Package-Chip Co-Design to Increase Flip-Chip C4 Reliability

Sheldon Logan, Matthew R. Guthaus
Department of CE, University of California Santa Cruz, Santa Cruz, CA 95064
{slogan,mrg}@soe.ucsc.edu

Abstract— **The magnitude of the I/O requirements for modern ICs continues to increase due to the growing complexity and size of ICs. The large I/O count found on most ICs have forced most designers to use flip-chip packaging instead of wire bonded packaging. Unfortunately, the solder bumps in flip-chip packages are susceptible to failure, especially in the presence of high temperatures which can cause large stresses and strains leading to mechanical failure of the bump. In this paper, we present a simplified stress/strain/fatigue model that can be used during floorplanning to optimize for package reliability. We also demonstrate a quadratic C4 bump placement method that can be used during floorplanning to increase C4 bump reliability. Our experimental results show that this co-optimization can increase the lifetime of C4 bumps by about $47\times$ with only a modest 3% increase in HPWL wirelength.**

I. Introduction

There are two main methods of connecting IC dies to package substrates, wire bonding and flip chip processing. Wire bonding consists of routing bond wires from the I/O pads located on the perimeter of the die to the package. The flip chip method, also called Controlled Collapse Chip Connection (C4), connects the die to the substrate using C4 solder balls located throughout the die area. A sideways view of a flip chip package is shown in Figure 1 highlighting the C4 solder ball connections between the chip and package.

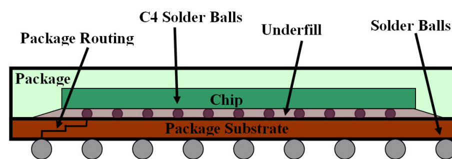


Fig. 1. Flip Chip Package

There are several advantages of flip chip designs over wire-bonded designs. Since the I/O connections can be located throughout the die, the length of wires connecting the die to the package can be greatly reduced leading to better performance. Also, since the I/O connections are not limited to the perimeter of the die, more I/O connections are feasible than in wire bonded designs. Lastly, wire bonded designs usually require larger packages due to the I/O pads and bonding wires. These advantages along with the increase in complexity and size of ICs have resulted in many modern designs using flip chip connections [1].

One of the major disadvantages of flip chip designs over wire bonded designs is the reliability of the C4 solder balls. Due to the Coefficient of Thermal Expansion (CTE) mismatch between the different layers in the package (substrate, underfill/C4 balls, and silicon die), the solder balls experience large stresses which, over many thermal cycles, can lead to crack formation and subsequent failure. A recent $150-200MM recall by NVIDIA [2] of certain GPUs due to solder ball reliability illustrates the severity of the problem.

Due to recent legislature banning the usage of hazardous lead in electronic components, lead-free solders have started being used in flip chip designs. Unfortunately, these lead-free solders are more susceptible to failure from thermal cycling than lead based solders, which means that C4 bump reliability will become increasingly important in future IC designs. One method of addressing the solder bump reliability problem during IC design is to co-optimize the placement of bumps and the chip which will be the focus of this paper.

The importance of chip package co-design are detailed in the work of McGrath et al [3]. However, most traditional floorplanning papers found in literature only consider wire bonded designs. There have been a few recent works that consider the flip chip packaging problem with some of these works focusing on C4 bump placement [4]. Others have focused on the co-optimization of the placement of I/O buffers, C4 bumps and blocks so as to minimize design metrics such as total wirelength and skew [5]–[7]. Others have considered flip chip routing and how it affects PCB escape routing [8]. In an alternate direction, there have been many works on modeling the reliability of solder balls and flip chip packages [9]–[12], but none of these works consider the co-optimization of C4 bump placement and block placement to increase package reliability. To the best of our knowledge, this is the first work to consider bump placement/floorplanning in flip chip designs to increase the lifetime/reliability of C4 bumps.

Our contributions in this paper are three-fold.

