**; in general, the closer the value is to **0%**, the better quality of the response surface.

Name	P20 - Strain Energy Maximum	P21 - Maximum Principal Stress Maximum
Goodness Of Fit		
Coefficient of Determination (Best Value = 1)	1	1
Maximum Relative Residual (Best Value = 0%)	0	0
Root Mean Square Error (Best Value = 0)	9.7005E-19	3.2706
Relative Root Mean Square Error (Best Value = 0%)	0	0
Relative Maximum Absolute Error (Best Value = 0%)	0	0
Relative Average Absolute Error (Best Value = 0%)	0	0

Figure 6.3 GOF Matrix

3. Root Mean Square Value: This is the square root of the average square of the residuals at the DOE points for regression methods. The best value is **0**; in general, the closer the value is to **0**, the better quality of the response surface.
4. Relative Root Mean Square Error: This is the square root of the average square of the residuals scaled by the actual output values at the points for regression methods. The best value is **0%**; in general, the closer the value is to **0%**, the better quality of the response surface.
5. Relative Average Absolute Error: This is the average of the residuals relative to the standard deviation of the actual outputs. This is useful when the number of samples is low (<**30**). The best value is **0%**; in general, the closer the value is to **0%**, the better quality of the response surface [43].

Figure 6.3 shows the GOF matrix values obtained in this optimization study.

6.2.4 Predicted versus Observed Chart

This chart represents for output parameters the value predicted from response surface versus the value observed from design points. This chart lets you quickly decide how well your response surface fits all design points as well refinement points. Closer the points from the diagonal identity line, better the response surface fit all points. All the output values are by default normalized [43].

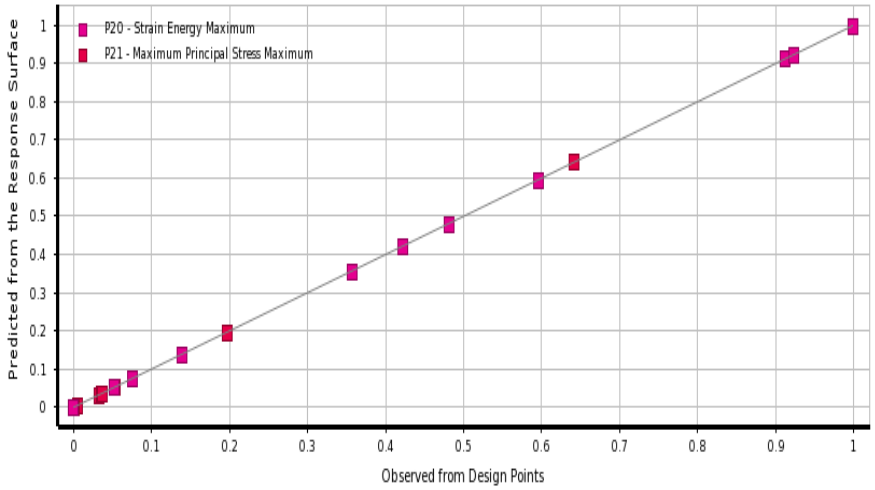


Figure 6.4 Predicted versus Observed Chart

6.2.5 Verification Point & Refinement Point

Initial response surface is built from design points available in DOE. As this response surface is built from limited number of design points, it needs a check for accuracy. RS algorithms are mainly an interpolation process which fits all the design points (such as kriging). In this case GOF parameters would indicate that response surface is accurate enough for given design points but does not indicate that current RS represents whole parametric solution. A better way to check this accuracy is creating a verification points. Verification points are solved on real model and their output is compared with predicted output from the current RS to check the accuracy of RS. To check the fitness of this verification points with current RS, there is a different GOF matrix for verification points with the same criteria. To quickly determine that,

predicted versus observed chart can be used. It is easy to differentiate original design points (square) with verification point (round) on this chart [43].

Through GOF matrix of verification points and predicted versus observed chart accuracy of current RS can be obtained. If this it's not acceptable, that means current RS need a refinement to improve accuracy and covering parametric solution. So idea is to add those verification points as a refinement points to refine current RS. In RS project window there is option of inserting these verification points as refinement points which lets you refine and update response surface [43].

