

Customer Service Note

Package-on-Package (PoP) User Guide

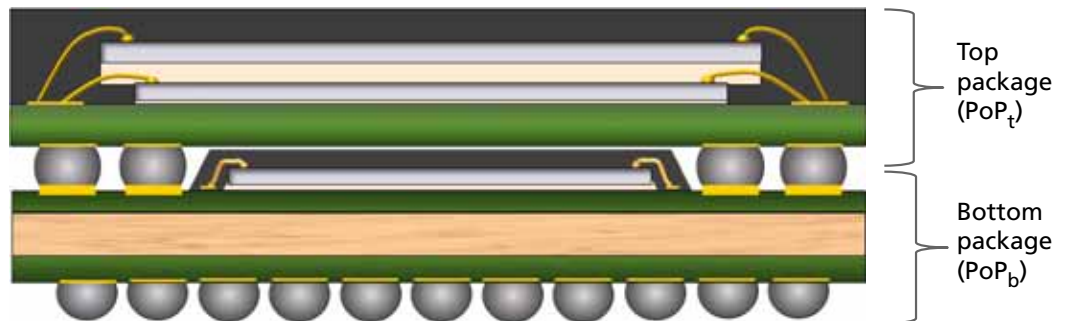
Introduction

Package-on-package (PoP) semiconductor packaging technology requires unique considerations in both the up-front design and in the manufacturing process. This document provides several well-established guidelines for PoP design and assembly, but it does not cover all possible variables that can affect the performance of PoP design or manufacturing yield and should not be interpreted as a recipe. The end user must discuss PoP assembly considerations and implementation with all the parties involved in the development process, from beginning to end.

PoP Overview

PoP is a packaging technique used to integrate compatible chip modules in a cost-effective manner while maintaining thinness and a reduced footprint. While PoP systems can be composed of numerous package stacks and chip-module combinations, the market has precipitated a blueprint involving two chip modules: a top memory package (PoP_t) and a bottom applications processor package (PoP_b). As shown below, this configuration enables a large area to route signals to the PCB for the applications processor while maintaining short routing paths to the upper memory module.

Figure 1: Conventional PoP System



Note that the PoP system is composed of markedly dissimilar packages, often from different vendors, and that these packages are physically interfacing with one another. To avoid conflict in the assembly process, it is critical to have clear, concise, and accurate communication among vendors.

PoP Benefits

PoP offers system designers several key advantages over other memory packaging solutions. Primarily, these are an XY-size PCB footprint that is smaller than a two-package, side-by-side alternative; a simpler PCB design; and improved frequency performance enabled by the direct connection of the memory package to the logic package. These same benefits can also be achieved by integrating logic and memory functionality into a single-die design (system-on-a-chip, or SOC) or by connecting the logic and memory die together in one package (system-in-a-package, or SIP).

However, like the conventional two-package, side-by-side alternative, the PoP solution isolates the design, manufacturing, and testing of the logic and memory components into separate activities. This separation provides a much faster time-to-market for the final logic-plus-memory solution, more flexibility in the choice of memory type and density, and a lower total cost by avoiding the compounding yield-loss impact often seen in SOC or SIP alternatives.

Overall, PoP offers small XY size and electrical performance benefits similar to SOC or SIP, but also provides the flexibility, faster time-to-market, and lower-yielded cost benefits of two separate package solutions.

PoP Design Considerations

Micron PoP_t products may be multichip packages (MCPs), combining various combinations of DRAM, phase change memory (PCM), NAND, and NOR die to provide a complete memory system. The PoP_t products are tailored for specific logic chipsets with JEDEC or a custom electrical signal assignment matrix. The system designer's preliminary selection of a specific logic chipset and the complementary memory types and densities is a complex process that is beyond the scope of this CSN, except to say that many of the initial considerations are the same, regardless of whether the system is implemented with a PoP solution or with separate logic and memory packages.

The logic package XY size is often determined by the logic die size and the count and pitch of the signals necessary to pass down to the PCB. This size, in turn, puts limitations on the count and pitch of the signals passed up to the memory and impacts which memory types and architectures can be located in the upper PoP.

In practice, XY sizes and memory interconnect signal pitches are emerging as defacto mechanical PoP standards. However, the electrical implementation details for each size still vary from logic supplier to logic supplier. For example, a 14mm x 14mm PoP with 152 memory balls at a 0.65mm pitch is meeting with marketplace acceptance, but currently has at least three different electrical pinouts for various memory configurations, as shown below.

Table 1: Common Top Package Configurations

Package Size (mm)	Pitch (mm)		
	0.65	0.5	0.4
	Total Ball Count		
12 x 12	128	168	216
14 x 14	152	240	N/A
15 x 15	160	N/A	N/A

PoP Signal Integrity

Signal integrity is not unique to PoP. (For more information, refer to Micron technical note “TN-00-20: Understanding the Value of Signal Integrity Testing.”) One advantage of PoP is its ability to operate at higher speeds, but this requires careful simulation to verify proper operation. The simulations are complex; signal lines can be long and run through various package interfaces, making it difficult to achieve stable power. To complicate matters, physical constraints can limit ground plane options.

Both PoP_b and PoP_t vendors perform signal integrity assessments and analysis on their respective products. It is best to involve both package vendors early in the design process to ensure that signal and power integrity concerns have been properly addressed.

