REVIEW

Signal and Power Integrity Challenges for High Density System-on-Package

Nathan Totorica   Feng Li*
Department of Electrical and Computer Engineering, University of Idaho, Moscow, Idaho, 83844, USA

ARTICLE INFO

Article history
Received: 18 February 2022
Revised: 17 March 2022
Accepted: 28 March 2022
Published Online: 30 May 2022

Keywords:
System on package (SoP)
System in package (SiP)
System on chip (SoC)
Through silicon via (TSV)
Signal integrity
Power integrity
Thermal integrity

1. Introduction

Portable electronics, IoT, and 5G continue to drive demand for smaller form factors with increased performance and reliability. While this trend is not new, many niche solutions have attempted to address it over the years. In the heyday of Moore’s law, integrated circuit (IC) scaling was doubling roughly every two years. This gave rise to integrating whole modularized systems on a single IC, known as system-on-chip (SoC). While SoCs are widely used in many applications today, SoC comes with a lot of scaling and performance issues as Moore’s law continues to slow down and integration of digital, RF, optical, and MEMS-based components into a single IC creates a much more complex fabrication process. In some specific market sectors, such as aerospace, the need for high-performance systems in small footprints is high. Still, low volume demand may not justify SoC solutions’ considerable R&D and manufacturing costs. While modern transistor densities are at historical peaks, they have continued to

*Corresponding Author:
Feng Li,
Department of Electrical and Computer Engineering, University of Idaho, Moscow, Idaho, 83844, USA;
Email: fengli@uidaho.edu

DOI: https://doi.org/10.30564/ssid.v4i2.4475
Copyright © 2022 by the author(s). Published by Bilingual Publishing Co. This is an open access article under the Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0) License. (https://creativecommons.org/licenses/by-nc/4.0/).
drift from Moore’s law. Apple’s A14 mobile processor, for example, is reported at 11.8 billion transistors in an 88 mm² footprint (~134 million transistors per mm²), and even with these impressive densities this SoC comes in less dense than what a 5 nm node would predict \[1\]. Innovation in package and system design has become increasingly important to address “Moore’s Stress” \[2\].

Several decades ago, a modularized approach was taken to mimic SoC like behavior that instead combined several discrete components into a miniaturized system known as a multichip module (MCM) \[3\]. Innovation based on this idea has opened the door to many other types of advanced packaging technologies, such as system-in-package (SiP) and system-on-package (SoP) that have shown as promising solutions to addressing “Moore’s stress”. Taking the MCM idea a step further, SiP brings multiple, potentially dissimilar, die into a single package, and expands the packages responsibility to include interconnection. This has enabled many creative solutions to package dissimilar ICs into single discrete packages \[4-6\]. Continued integration brought about SoP, bringing even more functionality into the package itself. SoP contains all interconnection, as well as embedded passive components on a single substrate. Packaging solutions are a new form of Moore’s law that will continue to miniaturize systems and increase performance without relying strictly on transistor scaling. The role of packaging has expanded into new territory, with additional responsibilities on top of thermal dissipation, electromagnetic interference (EMI) shielding, and mechanical and environmental protection. These attributes have become especially desired in market sectors, such as medical, aerospace, and mobile devices, where size and weight are critically important.

Fundamental design challenges become increasingly detrimental to system performance as complexity grows, all while in ever-smaller footprints. Higher density layouts, along with faster edge rates, result in increased EMI that can generate crosstalk in the connector and signal traces. While SoP addresses many of the shortcomings seen amongst packaging solutions presented here, careful attention must be taken to address the fundamental signal, thermal, and power integrity issues.

2. Comparison of Advanced Packaging Solutions

Advances in package technology, such as MCM and SoC, originally pursued options that integrated functionality in a two-dimensional (2D) way. Continued innovation led to SiP and SoP expanding into three-dimensional (3D) territory by stacking discrete die vertically \[7\]. The history and background of the leading advanced packaging solutions are described in detail and compared in Table 1.