- First, we outline a simplified stress/strain/fatigue model for C4 solder balls.
- Second, we develop a quadratic C4 ball placement algorithm that optimizes for both wirelength and reliability which can be used during floorplanning.
- Third, we show that by doing package chip co-optimization, solder ball reliability can be increased without significant loss in the previous performance metrics.

Our work proceeds as follows: Section II introduces the theoretical and experimental motivation for our work. Section III discusses how stress and strain affect C4 reliability. Section IV

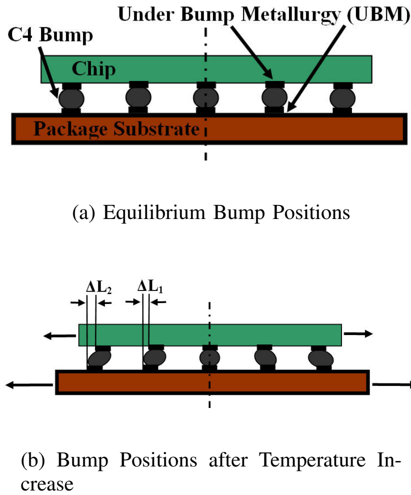
introduces our quadratic C4 ball placement algorithm for improved reliability. Section V provides the experimental setup and then the results obtained from our floorplanner are presented in Section VI. Finally, Section VII concludes the results.

II. Thermal-Mechanical Stress

In this section we provide a brief explanation of the stress/strain models used in our experiments.

A. Shear Strain in C4 Solder Bumps

Strain in C4 bumps are caused by the CTE mismatch between the silicon chip and the substrate to which the chip is attached. The CTE mismatch causes the chip and the substrate to expand by different amounts when subjected to changes in temperature which results in shear strains in the solder balls as shown in Figures 2(a)and 2(b). Figure 2(a) shows the substrate and chip in the equilibrium state and Figure 2(b) shows the substrate and chip after a increase in temperature.



(a) Equilibrium Bump Positions

(b) Bump Positions after Temperature Increase

Fig. 2. Shear Strain in C4 Bumps

B. Strain and Stress Calculation

1) Strain: The shear strain caused by the CTE mismatch can be approximated using the following formula [13]:

$$\gamma = r\,(\alpha_s - \alpha_c)\,\Delta T \tag{1}$$

where r, is the scaled distance from the centroid of chip/substrate, α_s is the CTE for the substrate, α_c is the CTE for the chip, ΔT is the difference in temperature from the equilibrium temperature value and γ is the shear strain. It should be noted that the maximum shear strain occurs at the edge of the chip and no shear strain occurs at the center as illustrated in Figure 2(b). Equation 1 assumes that there is no underfill between the substrate and the chip. Underfill provides mechanical reinforcement between the die and substrate, consequently reducing the strains and stresses experienced by C4 balls. Modelling the underfill effects however is a complex task and usually requires Finite Element Analysis(FEA). For our calculations we chose to ignore the underfill effects, hence all the strains calculated can be viewed as an upper limit to the actual value. For our experiments, we used 2.3 and 25 for the values of α_c and α_s respectively.

2) Stress: If we assume that the C4 solder material obeys Von Mises' criterion, the equivalent strain can be approximated from the following formula

$$\begin{aligned}\epsilon_e \;=\; & \frac{\sqrt{2}}{3}[(\epsilon_{xx}-\epsilon_{yy})^2+(\epsilon_{yy}-\epsilon_{zz})^2+(\epsilon_{zz}-\epsilon_{xx})^2 \\ & +\frac{3}{2}(\gamma_{xy}^2+\gamma_{yz}^2+\gamma_{xz}^2)]^{\frac{1}{2}}\end{aligned} \tag{2}$$

where ϵ and γ are the various normal and shear strains in the solder material. Since the dominant shear strain in solder balls is in the xy plane, we set all the normal and shear strains to 0 except γ_{xy}. The equivalent stress can be estimated from the equivalent strain using the stress-strain relationship

$$\sigma_e = E\epsilon_e \tag{3}$$

where σ is the equivalent stress and E is the Young's modulus of the material. The Young's modulus of solder is, however, temperature dependent. For our experiments, we approximate the Young's Modulus for the SnAgCu solder as

$$E = 52708 - 67.14T - 0.0587T^2 \tag{4}$$

where the temperature is measured in Celsius. This relation was obtained from experiments presented in [9].