6.2.6 Optimization

In Design Exploration tool there are two options to perform Goal Driven Optimization (GDO): Response surface optimization and direct optimization. Direct optimization is single component system which uses a real solve while response surface optimization depends on its own response surface to provide optimized candidate, so accuracy of optimized candidate depends on accuracy of RS [43]. In this study response surface optimization technique was used to perform optimization. There are several algorithms available for performing this optimization such as screening, Multi-Objective Genetic Algorithm (MOGA), Nonlinear Programming by Quadratic Lagrangian (NLPQL), and Mixed-Integer Sequential Quadratic Programming (MISQP) [43].

In this study screening algorithm has been considered to perform optimization. It allows us to create new numbers of samples and then sort them by using objective function or constrains.

In this algorithm you can enter as much number of samples you want as it does not take much time. This method is RS based, so no need of real solve. After that you get a chance to specify your objective for optimization and constrains, if any. In case of multiple objectives, you can provide weight (preference) to each one by defining higher or lower relevance. You also get an option of how many optimized candidate you want [43]. There is also an option of

verifying optimized candidate with verification point with real solve. If the difference between both points is not acceptable then again refinement can be done [43].

6.2.7 Flow Chart

Optimization process in ANSYS contains of different individual steps like creating DOE, RS, perform GOF check, verifying accuracy of current RS and then perform optimization as well verify the optimized candidate which follows the below mentioned flow chart.

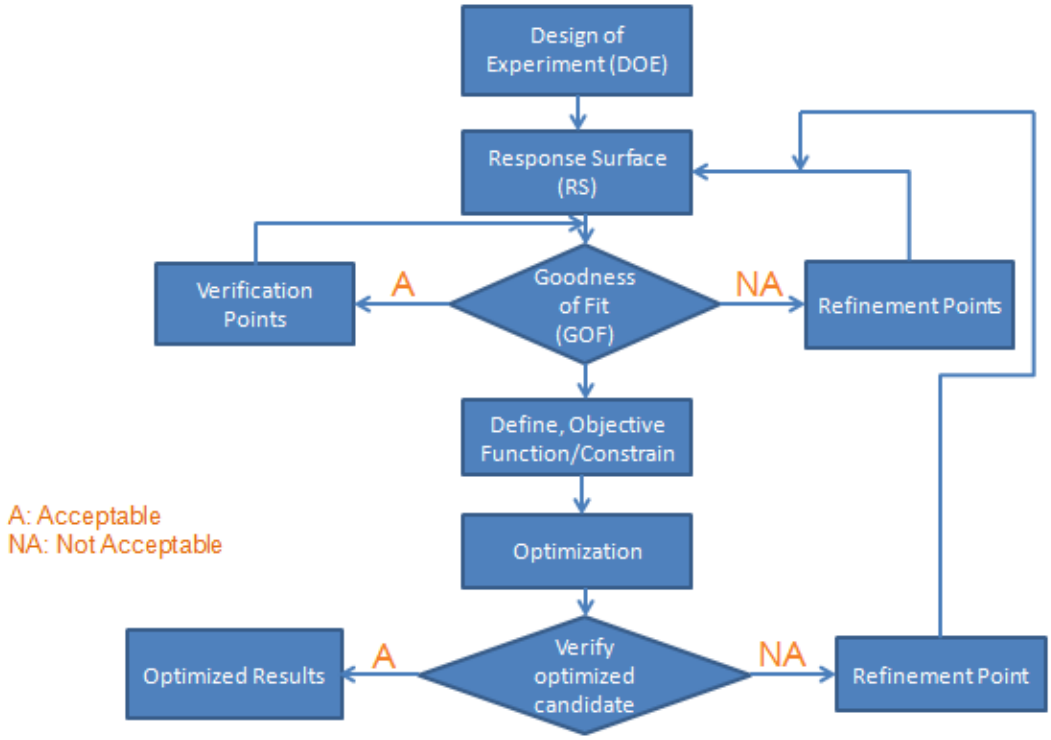


Figure 6.5 Optimization Process Flow Chart

6.3 Results and Discussion

Multi-Objective Design Optimization has been carried out to reduce the damage in BEoL and *f*BEoL region on flip chip package during the chip attach (solder bump reflow) process. This optimization approach is based on DOE and Response surface and whole optimization process has been followed as stated in figure 6.5. Results obtained by this multi-objective design optimization study are as shown in table below.

Table 6.1 Optimization Results

Case	Substrate Thickness (μm)	Solder bump height (μm)	Strain Energy (J)	Max. Peeling stress (MPa)
Nominal	300	62.5	9.81e-10	219
Optimal	242.3	69.8	7.12e-10	192

Comparing the output parameters values between nominal candidate and optimal candidate from the table above, it seems that there is an improvement of 28% in strain energy of solder bump which is the part of *f*BEoL and improvement of 10% in peeling stress in low-k layers which is a part of BEoL.