2.1 Multichip Module (MCM)

Innovation of Multichip modules was mainly driven by the needs of the aerospace industry, often called hybrid circuits where small footprints and high performance are a must \[9\]. Due to the low volume demand, novel solutions like MCM may be used instead placing discrete components onto a single substrate, two-dimensionally, facilitating interconnection. This technique is used when specialized components may be required and are not practical to integrate into a single, monolithic form. Components are physically positioned closer, which results in better signaling. Interconnects are typically built on ceramic (MCM-C), thin-film (MCM-D), or laminate (MCM-L) substrates, depending on the targeted application. The earliest form, MCM-C, is often used due to its higher capacitance and better power rail support. Improvements in areas like performance and cost quickly followed with deposited MCM-D and cheap laminates MCM-L. MCM-D allows use of reduced permittivity dielectrics that enable lower resistivity of interconnects, which in turn boosts performance. Additionally, since MCM-D uses photolithography methods smaller feature sizes are possible that enable higher density signal interconnections. Lastly the MCM-L approach utilizes commonly uses printed circuit board materials like FR-4 for interconnection and is generally the simplest approach. MCM proved very useful as a low-cost solution without forfeiting significant performance requirements. MCM modules employing low-temperature cofired ceramics (LTCC) and organic thin film have been remained especially popular in recent decades in RF and millimeterwave applications \[9-11\]. Multichip modules are the most generalized of the other advanced packaging solutions.

2.2 System-on-chip (SoC)

Integrating more functionality onto a single IC certainly has an understandable appeal. Smaller footprint, lighter weight, and lower power consumption are just a few of the desirable attributes that come with SoC designs. Execution speeds are extremely fast with minimal trace lengths and propagation latency between logic blocks. If implemented effectively, SoC offers the most cost-effective system solution that is high volume manufacturable. While transistor scaling remained closely in step with Moore’s law, this seemed to be the solution of the future. However, as systems grow more complex many challenges that accompany SoC are realized when trying to incorporate a variety of functionalities, such as RF, analog, digital, or optics, all into a single IC. Further, integrating simple passives onto IC incurs a large die size hit, and fabricating
high Q-factor passives for RF applications is extremely difficult \[12\]. Other drawbacks include the fact that complex, low-yielding fabrication processes in certain blocks of the SoC can cause entire chip to be scrapped. Despite these drawbacks, there are many advantages to combing similar blocks into SoC, in conjunction with other SiP or SoP systems to realize full system functionality.

2.3 System-in-package (SiP)

SiP can incorporate 2D or 3D layouts, with the latter configuration reaching near chip-scale package sizes. 3D stacking techniques are often used with highly regular die as is generally the case for memory ICs \[13,14\]. The SiP solution can also exploit the ability to place dissimilar, high-yielding die into a single discrete package. Commonly, high-density memory like High Bandwidth Memory (HBM) is packaged along with computational Application Specific Integrated Circuits (ASIC) \[15\]. This stacked configuration can use a variety of high-density interconnect (HDI) technologies, from wire bonding, through silicon via (TSV) or flip-chip techniques for I/O interconnection. Many 3D stacking configurations rely on TSV or wire bonding, but as need for more I/O pin counts increase, fanout techniques are being employed to meet these high bandwidth requirements using wafer level packaging (WLP) \[16\]. Some other techniques exploit cheaper redistribution layers (RDL) to facilitate interconnection \[17\], which eliminates need for expensive WLP equipment and removes costly interposer TSV process steps. The flexibility to customize SiP systems to the application and manufacturing needs is a major reason for its widespread adoption. SiP designs are often able to be treated like a normal, single packaged IC which poses a big advantage in assembly simplification compared to MCM.

2.4 System-on-package (SoP)

Taking a step further, SoP combines SiP style stacked die with integrated passive elements, overcoming many of the disadvantages faced by all three MCM, SoC, and SiP techniques. Integrated passive components can utilize thin-film deposition techniques that allow thinner, more reliable performance than typical discrete passives. SoP effectively creates a complete PCB functionality into a single discrete footprint. The reduced signaling path between die decreases trace losses, which is exacerbated in RF applications \[3\]. Generally, inductors, capacitors, and antennas are much more easily realized on package substrate than on silicon, especially for high Q-factor requirements. Additionally, this allows more flexibility than in an SoC design. Ultimately, many of the aspirations of SoC can be achieved using SoP techniques. While SoP is very similar to SiP in concept, SoP promises much more functionality. Subsequently, this increased complexity means additional challenges must be addressed to realize these benefits \[18\]. Due to these challenges SoP has been slow to take hold in industry. In recent years however, 5G has created a resurgence of attention around SoP, with many applications exploiting mmWave based 5G frequencies to build extremely compact RF components for a highly integrated SoP solution \[19-21\].