III. C4 Ball Failure

As stated in the introduction, SnAgCu C4 solder balls are susceptible to low cycle thermal fatigue failure as a result of their viscoplacsiticity. Failure in C4 balls is usually due to fracture which is caused by crack formation and propagation caused by cyclic thermal induced stresses. Figure 3 obtained from [13] illustrates the magnitude of the cracks that can be created in solder balls from thermal cyclic loading.

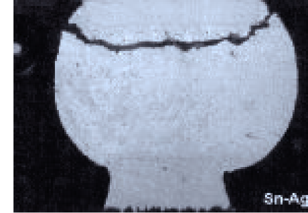


Fig. 3. Fracture of SnAg Solder Ball [13]

There are several methods of estimating the mean cycles to failure for C4 solder balls. For our experiments we used the Knecht-Fox model due to its simplicity. The model is described in the following equation:

$$N_f = \frac{C}{\Delta\epsilon_c} \tag{5}$$

where N_f is the number of cycles to failure, C is an empirical constant and $\Delta\epsilon_c$ is the creep strain range. For our experiments C was set to 8.9 as reported in [12]. Details on how to calculate the creep strain range for a C4 solder bump can be found in Section III-A below.

A. Creep Analysis

Creep is defined as the slow deformation of a material subject to high stresses. It is positively correlated to temperature in that large temperatures increase the rate of the creep deformation. The creep strain rate in SnAgCu C4 solder balls can be calculated by using the Garofalo-Arrhenius hyperbolic sine law:

$$\frac{d\gamma}{dt} = C\left(\frac{G}{\Theta}\right)\left[\sinh\left(\omega\frac{\tau}{G}\right)\right]^n \exp\left(-\frac{Q}{R\Theta}\right) \quad (6)$$

where $\frac{d\gamma}{dt}$ is the creep shear strain rate, γ is the creep shear strain, C is a material constant, Θ is the absolute temperature, G is the shear Modulus, τ is the shear strain, Q is the activation energy, R is Boltzmann's constant, n is the stress exponent and ω is the stress level. Since we assume the C4 bumps obey Von Mises' criterion, Equation 6 can be rearranged to:

$$\frac{d\epsilon}{dt} = C_1\left[\sinh\left(C_2\sigma\right)\right]^{C_3} \exp\left(-\frac{C_4}{T}\right) \quad (7)$$

where $C1$, $C2$, $C3$, and $C4$ are material constants, σ is the equivalent stress, T is temperature, $\frac{d\epsilon}{dt}$ is the equivalent creep strain rate. The equivalent stress is calculated using Equation 3. In our experiments the bump temperatures were obtained through Hotspot [14] simulations. A detailed explanation of how these temperatures are estimated is given in Section III-B. For our experiments, we set $C1$ as 501.3, $C2$ as 0.031, $C3$ as 4.96 and $C4$ as 5433.5 which are the material constants reported in [15]. Given a creep strain rate, the creep strain range can be calculated using the following equation:

$$\Delta\epsilon_c = t\dot{\epsilon_c} \quad (8)$$

where $\dot{\epsilon_c}$ is the creep strain rate and t is the length of time a bump is subjected to a high stress (i.e. the time the chip is in an active state causing high temperatures and consequently high stresses). In our experiments we assumed the chip would be in its peak active mode for 1s intervals.