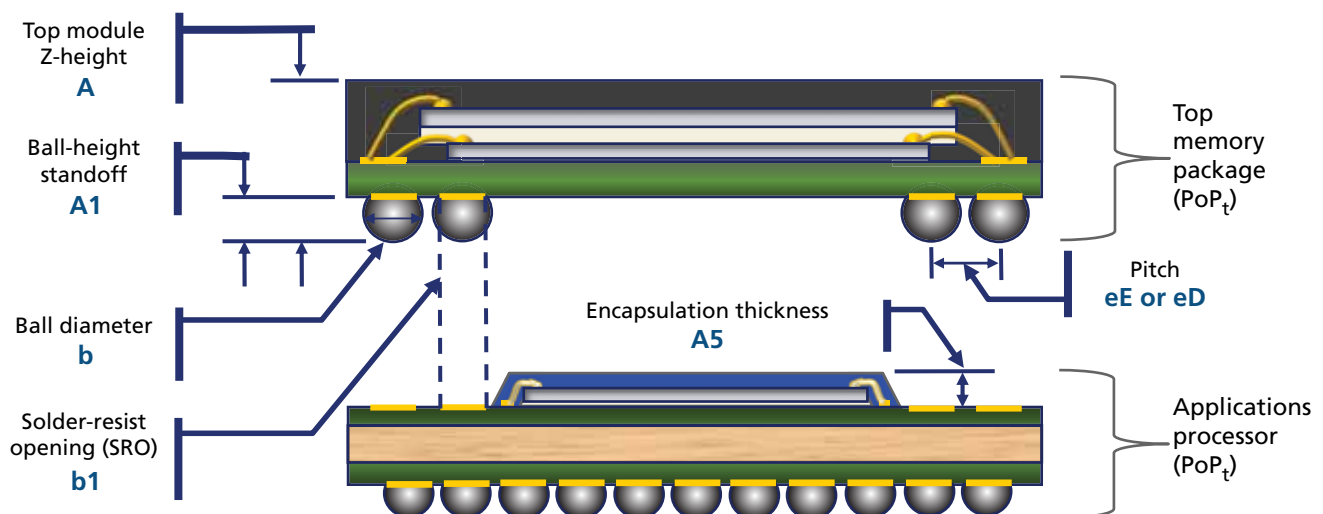
PoP Z-Height Calculations

Simply stacking one package on top of another would save XY space, but overall thinness would be sacrificed. PoP benefits from significant advancements in thinning package structures, including the silicon. (For more information on silicon thinning, see Micron technical note “TN-00-19: Thinning Considerations for Wafer Products.”)

Ultimately, end users do not physically see the internal structure advancements; they are not something end users need to be concerned about. On the other hand, it is critical that end users understand the external features of the package and how they fit together with each other in the PoP application to ensure good SMT yield.

PoP has two states: unassembled state (uncollapsed profile), shown in Figure 2, and assembled state (collapsed profile height), shown in Figure 3. The unassembled state dimensions are easily determined with accurate drawings provided by the package vendor.

Figure 2: Dimensional Elements of an Unassembled PoP

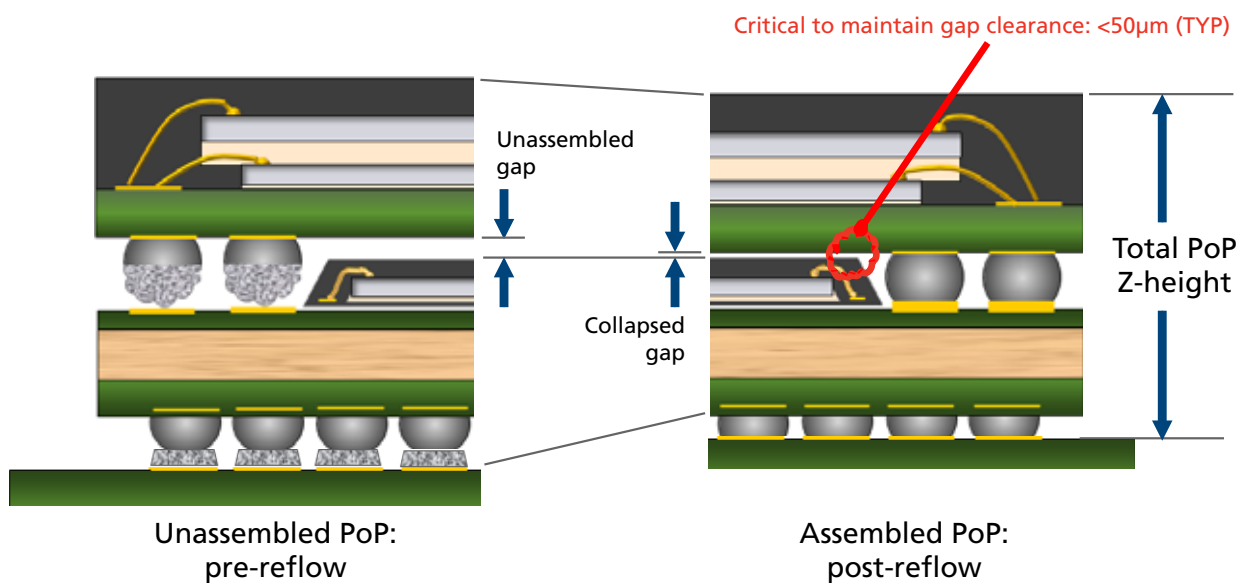


The assembled state PoP thickness, or Z-height, is determined by the post-reflow height of the soldered device or second ball collapse. Second ball collapse is defined as the solder joint collapse after PCB assembly reflow. This value is not present on package drawings because its dimension is dependent on the interaction of the PoP components

and the assembly process. The collapsed profile height and gaps can be assessed using ball collapse calculators. Work with your chip suppliers to understand ball collapse expectations for your application.