### Table 1. Electronic package technology comparison

<table>
<thead>
<tr>
<th>Package Technology</th>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
<tbody>
<tr>
<td>MCM</td>
<td>Modularized</td>
<td>Larger footprint</td>
</tr>
<tr>
<td></td>
<td>Robust</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cost effective</td>
<td></td>
</tr>
<tr>
<td>SoC</td>
<td>Smallest footprint</td>
<td>Long design/verification cycles</td>
</tr>
<tr>
<td></td>
<td>Highest speed</td>
<td>Difficult testability</td>
</tr>
<tr>
<td></td>
<td>Low power</td>
<td>Potentially low yield</td>
</tr>
<tr>
<td>SiP</td>
<td>Simpler design &amp; verification</td>
<td>Complex packaging equipment/assembly</td>
</tr>
<tr>
<td></td>
<td>Shorter time to market</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2.5/3D stacking</td>
<td></td>
</tr>
<tr>
<td>SoP</td>
<td>Flexibility</td>
<td>More complex than SiP</td>
</tr>
<tr>
<td></td>
<td>Small system footprint</td>
<td>SI/PI challenges</td>
</tr>
<tr>
<td></td>
<td>2.5/3D stacking</td>
<td></td>
</tr>
</tbody>
</table>

3. Signal and Power Integrity Issues

Along with the desire for smaller, more compact electrical systems, demand for faster speeds is also increasing. Faster speeds, and in turn, faster signal rise times bring a myriad of signal integrity issues. Commonly referred to as dV/dt and dI/dt noise, the faster signals change, the more EMI and subsequent noise is generated. Along with these challenges, designers have less and less space for routing traces. Traces are packaged closer together in package and PCB, resulting in very high-density routing and higher coupling between signal traces. Classic signal integrity problems become more apparent as systems get larger and timing margins get smaller. The most common issues are discussed in the following sections.
3.1 Reflections

Through a signal’s entire path, it will see impedance at every point along the way. One of the most common signal integrity issues that arise in high-speed systems is signal reflection [22]. Reflections occur when there is a sudden change in impedance and is commonly simplified as an equivalent transmission line feeding a load impedance, as seen in Figure 1.

\[ \tau = \frac{Z_L - Z_0}{Z_L + Z_0} \]  

Additionally, this can be caused by any change in signal medium or trace geometry. Typical examples include vias, edge connectors, or solder joints. Signals that see an extremely high impedance, like an open, will have 100% reflection with a 0° phase shift in the signal. The other extreme is that if the signal sees zero impedance, like a short, 100% of the signal will be reflected, but this time with a 180° phase shift. Since reflections are a well-known issue, designers have come up with different techniques to minimize them. The primary way is to add termination to signaling paths. A few examples of terminations schemes include series, parallel, or double termination. For series termination, reflections will bounce off the receiver side and terminate at the transmitter side. Series termination can be useful to eliminate crosstalk effects. Parallel termination is a very common termination scheme where the transmitted signal will terminate at the receiving side and can be done by either terminating transmission line to the power supply VDD or to the ground. Double termination can be parallel terminations on both source and receiver side, or a combination of series and parallel termination schemes. Double termination schemes are often used in high-speed, high performance serial link designs [23].