B. Temperature of Bumps

The temperature of the bumps are determined by two factors: 1)The self-heating generated in the bumps as current flows through them and 2)The temperature of the blocks within the vicinity of the bumps. We used multi-layer simulations in Hotspot to generate accurate pin temperatures. For our experiments, we used three layers. Layer 1 is the pin layer, layer 2 is the silicon layer and the final layer is the thermal interface material (TIM) layer. The resistivity and the current through the bump was used to determine the self heating power of an individual bump according the following equation:

$$H_{bump} = i^2\left(\rho\frac{h}{\pi r^2}\right) \quad (9)$$

where H_{bump} is the heating power of the bump, i is the current flowing through the bump, ρ is the resistivity of the bump, h is the height of the bump and r is the radius of bump. The current in the bumps were assumed to be between 0.1A and 1A which are the values used in the experiments in [16]. The diameter of the solder bumps were set to 50μm, the height of solder balls were set to 60μm, and the minimum pitch was set to 100μm as reported in [17]. The thermal conductivity and electrical resistivity of the balls were assumed to be that of Tin(Sn) and set to 67 W/mK and $1.09 \times 10^{-7}\Omega \cdot m$, respectively.

IV. Floorplanning for Reliability

In this section we outline methods to increase the lifetime of C4 bumps during floorplanning using the previously mentioned fatigue model and our quadratic C4 ball placement algorithm.

Analysing Equation 7, it is clear that the reliability/lifetime of a C4 bump is affected by 1) temperature, 2) stress and 3) the material properties of the solder balls, package and chip. Factors 1 and 2 can be controlled during floorplanning/placement since the placement of blocks affects the temperature map of the chip and the placement of bumps will affect the bump distance from the center of the chip and subsequently its stress.

A. Quadratic Ball Placement

Our proposed quadratic ball placement algorithm is shown in Algorithm 1.

Algorithm 1 Quadratic Stress-Aware Ball Placement

Require:
Stress-Aware I/O Ball Placement.

Ensure:
All balls satisfy a minimum number of cycles to failure.

1: Calculate optimal position for balls via quadratic optimization.
2: Create a grid of possible ball locations.
3: Calculate the failure rate at all possible ball locations.
4: Prune possible ball locations.
5: Greedy legalize ball positions.

The first step of the algorithm consists of determining the optimal location of C4 balls using quadratic wirelength optimization. This consists of solving the following quadratic program:

$$\min \frac{1}{2}\sum_{i}^{n}\sum_{j}^{m} w_{ij}\left(px_i - bx_j\right)^2 + \left(py_i - by_j\right)^2 \quad (10)$$

where w_{ij} is the weight of the connection between block j and ball i obtained from the device net-list, n is the number of pins, m is the number of blocks, px_i and py_i are the x and y coordinates of ball i, and bx_i and by_i are the x and y coordinates of block j.

If a net in the device net-list only contains a single block and pin, then the weight of the connection between the block and pin is 1. However if the size of the net is greater than 2, then the clique model [18] is used to calculate the weight between each block and pin within the net.

After the optimal pin locations have been found the algorithm creates a grid of possible C4 ball locations using the minimum pitch distance between C4 balls, and the length and width of the chip. Once the grid has been created, the failure rate due to CTE mismatch between the substrate and die is calculated for every possible C4 location using a temperature

map generated from the block layout and the fatigue models detailed in Section III. The next step of the algorithm consists of removing all possible ball locations that have potentially low reliability. There are two methods that can be used for pruning bad ball locations. The first method consists of specifying a minimum number of cycles to failure for all the balls in the design. Given this value the algorithm will remove all possible ball locations with a value below that specified value for being a candidate location for a data I/O ball. The other method of pruning consists of specifying a number of possible ball locations (N) that will not be used. The algorithm then removes the N possible ball locations with the lowest number of cycles to failure. The final step of the algorithm consists of legalizing the ball placement from the first step. Legalization consists of assigning a ball to one of the possible ball locations on the grid array calculated in step 2 minus the locations removed from step 4 of the algorithm. The legalization employed by our algorithm is a greedy procedure which consists of 2 steps. In the first step, for each possible ball location the number of balls closest to that location is tabulated and placed in a bin. The possible ball location bins are then sorted so that the bin with the largest number of balls is first in the list. Finally, for each bin in the list, the balls are greedily place as close to its optimal location avoiding overlaps with other pins and the locations that have been removed for reliability issues.

B. Floorplanning

The location of I/O C4 balls affects the floorplan and vice-versa. Consequently it is important to co-optimize floorplanning and pin placement.