Note: An important note to end users and bottom package vendors: Second ball collapse standoff height must be considered not only at the time of design and component selection, but throughout the PoP development. Any manufacturing change can impact the final fit. A mounted PoP may have clearances of only a few micrometers. If second ball collapse is not accurately utilized in calculating assembled state clearances, it can result in significant SMT yield loss and a dramatic reduction in board-level reliability. Micron can work with end users and bottom package vendors to ensure proper package fit.

Figure 3: Dimensional Elements of an Assembled PoP



Caution Engineering analysis is recommended for each application to avoid the risk of package-to-package interference or clash.

PoP Assembly Flow Options

The two most common methods for assembling PoP are single-pass reflow (SPR) and two-pass reflow (commonly known as pre-stack). Each method has pros and cons that should be considered to achieve the best fit for manufacturing flow and supply chain concerns.

The SPR method, shown in Figures 4 and 5, is the simplest and perhaps least costly approach. The application processor (PoP_b) is placed onto the paste-printed PCB, and then the top memory package (PoP_t) is flux- or paste-dipped, placed onto the mating PoP_b, and finally sent through the reflow oven with the completed PCB.

Figure 4: Single-Pass Reflow Sequence

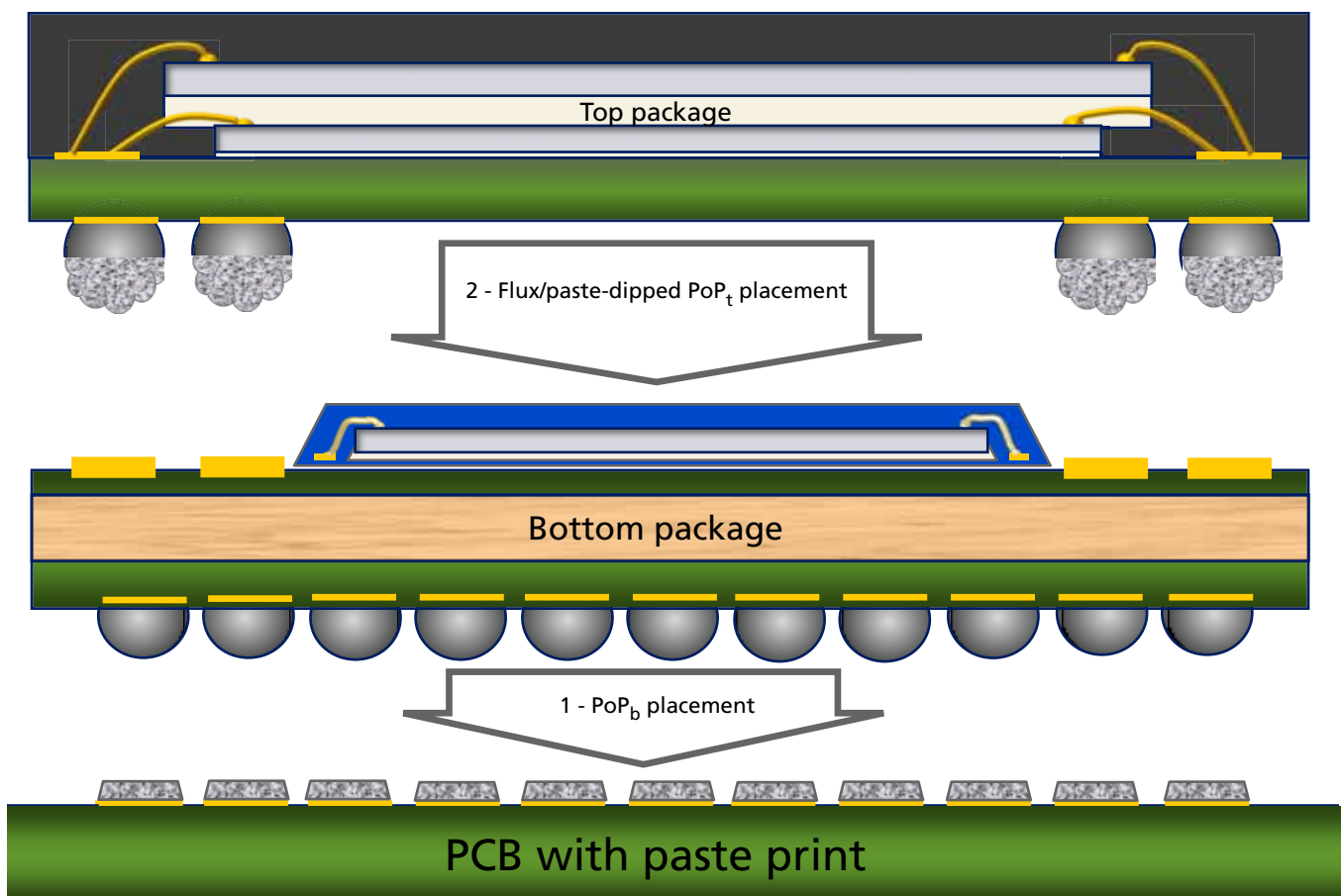
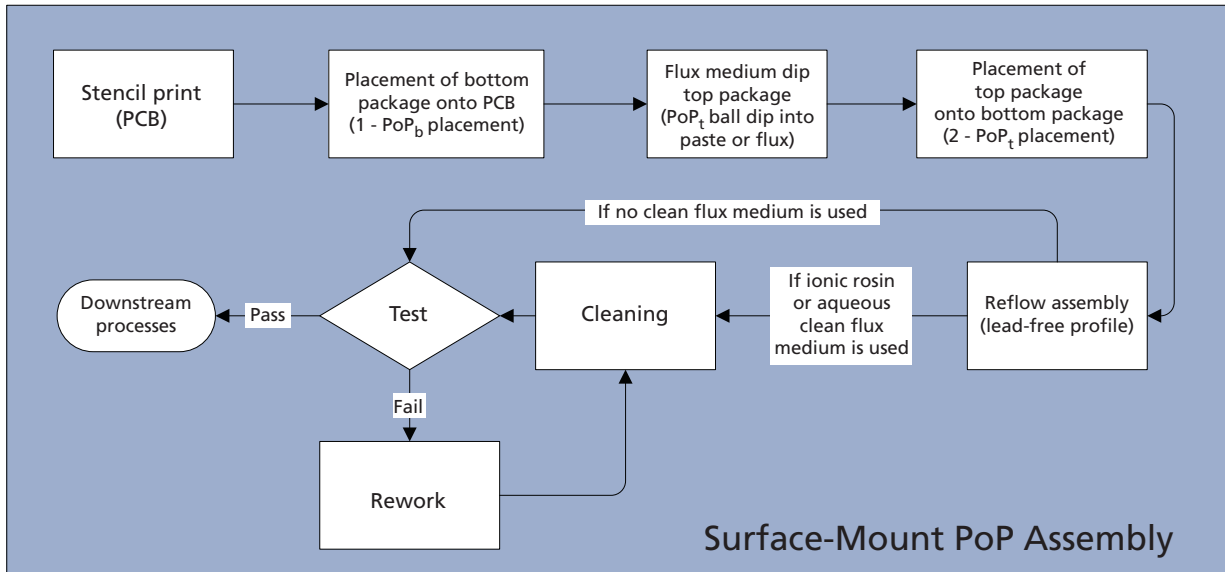