3.2 Crosstalk

Change in current or voltage through a conductive medium will create changing electric and magnetic fields. If a signal is propagating down one line, known as the aggressor line, and the magnetic field lines intersect another conductive medium, known as the victim line, a voltage and resulting current will be generated on the victim line. The intersecting magnetic field creates a relationship known as mutual inductance \((L_m)\) between the two conductors, and the intersecting electric field creates mutual capacitance \((C_m)\). The resulting signal generated on the victim line through coupling to the aggressor line is known as crosstalk. This can cause issues when the layout for two sensitive signal paths is extremely close together. The individual voltages are defined by Equations (2) and (3).

\[ \Delta V_L = L_m \times \frac{di}{dt} \]  
\[ \Delta V_c = C_m \times \frac{dv}{dt} \]  

Typically for transmission lines with long, wide return paths, the amount of mutual inductance and mutual capacitance is comparable. However, for non-uniform structures, like connectors or pins, inductive coupling dominates [24]. In the case of transmission lines, the crosstalk manifests differently at the near end versus the far end. This is known as near-end crosstalk (NEXT) versus far-end crosstalk (FEXT).

The polarity and magnitude can be different between the NEXT and FEXT noise signals, demonstrated in Figure 2, and is dependent on which type of EMI is dominant in generating the noise. The mutual inductive coupling will create a wave that propagates in either direction, one towards the far end and one towards the near end that are opposite in polarity. The mutual capacitance will also generate a wave in either direction but with the same polarity. The combination of these interactions will describe how the FEXT and NEXT noise will look at either end of the victim transmission line.
3.3 Power Distribution Network (PDN)

The power distribution network is one of the most important pieces of the system. Careful attention must be paid to the PDN to ensure a low impedance profile across a broad frequency spectrum. Ideally, a chip expects a constant near-flat DC voltage supply. High inductance in power and ground connections can cause unwanted power rail fluctuations due to the inability for PDN to respond to quick current draw, causing a voltage drop. This is especially true with many drivers toggling simultaneously, often called simultaneous switching noise (SSN) or delta I noise. The delta I noise is described by Equation (4) where $L_{\text{off}}$ is the PDN effective inductance and $N$ is the number of switching drivers.

$$\Delta V = N L_{\text{off}} \frac{\Delta I}{\Delta t} \quad (4)$$

This effect is nearly impossible to avoid but ensuring impedance remains below target values across the bandwidth of the design is the only way to guarantee correct system operation.

4. SoP Design Challenges

Regardless of the package methodology, traditional design issues like crosstalk, impedance mismatch, rail collapse, and thermal emission will be present, if not exacerbated with tighter layouts and faster clock speeds. Since SoP designs share a subset of design challenges as SiP, many real-world solutions are used interchangeably to illustrate how to mitigate these issues. For example, Qualcomm’s newer Snapdragon series expands SoC design to a SiP, supporting up to 1.8 GHz per core, with over 400 individual components. The Snapdragon SiP contains memory, RF, audio, and more all in one compact footprint, easily integrated into mobile phone systems [25]. For applications such as this, traditional management techniques like increased spacing or reduced signal edge rate may not be sufficient or possible under the size and speed specifications. Additionally, a well-known problem for packaged systems is the Known Good Die (KGD) issue. If SoP or other stacked die packages are built with one defective die, then the other good die and components may be scrapped along with it. This can pose yield issues for SoP designs, resulting in high fabrication expenses. Sufficient die level testing is necessary to avoid this issue. Many of the techniques discussed in this section will help mitigate some other common issues for a more robust SoP system design.

4.1 Signal Integrity

Inherent benefits of SoP include integration of interconnects into the package, thin-film construction of passive components, and easier integration of dissimilar die. Reducing overall signaling length helps to improve signal integrity. Interconnect paths can use a variety of options like wire bonding or TSV techniques. While TSV is more expensive than wire bonding, it offers shorter interconnects and improved signaling. Shorter channel lengths give less distance for parallel traces to couple, reducing crosstalk and insertion loss. As previously discussed, reduced signaling loss is a major benefit, especially for RF modules in mobile devices where prolonged battery life is a major benchmark. Additionally, SoP enables the ability to fabricate thin-film passive components. Thin films enable smaller, better-performing passive devices compared to standard discrete components. Fabricating high-Q passive components is extremely difficult on a die and takes a large amount of die real estate.