For our baseline floorplanner we used a simulated annealing algorithm that is similar to other modern floorplanners [19]. Our simulated annealer has three moves: interchange two blocks by swapping both sequence pairs, displace a single block by swapping one pair in a single sequence pair and rotation of a single block.

The cost function for the floorplanner is:

$$cost = \alpha \cdot \text{area} + \beta \cdot \text{HPWL} + \eta \cdot P \tag{11}$$

where area represents the floorplan area, HPWL is the half perimeter wirelength, P is a value calculated from the power density distributions on the chip, and α, β, η are the different weights associated with each value. The power density cost was added to the floorplanner cost function to determine the effects of thermal-aware floorplanning on C4 ball reliability.

During floorplanning calls are made to the quadratic stress-aware pin placement at a specified frequency so as to co-optimize for both wirelength and reliability. For our experiments, the floorplanner called the quadratic placement method at a 5% frequency rate. This value was chosen empirically. The stress-aware pin placement could be called for every move for the floorplanner, but this would lead to significant increases in runtime.

V. Experimental Setup

We implemented our floorplanner in C++. It uses a simulated annealing algorithm with a sequence-pair (SP) representation [20]. For thermal analysis, we integrated with HotSpot 5.0 using the default parameters. We use the more accurate grid mode for the temperature simulations for our final results. The conjugate gradient method was used for solving the quadratic program in our placement algorithm. Our results are run on a CentOS 5.1 Linux system with a 2.6GHz AMD Opteron processor and 8GB of memory.

We use the larger GSRC benchmarks for our experiments due to the small number of pins found in the smaller benchmarks. The GSRC benchmarks do not have actual dimensions or power information and both of these parameters dramatically affect the performance of the heat sink and overall chip cooling. For benchmark dimensions, we scaled all the benchmarks to be in range of medium to large area chips (0.5-2cm^2). To do this, we assume that the dimensions for the GSRC benchmarks are in tenths of a micron. The aspect ratio of soft blocks is constrained to the limits specified in the respective benchmarks (0.3 to 3.0).

For block power information, we randomly generate power numbers using design power densities similar to the predicted 65nm node in [21]. The power densities used are $750\frac{W}{cm^2}$, $250\frac{W}{cm^2}$ and $25\frac{W}{cm^2}$ with corresponding frequencies of 15%, 45% and 40%. The mean is therefore $235\frac{W}{cm^2}$ and there is approximately a $3.2\times$ difference between the average and maximum power density as observed in [21].

The positions of the C4 balls are constrained to a grid with a minimum ball pitch which of 100μm [17]. The number of C4 ball locations within the grid is actually larger than the number of I/O pins required by the chip so as to allow for some flexibility in their placement. Up to 10% missing balls is common in many commercial chips.

Since the temperature map of a floorplan is strongly dependent on the amount of whitespace available, we perform fixed-area floorplanning for fair comparisons. Consequently, the area cost during annealing is the area outside of the fixed area. We do not place constraints on the floorplan aspect ratio, however. For the experiments in the subsequent sections, we used a fixed area that is 10% larger than the total area of all blocks. All the results presented are mean values for 100 simulated annealing runs.

VI. Experiments

We conducted three sets of experiments to analyze the effectiveness of our reliability based floorplan methodology. The first experiment consisted of only doing HPWL optimization (η = 0, no pruning of candidate pin locations) to serve as a baseline for the other two sets of experiments. The results of the HPWL optimization are shown in Table I. The creep rate ($\dot{\epsilon_c}$) reported in the table is the maximum for all the bumps, and the number of cycles to failure (N_f) is calculated from that creep rate. The second set of experiments consisted of doing HPWL optimization in conjunction with temperature optimization (no pruning). The results of these

experiments are detailed in Section VI-A. The final set of experiments consisted of doing HPWL, temperature and reliability optimization concurrently. The results are summarized in Section VI-B.

