Importance of Integrated Electrical and Thermo-Mechanical Co-optimization for Future Package Designs

Aditya Vaidya
Intel Corporation
aditya.s.vaidya@intel.com

Yinglei Ren
Intel Corporation
yinglei.ren@intel.com

INTRODUCTION

The world is becoming increasingly impatient and demanding. There is need for fast internet, social media pages to load quickly, and have in-person feeling with colleagues while working from home. On a serious note, processing large amounts of data in the shortest amount of time has become the norm in a world dominated by artificial intelligence (AI), machine learning and large amounts of data in general. A direct consequence of this is on the bandwidth (BW) requirements for different communication protocols to and from a microprocessor. In recent years, the signaling speed for standard protocols like Memory and Peripheral Component Interconnect Express (PCIe) have been increasing at an exponential rate to get higher BW as shown in Figure 1 below [1,2].

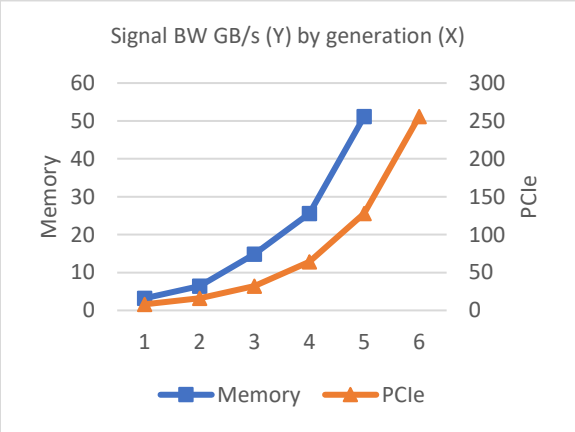


Figure 1 Evolution of signal speed.

With higher speeds, signal performance becomes more sensitive to energy reflection (e.g., caused by geometry discontinuities) and noise coupling. Previously, the geometries and material set were defined by the mechanical and reliability requirements for the package. However, in the future, the three branches (electrical, mechanical and reliability) must work together right from the beginning to ensure the product satisfies all the functional requirements. This article covers how the choices driven by mechanical and reliability considerations impact High-Speed IO (HSIO) performance and how a compromise needs to be made to meet demands from both areas.

MECHANICAL AND RELIABILITY CONSIDERATIONS

The increased demand for compute and speed is resulting in proportional growth of silicon area [3, 4]. The area scaling coupled by the coefficient of thermal expansion (CTE) mismatch between the silicon and the organic package causes higher deformation or warpage of the package when it sees different temperatures. For a ball grid array (BGA) package which is surface mounted (SMT) to a printed circuit board (PCB), the warpage at peak reflow temperature (~260°C) is of relevance. If this warpage is high, then there can be defects during the SMT process leading to rework or scrap. As a result, careful assessment is done to optimize the geometry and materials used in package to ensure that there is a robust SMT process. Some of the enablers to have low warpage are using thicker core in the package substrate, lower CTE materials, looser BGA pitch and larger BGA ball size. These enablers either lower the package warpage by stiffening it / lowering the CTE mismatch or buy more process margin at SMT to manage larger warpage of the package.

Beyond manufacturability, the second consideration is reliability. Some applications like industrial and automotive have longer life requirements (10-15 years) compared to personal computing (PC) (5 years); also, the communication and networking segment has a wider ambient temperature range (-40°C to 50°C, depending on local weather) compared to PC which is in the range of 24°C to 30°C. These harsh

requirements put different interfaces within the package at risk of cracking or delamination during the use. One such example is cracking of the solder joints between the package and PCB due to temperature cycling. At higher temperatures, the solder joints are in tension whereas at colder temperature they are in compression. This cyclic nature of stresses results in cracking of the joints and can cause functional failures in the field. To address this concern, some joints are assigned as sacrificial in a package. These joints are expected to fail during the life of the package and protect the joints which are required for proper functioning of the package. The number of sacrificial joints needed depends on the reliability requirements of the package as well as general architecture of the package like silicon/package size, and BGA pitch/pattern/size. In many applications the package or joints are reinforced by use of an adhesive. This adhesive can be dispensed just in the corner/edges of the package or cover the entire shadow of the package encapsulating all the solder joints as shown in Figure 2. The adhesives contain fillers in a polymer matrix which share some of the stresses on the solder joints and hence improve the reliability.

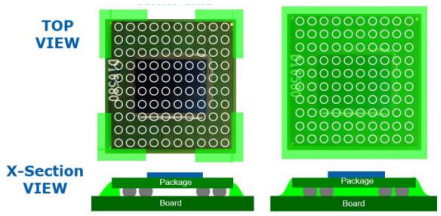


Figure 2 Image showing use of corner (left) or full (right) adhesive.