Figure 5: Single-Pass Reflow Process



Notes: 1. After all components are placed, the entire assembly is reflowed.

Pre-stacking is a method in which the PoP_b and PoP_t are mounted independently of the PCB, as shown in Figures 6 and 7. Compared to SPR, pre-stacking requires additional fixtures, process steps, manufacturing controls, and trained personnel. While the assembly is more complicated with the pre-mount step, this method does provide the ability to test the full functionality of the assembled PoP prior to final PCB assembly.

Figure 6: Pre-Stack Reflow Sequence

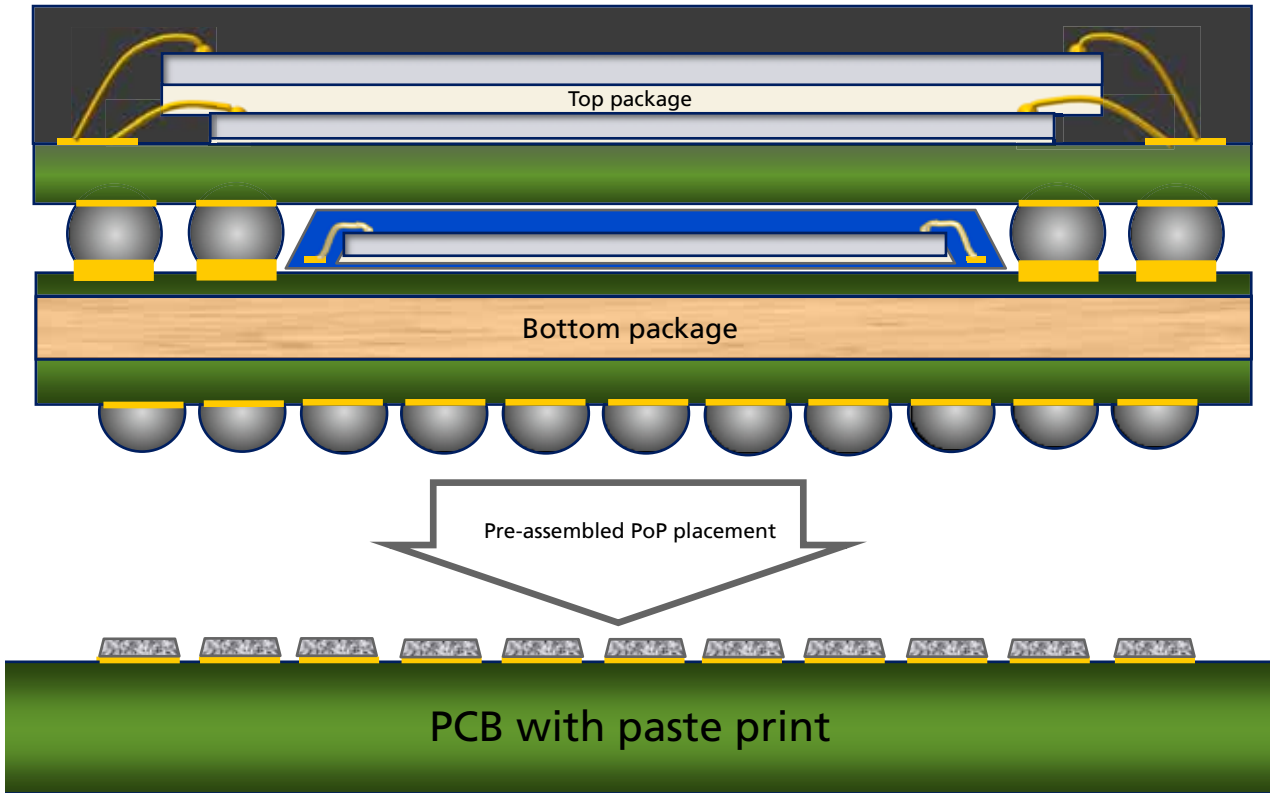
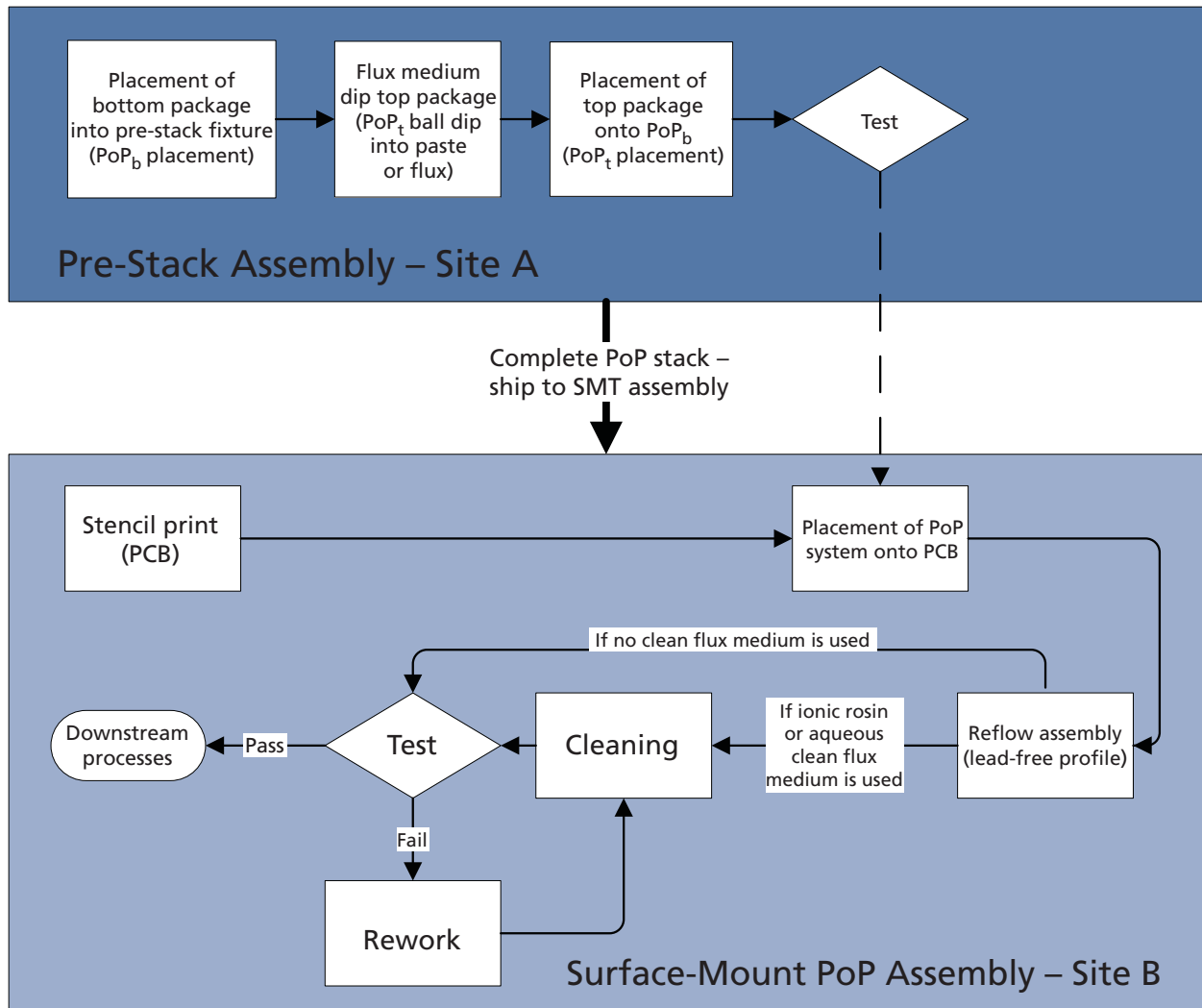


Figure 7: Pre-Stack Reflow Process



Notes: 1. The PoP is assembled separately, then placed onto the PCB for final assembly reflow.

Micron generally recommends single-pass assembly of the PoP due to its overall simplicity and history of achieving good manufacturing yields. Therefore, the guidelines presented in this CSN focus on SPR.

Top Package Mounting Preparation

A cross-company team of engineers has worked together to align all critical component dimensions. Nevertheless, component dimensions, unit placement, and thermal warpage are only geometric boundaries of the PoP assembly equation. Flux chemistry and convective heat also play an important role in reflow soldering. Even if the ball is placed on the pad throughout the reflow process (zero gap), good flux chemistry is still required to make a strong, high-yielding joint.

Micron recommends allocating sufficient resources and time to perform proper evaluation of different materials and process settings.