A. Temperature Optimization

The analysis presented in Section IV showed that large temperatures significantly affected the creep rate of C4 solder balls. To determine the significance of high peak temperatures on the creep rate of C4 balls, temperature optimization experiments (no pruning) were conducted. The results of these experiments, also depicted in Table I, show that temperature optimization during floorplanning can significantly increase the reliability of C4 balls. Temperature optimization was able to decrease the average maximum creep rate over all benchmarks by 85%, and increase the lifetime of C4 balls by a factor of 12× even with just a 3K decrease in peak temperature. This significant decrease can be explained by examining the location of hotspots in HPWL vs Temperature optimized floorplans. In HPWL optimized floorplans, blocks with large power densities have a possibility of being placed on the periphery of the chip causing significant hotspots due to the adiabatic boundary conditions. Consequently if a pin is placed in the vicinity of that block it will have a high creep rate, since it is located at the edge of the chip and is subjected to high temperatures. In temperature optimised floorplans, blocks with large power densities have a much lower probability of being placed at the periphery of the chip since that would lead to large temperatures. Consequently there are fewer hotspots located on the periphery of the chip, the area of the chip that is susceptible to large creep rates.

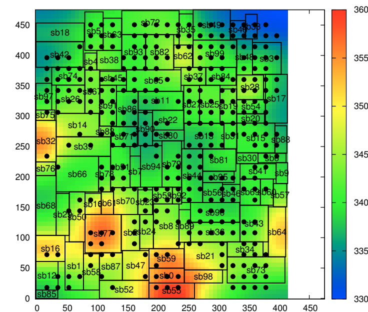
B. Reliability Optimization

C4 ball reliability optimization experiments were conducted to determine if pruning bad C4 ball locations could provide improvements over thermal-aware floorplanning. The results of these experiments, depicted in Table II, show that pruning bad C4 ball locations can lead to significant improvements in the C4 bump reliability, even when compared to thermal-aware floorplanning. The ball reliability optimization was able to decrease the average maximum creep rate over all benchmarks by a factor of 94%, and increase the lifetime of C4 balls by a factor of 47× as compared to HPWL only optimization. These improvements came without a significant increase in HPWL wirelength (3.6%) or runtime (1.36×) as compared to HPWL only optimization.

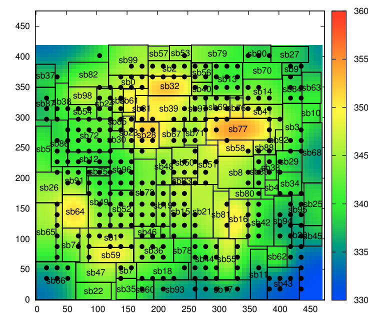
An example plot of n100 with and without reliability optimization is shown in Figure 4. Figure 4(a) corresponds to HPWL-only optimization. The HPWL for this placement is 176537, the maximum temperature is 358.5K, the maximum creep rate is 1.64E-3, and the runtime is 26.16s. The creep rate corresponds to a number of cycles to failure value of 5349. Figure 4(b) corresponds to HPWL, temperature and reliability optimization. The HPWL for this placement is 179834, the maximum temperature is 352.9K, the maximum creep rate is 1.40E-4 and the runtime is 29.86s. The creep rate corresponds to a number of cycles to failure of 63083. These figures show that by using reliability optimization along with temperature optimization, the lifetime of C4 balls can be greatly improved with a modest increase in HPWL and runtime. The reliability aware floorplanning tends to not place pins in hotspots, especially those located on the edge, while the HPWL optimization will tend to place pins closer to their respectively blocks even if it means placing a pin in a position where it can fail rapidly due to thermal cycling.