Post silicon issues are costly, making simulations key to proactive system debugging. Proper simulation allows issues to be found well before anything is physically made. Simulating entire systems can be difficult and resource-intensive, though. If that expense is not an option, using intuition built off the signal and power integrity concepts discussed previously can help pinpoint worst-case susceptible signal paths. As discussed by Yang et al. [26], looking at a handful of these most critical paths help build confidence that the overall design will meet specifications. Generally, all the most sensitive paths, such as clock trees, should be examined thoroughly in simulation.

A technique that is commonly used for transmitting signals over longer traces (i.e. chip to chip) is pulse shaping [27]. The idea is to shape the signal in such a way as to better interact with the characteristics of the channel and transmit more accurately to the receiver. This idea could potentially be implemented to decrease the signal rise time by using a more sinusoidal shaped wave to eliminate some of the highest order harmonics associated with a square wave. This could help reduce some of the dV/dt, dI/dt noise that is generated by faster signal switching. Generally, slower rise times that still meet specifications (e.g., setup and hold timings) are always better from an EMI perspective.

4.2 Power Integrity

For both board and package level PDN, the overall impedance must be kept under target value through the entirety of the operating frequencies. Especially, low return path inductance reduces the effects of SSN and ground bounce. Wider conductors for power and ground planes will help reduce loop inductance. The threat of rail collapse can be mitigated by using decoupling capacitors...
strategically sized and placed to lower PDN impedance, as well as compensate for voltage droop caused by simultaneous switching \[28\]. Typically, several different types of decoupling capacitors are used in PDN design (bulk, local, package, embedded, and on-die) to cover specific frequency ranges. A comprehensive pre and post route power integrity analysis is outlined \[29\] to provide better performance of PDN. A critical piece in the study is to include implications of capacitor effective series resistance (ESR) and effective series inductance (ESL); if left out of the analysis, these can have unforeseen effects at the high frequencies. The technique outlined by Venkatesan et al. \[14\] uses embedded TSV-Caps to further tune impedance below the target levels at the high frequency range. MIM or trench capacitors are often embedded into interposer to tune at these high frequency harmonics but require significant silicon area and additional process steps. The TSV-Cap approach embeds the capacitor around the TSV structure in the pitch between adjacent TSVs. The technique, shown in Figure 3, allows increase of bandwidth with minimal increase in silicon area required.

Additionally, the effects of the TSV-Cap on cross-talk were analyzed to ensure that these structures do not cause unwanted noise on signal lines due to coupling of power lines through the TSV-Cap structures. Comparing insertion and return loss metrics between device with and without TSV-Cap shows comparable performance for insertion loss at low frequencies and a general improvement at higher frequencies, with return loss generally improved across entire frequency range.

**Figure 3.** (a) ASIC-HBM packaging illustration. (b) The interposer elevation view. (c) Interposer cluster interconnect. (d) PI/SI and GND line \[14\].

**4.3 Thermal Management**

Thermal management is extremely important to consider when designing any system, and this is especially true as SoP die stacks grow higher and bandwidths meet 5G specifications. A widely noticed trend in CPU clock speeds is that newer generations are not making the same massive jumps in operating speed as seen in the past. This complex problem is deeply involved with transistor physics and physical size, but an additional factor is a difficulty cooling for clock speeds greater than 3 GHz-5 GHz. Thermal management has become an integral part of package and system design that cannot be ignored. Although thermal dissipation is one of the package’s most important roles, thermal considerations need to start with electrical design. Excessive current density generates heat, leading to electrical failure or even mechanical package warping \[28\].

While heat generation is impossible to avoid, novel thermal management solutions are pursued to ensure adequate cooling for stacked die. During the design phase, transient thermal simulations, as discussed by Yoo et al. \[30\], ensure systems have a relatively even heat distribution across workloads. Careful attention to the order of the stack up, as presented by Mathur \[31\], can mitigate the increased heat generated from 3D stack configurations. Using a 3 GHz CPU as a test vehicle, modularized chiplets are stacked three-dimensionally in a logic-over-memory configuration. Comparison between a stacked configuration where thermal implications are considered versus when they are not showing a significant difference in overall heat experienced. A pre-simulated stacked design may experience half as much temperature increase as compared to a design where thermal implications are not considered. When compared to a traditional 2D design, the thermal efficient 3D design may experience as little as a 6 °C increase while decreasing the footprint by 23%. For applications that can tolerate the temperature increase, this is a significant reduction in space occupied. In general, these studies echo the importance of thorough simulation to understand thermal implications.