Fluxing Medium

The top memory device is not attached to a PCB, but to a thin individual package. This introduces a significant limitation because paste cannot be effectively stencil-printed onto the PoP_b, leaving ball dipping as the reasonable process option.

The preferred fluxing medium varies throughout the industry and is dependent on the PoP pad metallurgy (such as solder or NiAu) and pitch. Each type of dipping medium (flux or paste) has advantages and disadvantages that must be balanced depending on the end user's process.

PoP dipping materials are continually being improved upon. Trials with both flux and paste are recommended to help focus the evaluations on materials that are designed for ball dipping. Fine solder-sphere size, high tackiness, and linear Newtonian rheology are generally favored for consistent material transfer.

Table 2: Dipping Paste vs. Flux Performance

Pitch (mm)	Medium	
	Paste	Flux
0.65	++	+-
0.5	++	+-
0.4	+-	++
Dip-inspect	Easy	Difficult
All	Smooth, uniform appearance when spread, no lumps or waves in dipping tray	

The purpose of the ball-dip step is to transfer flux or paste to each ball. The efficiency of material transfer is critical to minimizing defects. As with most processes, too little flux medium leads to electrical opens, while too much leads to other challenges, such as shorting and inspection. Foreign material on components or in the flux wells also causes issues.

A target dipping depth of 50% of the ball height is typical. When evaluating flux mediums, experiment with ball-dip heights to determine an appropriate dip height and process margin.

Placement

Final placement of the PoP_t to the PoP_b is important to yield. Micron does not recommend a specific limit of placement tolerance, as it is a function of flux medium performance. The more precise the placement, the better the results. Target placement at no

less than 80% ball on pad. Placement requirements are well within most all modern chipsetting equipment capabilities. Consider local board fiducials for PoP_b and PoP_t, if necessary.

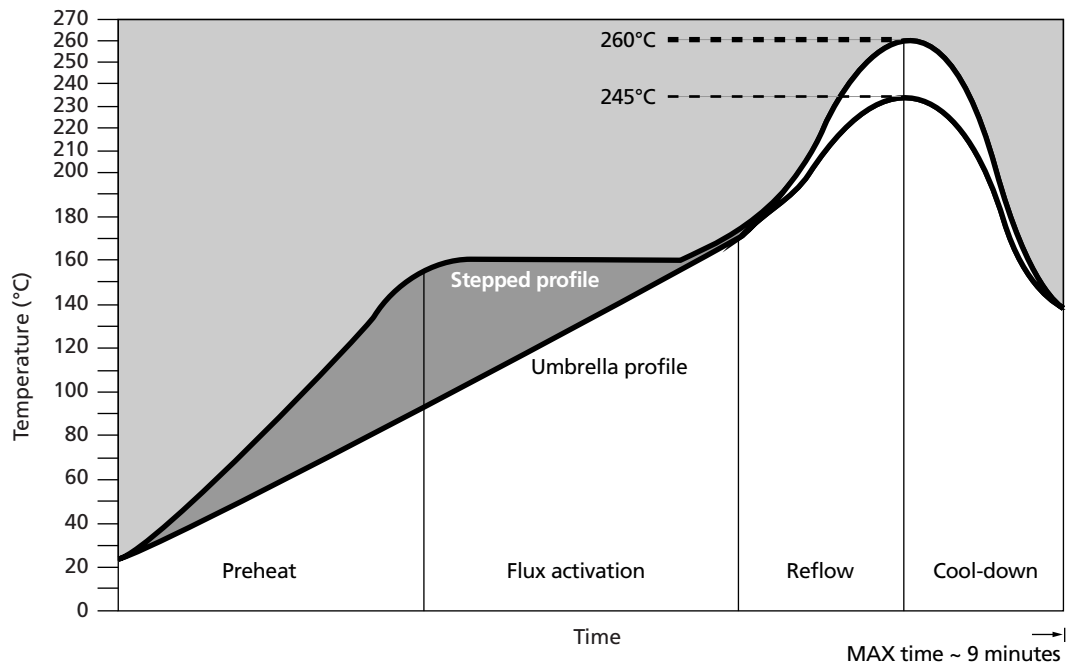
Reflow

PoP assembly reflow is similar to normal lead-free BGA profile requirements. The reflow specifics are largely dependent on the alloys and flux medium, as well as the board layout itself. It is important to recognize that the function of the profile is to activate the flux and wet the solder alloy to the pad. To achieve uniform board heat-up, the board layout should avoid pulling heat away from the solder pads, especially when designs implement small solder pads.

When adopting PoP technology, it is helpful to experiment with profiles. For PoP (and most chip-scale BGA), using a component profile with a 245°C peak temperature achieves two things. It minimizes thermal warpage interactions and preserves the PoP_t flux medium. Customers have had success with a wide range of profiles, but an umbrella profile (linear-ramp profile) tends to offer the greatest margin for SMT PoP yield. A stepped profile (soak-zone profile) is thought to exhaust most dip fluxes or pastes before the completion of wetting. Again, the acceptable profile is dependent on the specific board application and materials being used.

For more information, see Micron technical note “TN-00-15: Recommended Soldering Parameters Introduction.”

Figure 8: Reflow Profile Example



Reflow Atmosphere

Solder oxidation is the enemy of reflow soldering, and the fluxing medium must sufficiently protect the soldering surfaces from oxygen. When a flux medium is unable to support this function through the reflow profile, inert nitrogen atmospheres can offer

additional process margin, but not necessarily eliminate the defect. Fortunately, the industry is beginning to offer many dipping flux mediums that are easily capable of operating in an air atmosphere (nitrogen purge is not required).