IMPACT ON HIGH-SPEED IO SIGNALS

In the previous section, we discussed some of the manufacturing and reliability challenges along with the mitigation options employed, e.g., bigger BGA size, sacrificial joints, and adhesives. In this section, we will describe how those choices impact the electrical behavior of the package.

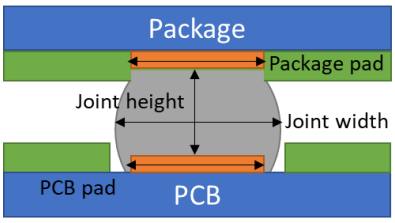
Signals connect two or more chips. The whole channel includes silicon, package, socket or BGA joints, PCB and in some cases cable connectors. At each interface, there is a transition in geometry and change in material properties which result in impedance discontinuity. Impedance is the ratio of the voltage and current waves at any one position on the channel and it is dependent on the frequency at which the signals operate. Impedance discontinuity results in wave reflection and energy loss in the transmitted signal. The bigger the impedance discontinuity, the bigger the energy reflection. This energy loss directly impacts signal performance at the receiver side which can be measured as eye height and width [5].

Geometry change at the solder joints at the package-PCB interface is one typical source of impedance continuity. Table 1 below shows an example of how ball size and adhesive impact HSIO performance.

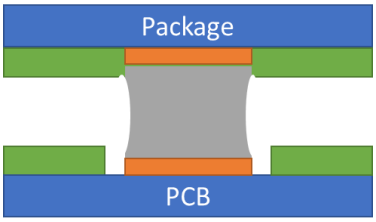
Table 1 Impact of ball size and adhesive on electrical performance of high-speed signals.

Signal	**20→24mil BGA**	**20→24mil BGA with adhesive**
Memory	No impact	Downgrade 1 speed bin
Ethernet	5mm lower route length	25mm lower route length
PCIe	13mm lower route length	43mm lower route length

20mil BGA size without adhesive (i.e., ball is surrounded by air whose dielectric constant is 1) is used as the benchmark. When ball size changes from 20mil to 24mil, lower impedance occurs at the BGA area, leading to bigger impedance delta between the BGA and package/PCB. The effect can be seen as shown in Figure 3. A larger ball, 24mil (Figure 3a) will cause more bulged solder joint whereas a smaller ball, 20mil (Figure 3b) will more likely cause an hourglass-shaped joint. Note that a larger ball is better for SMT, while a smaller ball is better from an HSIO perspective.



(a)



(b)

Figure 3a Cross-sectional view of a typical solder joint with the key dimensions marked, and 3b Ideal joint shape preferred from HSIO perspective.

The adhesive dielectric constant is approximately 3.5. When it is used around HSIO balls, the impedance becomes even lower than the package and PCB impedance.

In both cases, bigger energy reflection happens, degrading signal performance. For memory, speed bin downgrade happens for the case of 24mil ball with adhesive. For Ethernet and PCIe, to avoid signal transfer error, the maximum supportable trace length is reduced adding additional constraints to the package design.

Besides energy reflection, crosstalk (Xtalk) can also impact signal performance. Xtalk is unwanted signals in a communication channel caused by transfer of energy from another circuit (by leakage or coupling). When two transmission channels are close to each other, electromagnetic energy can couple from one channel (aggressor) to the other (victim) leading to energy loss on the aggressor and unexpected noise on the victim. The closer the two channels, the bigger the Xtalk. Having ground (GND) next to the aggressor can help reduce energy leakage as most energy remains between signal and GND.

Assuming same BGA ball pitch, the bigger-sized ball will be closer to the adjacent ball, resulting in higher Xtalk between different signals. As shown in Figure 3a, the bulged ball will have higher Xtalk risk versus an hourglass shaped joint in Figure 3b.

Another impact on Xtalk could come from the number of sacrificial joints needed. If more sacrificial joints are needed to meet the reliability requirements, the number of IO or critical power/ground pins will reduce assuming same package size and ball pitch. In some cases, the designers use the sacrificial joints as ground pins for shielding purposes. As the sacrificial joints are expected to crack during the life of the package, this strategy can increase the risk of Xtalk between signals next to sacrificial joints which are assumed as a ground.