Concentrated heat areas can cause issues, especially at the transistor level where elevated temperatures cause higher leakage currents and increased scattering. A major consideration for 3D configurations is the possibility of different max temperature ratings, as is commonly seen with stacked logic and DRAM. For example, an ASIC may be rated up to 125 °C while HBM is only rated to 95 °C \[32\]. Ensuring all devices are operating within specification is a must to avoid throttling and subsequent performance hits. Proper cooling will increase the longevity and reliability of the system.

Thermal bottlenecks tend to appear towards the center and lower stacked die in a 3D configuration. Solutions to avoid this type of bottleneck, like the ICE-SiP technique as presented by Kim et al. \[33\], use highly thermal conduc-
ative “chimneys” to dissipate heat from the inner layers to a heat sink, as depicted in Figure 4. In this example, a DRAM die is stacked on a computational ASIC. The ICE-SiP technique shifts the DRAM off center to provide room for an alternative, lower thermally resistant path made of silver (Ag) for heat to transfer instead of dissipating through the DRAM. Other configurations may use silicon (Si) instead of silver due to its higher thermal conductivity, ~120 W/mK versus ~50 W/mK, respectively. However, based on simulation results, the Ag-based chimney performed best, coupled with high-K epoxy compound molding (EMC). In contrast, the Si chimney required an adhesive to bond to the die and ultimately suffered from the low thermal conductivity of the adhesive.

Figure 4. Geometry of the “ICE-SiP” [33].

Increasingly, more advanced methods of thermal management are being researched for next generation electronics, including novel approaches in magnetic cooling [34] and nanofluids [35]. Nanofluids are a form of liquid cooling that employs nano particles suspended in a base fluid to increase thermal conductivity beyond what is achievable with traditional air or water mediums. The particles are often a metal but as described by Zumbühl [34], many different compounds have been researched, including carbon-based nanotubes. The base fluid is typically comprised of water, ethanol glycol, or oil. Generally, these results have shown very promising although this technique has not yet been widely adopted in applications.

5. Conclusions

The trend toward more compact and mobile systems is likely to continue, if not grow, in upcoming years. SoP innovation has offered a promising solution to help with these scaling trends. While the fundamental signal, power, and thermal challenges are only going to get harder to overcome, the solutions discussed here can help mitigate these problems for more robust, reliable products.

Conflict of Interest

There is no conflict of interest.

References

process for laser prototyping of microwave circuits in LTCC technology. IEEE Transactions on Components, Packaging and Manufacturing Technology. 5(6), 723-730.
DOI: https://doi.org/10.1109/TPCMT.2015.2434273

DOI: https://doi.org/10.1109/RFC.2008.4561409

DOI: https://doi.org/10.1109/IMW.2017.7939084


DOI: https://doi.org/10.1109/S3S46989.2019.9320698

DOI: https://doi.org/10.1109/ECTC.2016.65

DOI: https://doi.org/10.1109/ECTC.2018.00199

DOI: https://doi.org/10.1109/ICEPT.2008.4606942

DOI: https://doi.org/10.1109/TVT.2020.2989451

DOI: https://doi.org/10.1109/TMTT.2021.3076066

DOI: https://doi.org/10.1109/MWSYM.2019.8700931


DOI: https://doi.org/10.1109/ICEPT52650.2021.9567976


DOI: https://doi.org/10.1109/PCSPA.2010.106

DOI: https://doi.org/10.1109/ICCCREC.2015.7337052

DOI: https://doi.org/10.1109/ITherm45881.2020.9190945

technology. IEEE 70th Electronic Components and Technology Conference (ECTC). pp. 541-547.
DOI: https://doi.org/10.1109/ECTC32862.2020.00091
DOI: https://doi.org/10.1109/EDAPS47854.2019.9011675
DOI: https://doi.org/10.1016/j.enconman.2018.07.047