Handling and Rework

As with most chip-scale packages (CSPs), PoP devices should be free of moisture and stored and handled in accordance with J-STD-033 and J-STD-020 procedures. Also, the maximum number of reflows that each device can withstand is a total of three (including rework and/or reballing) with the parts adequately reflowed.

When properly controlled, PoP systems can be reworked. Special rework and reball tools and fixtures have been designed to simplify and help control this process. Instructions are provided by the relevant manufacturers. It is critical that a keep-out zone area around the rework component is maintained to thermally isolate the adjacent components during the reflow process and to ensure that parts are free of moisture per supplier labels and instructions. However, as for all BGA packages, manual soldering, manual rework, and reballing are not recommended because handling is too operator-dependent.

Above all, reballing must avoid excessive heat during the redress process because it can damage the functionality of the semiconductor.

Thermal Warpage

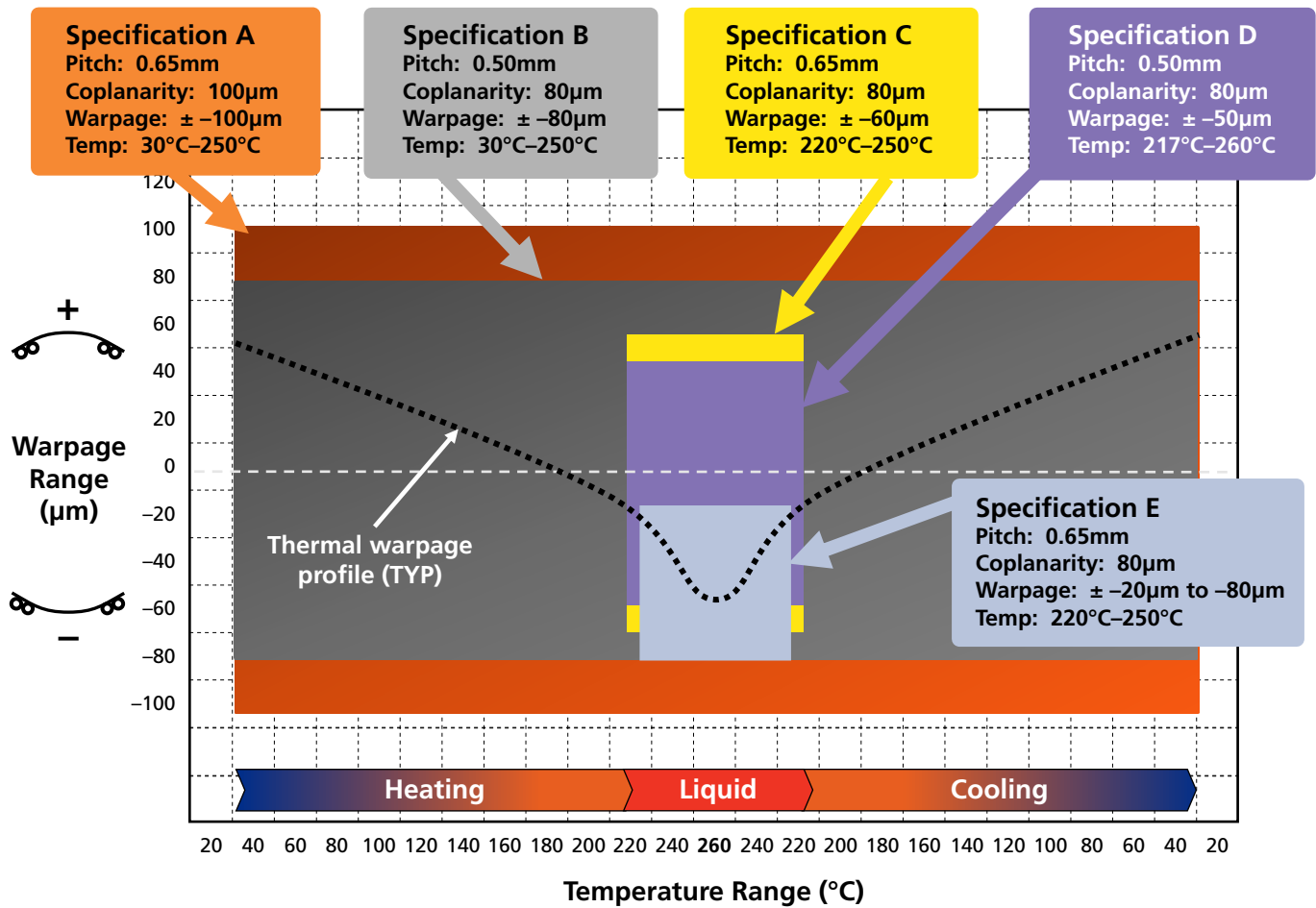
The semiconductor packaging industry has made a significant effort to understand package warpage, package warpage measurement, modulators of PoP warpage, and the impact on the final assembly yield.

In general terms, package warpage criteria are assigned a positive (convex) or negative (concave) value, representing the general package shape trend. When interpreting reported thermal warpage values, it is important to realize that this specification is a simplification of the three-dimensional package shape at various temperatures. Industry standards such as JEDEC JESD22-B112 or JEITA ED-7306 provide a good description of package warpage.

While a range of warpage values is specified to minimize the potential gap between the PoP_b pad and the PoP_t ball during reflow, it is probably more critical that the shape of the PoP_b and PoP_t match. The maximum allowable gap between the pad and ball is largely dependent on the flux medium chemistry selected; some will perform better than others.