CHAPTER 7

CONCLUSION

7.1 Conclusion

A Multi-level Finite Element submodeling technique has been implemented to predict the Chip Package Interaction (CPI) during chip attachment process and also to bring down the damage in Back End of Line (BEoL) and Far Back End of Line (*f*BEoL) region. This whole study has been carried out in two phase. In first phase, parametric study has been done by considering the two critical design variable (substrate thickness & solder bump foot print). Results of this parametric study exhibited that damage in BEoL and *f*BEoL region are qualitatively opposite direction with the variation in design parameters. So there is a trade off in BEoL and *f*BEoL damage. Even the response is unknown when these design parameters are varied together. These evidences provide the premise to perform multi-objective design optimization keeping objective in focus to reduce the damage in both BEoL and *f*BEoL, which is the other phase of the study. This multi-objective design optimization demonstrates significant improvement in the thermo-mechanical response at BEoL and *f*BEoL region. This work is of immense importance from process integration standpoint. It can provide a quantitative upstream guideline to the process/electrical team on the BEoL/fBEoL damage. This proposed process can help in improving the reliability and eliminating critical reliability issues from the product development phase itself.

7.2 Future Work

- Similar parametric and optimization study can be done by taking time-temperature dependent material properties for low-k materials to predict the exact chip package interaction during thermo-mechanical loading.

- Minimize fracture in BEoL by introducing crack stopper and its location and even by varying the chip size and thickness.

REFERENCES

1. Jonathan Luke Drake “THERMO-MECHANICAL RELIABILITY MODELS FOR LIFE PREDICTION OF BALL GRID ARRAYS ON CU-CORE PCBs IN EXTREME ENVIRONMENTS” Auburn University, 2007.
2. Hocheol Kwak, Dr. Todd Hubing “ An overview of advanced electronic packaging technology” Technical Report: CVEL 007-001.
3. Back End of Line: www.wikipedia.com/backend of line.
4. www.analysistech.com.
5. Lopez-Buedo, S., Boemo, E. “Electronics Packaging Technologies”.
6. Zhang, X. “Chip package interaction (CPI) and its impact on reliability of flip-chip packages” The University of Texas, December 2009.
7. Moore’s law: http://www.cymer.com/moores_law/.
8. Front End of Line: http://en.wikipedia.org/wiki/FEOL.
9. IC Fabrication: http://en.wikipedia.org/wiki/Integrated_circuit#Fabrication.
10. Back End of Line: http://en.wikipedia.org/wiki/Back_end_of_line.
11. http://www.cemag.us/articles/2001/12/cu-integrated-low-k-dielectrics-future-now#.UfBBso3VA0E.
12. T. Raman, “ASSESSMENT OF THE MECHANICAL INTEGRITY OF CU/LOW-K DIELECTRIC IN A FLIP CHIP PACKAGE” University of Texas at Arlington, December 2012.
13. http://www.hal.ipc.i.u-tokyo.ac.jp/research/titac/perspect/perspect-j.html.
14. X.H. Liu, T.M. Shaw, G. Bonilla. “Mechanical Reliability Outlook of Ultra low-K Dielectrics” Advanced Metallization Conference, New York, 2010.
15. M. Bundalo, A. Baig, “Fundamentals of Design for Reliability”.