TABLE II
PIN PLACEMENT OPTIMIZATION RESULTS

	HPWL, Temperature and Reliability Optimization				
Bench.	HPWL	Max T	$\dot{\epsilon_c}$	N_f	Time(s)
n30	59230	357.49	1.06E-04	82925	10.58
n50	107321	364.29	3.89E-04	22597	17.35
n100	182905	353.59	1.94e-04	45310	33.98
n200	379144	355.14	3.26e-04	26964	142.63
n300	600652	358.38	4.90e-04	17939	327.58
Mean	3.6%	3.96 K	**0.06**×	**47**×	1.36×



(a) HPWL Optimization Only



(b) HPWL, Temperature and Pin Optimization

Fig. 4. Example Results for n100

VII. CONCLUSIONS

Modern ICs have large I/O requirements due to the steady increase in their size and complexity. As a result, more modern ICs are using flip chip package solutions as opposed

TABLE I
TEMPERATURE OPTIMIZATION RESULTS

	HPWL Only					HPWL and Temperature Optimization				
Bench.	HPWL	Max T(K)	$\dot{\epsilon_c}$ (1/s)	N_f	Time(s)	HPWL	Max T(K)	$\dot{\epsilon_c}$ (1/s)	N_f	Time(s)
n30	54420	364.64	1.44E-02	609	9.39	54899	360.29	5.37E-04	16367	9.69
n50	103339	368.68	2.62E-02	336	14.57	104039	365.42	1.44E-03	6100	16.67
n100	177711	357.87	3.17E-03	2770	27.46	179513	355.37	5.83E-04	15077	32.56
n200	371960	357.68	2.65E-03	3312	95.96	374971	356.21	7.36E-04	11946	128.15
n300	596528	359.49	4.28E-03	2054	185.94	603540	358.34	8.88E-04	9903	283.84
Mean	0%	0K	0 1/s	0×	0×	1.7%	2.55K	**0.15**×	**12**×	1.25×

to wire bonded solutions to meet the I/O requirements. One major disadvantage of flip chip solutions is the reliability of solder balls since they are susceptible to failure from thermal cyclic fatigue. In this paper we propose a thermal fatigue model for solder balls in flip chip packages that can be used for chip package co-optimization. We used our model to quickly evaluate candidate C4 ball locations to guide reliability floorplanning. Using our reliability floorplanner on GSRC benchmarks, we were able to significantly increase the lifetime of C4 balls, even when compared to thermal-aware floorplanning. Thermal-aware floorplanning was able to increase the lifetime of C4 balls by $12\times$ on average compared to only HPWL optimization, which was congruent with our expectations due to the large dependency of creep rate on temperature. However, our proposed quadratic pin placement algorithm was able to improve on thermal-aware floorplanning significantly, as the increase in the lifetime of C4 balls was $49\times$ on average compared to only HPWL optimization. These improvements came with a modest 3% increase in wirelength and a $1.36\times$ increase in runtime. In the future we would like to verify our model against FEA simulations and also to extend our work to consider multiple temperature maps.

ACKNOWLEDGMENTS

This work was supported in part by the National Science Foundation under grant 0720913 and a Special Research Grant from the University of California, Santa Cruz. Any opinions, findings, and conclusions or recommendations expressed herein are those of the authors and do not necessarily reflect the views of the NSF.

REFERENCES

[1] G. Pascariu, P. Cronin, and D. Crowley, "Next generation electronics packaging utilizing flip chip technology," in *Electronics Manufacturing Technology Symposium, 2003. IEMT 2003. IEEE/CPMT/SEMI 28th International*, July 2003, pp. 423–426.

[2] http://www.nvidia.com/object/io_1215037160521.html.

[3] K. Sheth, E. Sarto, and J. McGrath, "The importance of adopting a package-aware chip design flow," in *Design Automation Conference, 2006 43rd ACM/IEEE*, 0-0 2006, pp. 853–856.

[4] R.-J. Lee, M.-F. Lai, and H.-M. Chen, "Fast flip-chip pin-out designation respin by pin-block design and floorplanning for package-board codesign," in *Design Automation Conference, 2007. ASP-DAC '07. Asia and South Pacific*, Jan. 2007, pp. 804–809.

[5] C.-Y. Wang and W.-K. Mak, "Signal skew aware floorplanning and bumper signal assignment technique for flip-chip," in *Design Automation Conference, 2009. ASP-DAC 2009. Asia and South Pacific*, Jan. 2009, pp. 341–346.