ELECTRICAL-MECHANICAL CO-OPTIMIZATION EXAMPLE

Figure 4 shows the package SMT capability (Y-axis) versus the different BGA sizes (X-axis). SMT capability is the package warpage value (including magnitude and shape) up to which we can have an SMT process with reasonable yield. The graph also shows the modeled warpage for the package under study (red line). Based on initial mechanical modeling, a 24mil BGA size is recommended for the package so that the capability matches the modeled warpage. However, as seen earlier, there is an impact of using a 24mil BGA on the allowable routing length for Ethernet/PCIe, and a 20mil BGA is preferred. However, with a 20mil BGA the SMT capability is much lower than the modeled warpage which puts the SMT process at risk. Based on this, a compromise of using a 22mil BGA works for both electrical and mechanical purposes. This brings a slightly better SMT margin and is also a middle ground from an electrical perspective. For the 22mil BGA, the modeled warpage is also lowered (green line) by making some geometric changes to the package which does not impact the electrical performance.

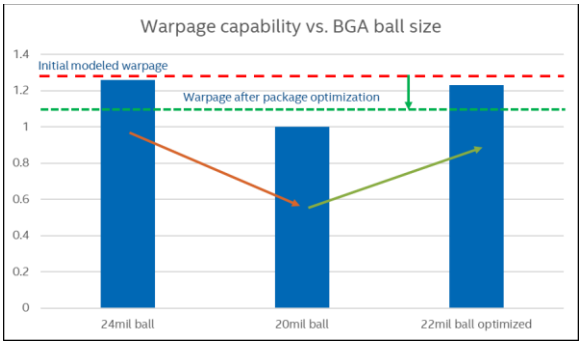


Figure 4 Warpage capability versus BGA size

In the example above, there is a compromise achieved between the electrical and mechanical teams to ensure that all the key performance metrics are met. However, it is not always true that there will be a middle ground in all situations. This emphasizes the fact that the mechanical and electrical teams need to be coordinated from the early definition of the package to identify enablers that would help each area without compromising the other.

SUMMARY AND CALL TO ACTION

This article discusses the parameters which impact mechanical and electrical performance of a package. In the past, when the signal speeds were low, these two areas could be in silo as there was plenty of margin. However, in future as the specifications get tighter and the silicon/package area increases, it is paramount that both areas are well aligned with each other right from the start of the design. This ensures that a thorough assessment of choices one team makes is analyzed by the other team to come to an optimized solution. This article only covered one example of this optimization, but there are several areas of collaboration all the way from the silicon to the PCB and everything in between.

Beyond collaboration, there is a call to action to the packaging community to help with this co-optimization problem. As mentioned in the earlier section, solder joint reliability can be a challenge for packages going into harsh-use conditions. For such requirements, it is necessary to have high reliability solder alloys so that the number of sacrificial joints can be reduced. Use of adhesives can lower this count today, but that degrades the electrical performance due to the high dk of these materials. So, it is prudent to develop adhesives which have lower dk without compromising the strength. The other aspect which was not explicitly discussed in this article is the overall package thermal management. The increased reliability risk for different interfaces in the package are due to CTE-mismatch-driven flexure as the package goes through different temperature cycles during the life. The source of heat is silicon and the ambient conditions which can increase the magnitude of the temperature change. Having efficient cooling of the package will lower the overall temperature fluctuations and hence improve reliability. Drawing more heat from the silicon side of the package will keep the rest of the package (BGA balls) cooler and improve their reliability.

ACKNOWLEDGMENTS

The authors would like to thank our colleagues from Intel Corporation - Chenghai Yan, Maoxin Yin, Xinjun Zhang, Jianting Li, Kusuma Matta, Puneesh Puri, Gopal Mundada, Gautam Medhi, Marco Cartas, Guru Arakere, Weizhe Li, Yanjie Zhu, James Mccall, Oluwafemi Akinwale and Ravi Mahajan - for their inputs and contribution to this article.

REFERENCES

[1] https://www.crucial.com/articles/about-memory/difference-among-ddr2-ddr3-ddr4-and-ddr5-memory.
[2] https://www.tomshardware.com/reviews/pcie-definition,5754.html.
[3] Stojcev Mile, Toki Teufik, and Milentijevi Ivan. (2005). The limits of semiconductor technology and oncoming challenges in computer microarchitectures and architectures. Facta universitatis - series: Electronics and Energetics. 17. 285-312. 10.2298/FUEE0403285S.
[4] Sun Yifan, Agostini Nicolas, Dong Shi, and Kaeli David. (2019). Summarizing CPU and GPU Design Trends with Product Data.
[5] Stephen Hall and Howard Heck. Advanced Signal Integrity for High-Speed Digital System Design. Hoboken, NJ: Wiley, 2009, ch. 4.