16. R.J.O.M. Hoofman, “Reliability challenges accompanied with interconnect downscaling and ultra low-k dielectrics” IEEE, 2005.
17. M. Brillouet, “Shifting Challenges in The Integrated Interconnection Systems” France, IEEE, 2005.
18. L. Fu, M. Su, F. Kuechenmeister, W. Huang, “Chip Package Interaction (CPI) Reliability of Cu/Low-k/Ultra-Low-k Interconnect” IEEE, 16th IPFA, China, 2009.
19. C.J. Uchibori, “Impact of Chip-Package Interaction on Reliability of Cu/Ultra Low-k Interconnects for 65nm Technology and Beyond” IEEE, 2006.
20. G. Wang, P.S. Ho, S. Groothuis, “Packaging Effect on Relibility of Cu/Low-k Interconnects” IEEE, 2004.
21. C.J. Zhai, U. Ozkan, A. Dubey, Sidharth, R.C. Blish II, R.N.Master, “Investigation of Cu/Low-k Film Delamination in Flip Chip Packages” Electronics Components and Technology Conference, 2006.
22. C.T. Kuo, M.C. Yip, K.N. Chiang, C. Tsou, “Characterization Study of Time- and Temperature- Dependent Mechanical Behavior of Polyimide Materials in Electronics Packaging Applications” Journal of ELECTRONICS MATERIALS, 2005.
23. T.H. Wang, Y.S. Lai, “Submodeling Analysis for Path-Dependent Thermomechanical Problems” Journal of ELECTRONICS PACKAGING, 2005.
24. F.X. Che, X. Zhang, W.H. Zhu, T.C. Chai, “Reliability Evaluation for Cu/Low-*k* Structures Based on Experimental and Numerical Methhods” IEEE, 2008.
25. F.X. Che, H.L.J. Pang, W.H. Zhu, W. Sun, A.Y.S. Sun, C.K. Wang, H.B. Tan, “Development and Assessment of Global-Local Modeling Technique Used in Advanced Microelectronics Packaging” United Test & Assembly Centre Ltd., Nanyang Technological University.

26. Y.S. Lai, T.H. Wang, “Verification of submodeling technique in thermomechanical reliability assessment of flip-chip packaging assembly” Microelectronics Reliability, 2005.
27. M. Lofrano, C.J. Wilson, K. Croes, B. Vandevelde, “Thermo-Mechanical Modeling of Stress-Induced-Voiding in BEOL Cu Interconnect Structures” EuroSimE, 2009.
28. J. Auersperg, D. Vogel, M.U. Lehr, M. Grillberger, B. Michell, “Crack and Damage Evaluation in Low-k BEoL Stacks Under Chip Package Interaction Aspects” EuroSimE, 2009.
29. C.S. Lau, M.Z. Abdullah, C.Y. Khor, “Optimization of reflow soldering process with multiple quality characteristics in ball grid array packaging by using the grey- based Taguchi method” Microelectronics International, 2012.
30. L. Xu, T. Reinikainen, W. Ren, B.P. Wang, Z. Han, D. Agonafer, “A Simulation –based multi-objective design optimization of electronics packages under thermal cycling and bending” Microelectronics Reliability, 2004.
31. M.M. Hossain, S.G. Jagarkal, D. Agonafer, M. Lulu, S. Reh, “Design Optimization and Reliability of PWB Level Electronic Package” Journal of Electronic Packaging, 2007.
32. J.N. Reddy, “Introduction to Finite Element Method”.
33. “The Finite Element Method and Applications in Engineering” http://www.springer.com/978-0-387-28289-3.
34. “Introduction to ANSYS Meshing” Lecture 3, 14.5 Release.
35. Chapter- 07 “FEM Modeling: Mesh, Loads and BCs”.
36. Sub modeling: http://mostreal.sk/html/guide_55/g-adv/GADV4.htm
37. Flipchips.com: http://flipchips.com/tutorial/process/under-bump-metallization-ubm/
38. Flipchips.com: http://flipchips.com/tutorial/assembly/redistribution-layers/

39. S.G. Chow, Y. Lin, E. Ouyang, B. Ahn, “A Finite Element Analysis of Board Level Temperature Cycling Reliability of Embedded Wafer Level BGA(eWLB) Package” Electronics Components and Technology Conference, 2012.
40. O.M. Guan Jimmy, “Structural Design and Optimization of 65nm Cu/low-k FlipChip Package” National University of Singapore, 2008.
41. G. Wang, P.S. Ho, “Packaging Effect on Reliability for Cu/Low-k Damascene Structures” TRC 2003, The University of Texas at Austin.
42. S.B. Park, I.Z. Ahmed, “Shorter Field Life in Power Cycling for Organics Packages” Transaction of ASME, 2007.
43. ANSYS user guide: “Workbench_Design Exploration” Release 14.5.

BIOGRAPHICAL INFORMATION

Hardik Parekh completed his Bachelors of Engineering in Mechatronics Engineering from Sardar Patel University, Gujarat, India. He worked at Ferromatik Milacron India Pvt Ltd. as a design engineer for a year. He completed his Masters of Science degree in Mechanical Engineering in December 2013. He was part of the Electronics, MEMS, Nanoelectronics Systems Packaging (EMNSPC) research team from January 2013. His research primarily deals with device level packaging of a flip chip with BEoL/*f*BEoL interconnects. He has authored and coauthored four conference papers related to electronic packaging.