[6] H.-Y. Hsieh and T.-C. Wang, "Simple yet effective algorithms for block and i/o buffer placement in flip-chip design," in *Circuits and Systems, 2005. ISCAS 2005. IEEE International Symposium on*, May 2005, pp. 1879–1882 Vol. 2.

[7] C.-Y. Peng, W.-C. Chao, Y.-W. Chang, and J.-H. Wang, "Simultaneous block and i/o buffer floorplanning for flip-chip design," in *Design Automation, 2006. Asia and South Pacific Conference on*, Jan. 2006, pp. 6 pp.–.

[8] J.-W. Fang, I.-J. Lin, Y.-W. Chang, and J.-H. Wang, "A network-flow-based rdl routing algorithmz for flip-chip design," *Computer-Aided Design of Integrated Circuits and Systems, IEEE Transactions on*, vol. 26, no. 8, pp. 1417–1429, Aug. 2007.

[9] H.-C. Tsai and W.-R. Jong, "The significance of underfill on the ic packages subjected to temperature cyclic loading," *Journal of Reinforced Plastics and Composites*, vol. 26, no. 12, pp. 1211–1223, 2007. [Online]. Available: http://jrp.sagepub.com/cgi/content/abstract/26/12/1211

[10] J. H. Lau and S. H. Pan, "Creep behaviors of flip chip on board with 96.5sn-3.5ag and 100in lead-free solder joints," *The International Journal of Microcircuits and Electronic Packaging*, vol. 24, no. 1, 2001.

[11] A. Schubert, R. Dudek, H. Walter, E. Jung, A. Gollhardt, B. Michel, and H. Reichl, "Reliability assessment of flip-chip assemblies with lead-free solder joints," in *Electronic Components and Technology Conference, 2002. Proceedings. 52nd*, 2002, pp. 1246–1255.

[12] J. Pang and D. Chong, "Flip chip on board solder joint reliability analysis using 2-d and 3-d fea models," *Advanced Packaging, IEEE Transactions on*, vol. 24, no. 4, pp. 499–506, Nov 2001.

[13] D. R. Frear, J. W. Jang, J. K. Lin, and C. Zhang, "Pb-free solders for flip-chip interconnects," *JOM Journal of the Minerals, Metals and Materials Society*, vol. 53, no. 6, pp. 28 – 33, 2001.

[14] M. Stan *et al.*, "Hotspot: A dynamic compact thermal model at the processor-architecture level." *Microelectronics Journal*, pp. 1153–1165, 2003.

[15] C. Y. Lee, C. Wong, and E. Suhir, *Micro- and Opto-Electronic Materials and Structures: Physics, Mechanics, Design, Reliability, Packaging*, 1st ed. Springer, 2007.

[16] D. Chau, C.-P. Chiu, J. Torresola, S. Prstic, and S. Reynolds, "Experimental method of measuring c4 die bump temperature for electronics packaging," in *Thermal and Thermomechanical Phenomena in Electronic Systems, 2004. ITHERM '04. The Ninth Intersociety Conference on*, June 2004, pp. 91–95 Vol.1.

[17] K.-N. Tu, *Solder Joint Technology: Materials, Properties, and Reliability*, 1st ed. New York: Springer, 2007.

[18] Y. Zhan, Y. Feng, and S. S. Sapatnekar, "A fixed-die floorplanning algorithm using an analytical approach," in *ASP-DAC '06: Proceedings of the 2006 Asia and South Pacific Design Automation Conference*. Piscataway, NJ, USA: IEEE Press, 2006, pp. 771–776.

[19] J. Cong, J. Wei, and Y. Zhang, "A thermal-driven floorplanning algorithm for 3D ICs," in *ICCAD*, 2004, pp. 306–313.

[20] H. Murata, K. Fujiyoshi, S. Nakatake, and Y. Kajitani, "VLSI module placement based on rectangle-packing by the sequence-pair," *TCAD*, vol. 15, no. 12, pp. 1518–1524, Dec 1996.

[21] G. M. Link and N. Vijaykrishnan, "Thermal trends in emerging technologies," *ISQED*, pp. 625–632, 2006.