Through reflow, the package experiences a wide range of temperatures. The solder balls' liquidus temperature range is the region where the PoP_t ball array will wet to the PoP_b pads, making this region the most important in controlling warpage. Figure 7 on page 8 shows some representative end user and PoP_b supplier specifications. Consult with industry experts and evaluate PoP processes to identify a cost-effective solution for your application.

Figure 9: Thermal Warpage Specifications for PoP Systems



Memory Die Temperature

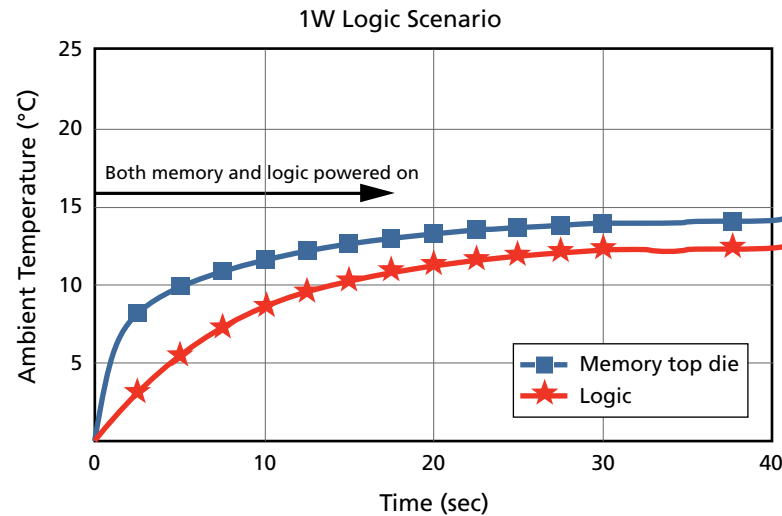
The silicon devices packaged in a PoP have operational temperature limitations; see Micron technical note “TN-00-08: Thermal Application Note.”

Unlike typical PCB layouts, in which heat-generating packages can be isolated to separate areas of the PCB, the PoP system locates the heat-generating applications processor in close proximity to the heat-generating memory. It is important to consider the implications of this configuration.

Table 3: Power and Temperature Relationships

Processor Power (W)	Memory Power (0.1W each die)	Processor Die Temperature (°C)	Memory Die Temperature (°C)	ThetaJA (C/W)
0.6	0.2	92	90	16.4
0.8	0.2	98	95	16.5
1	0.2	104	101	16.7
1.2	0.2	111	107	16.8
1.5	0.2	120	115	16.8
1.7	0.2	126	121	17

Figure 10: Modeled Memory Temperatures as a Function of Logic Power



In a typical BGA, heat is dissipated through the PCB and the surrounding air. In a PoP system, the two packages are in close proximity to each other, resulting in a small temperature gradient between them. (Further information can be found in Micron technical note “TN-10-08: LPDDR – Thermal Implications for Die Stacks.”) Also, the PoP_T-generated heat must sink through the PCB and ambient air (or case, if a thermal interface is placed between the PoP_T and the case).

Table 3 and Figure 10 show an example of the thermal modeling results of a PoP structure in a 92°C ambient environment. The memory die temperature is correlated to the logic die power level, reaches equilibrium in ~30 seconds, and can reach 105°C with logic power levels in the range of 1W.

The usage model of the particular system (% of time in the fully active mode, standby mode, or power-down mode and the duration of time spent in each mode) ultimately determines the temperature-performance capability required of the memory solution. Refer to the Micron technical note “TN-00-18: Upgrading Semiconductors for High-Temperature Applications” for further information.

Underfill for Board-Level Reliability

Inherently, PoP generally meets most board-level reliability criteria for most applications without the use of underfill. However, some applications do require additional margin. Micron does not generate underfill-specific data due to the wide variety of materials, applications, and objectives that need to be met by underfill. The industry does have publicly reported data on the underfilling of PoP systems. It is clear that the proper underfill must be selected in order to provide benefit; improper underfill can actually be a detriment.

Underfill trends can be summarized in three primary groups:

- **Non-underfill PoP system** – Joint failures tend to occur on the bottom package joints.
- **Underfill of PoP_b only** – The failure mechanisms shift to the PoP_t joints and provide minimal improvement to the overall PoP system reliability.
- **Simultaneous underfill of PoP_b and PoP_t packages** – Offers the greatest benefit.

Micron recommends that if underfill is to be performed, it should be performed on both the PoP_t and PoP_b packages.

Conclusion

Package-on-package technology has proven to be a flexible, cost-effective solution for the integration of logic and memory in a small XY space. While it simplifies the task of system PCB design, it effectively shifts integration into a shared responsibility between logic and memory suppliers. This integration is facilitated by the use of various specifications and standards, addressing elements such as physical dimensions, dimension changes across temperature, and electrical signal locations. PoP also uses advanced wafer- and package-thinning technologies, which can result in higher memory die operating temperatures.

Surface-mount recipe formulation requires experimentation with materials and heat, and adequate time should be allowed for this experimentation. System designers can reap the benefits and avoid the pitfalls of PoP by selecting memory and logic suppliers who are active in defining these specifications and standards and who have the existing tooling, experience, and expertise to deliver compatible, reliable PoP solutions.



Revision History

Rev. B	8/11
<ul style="list-style-type: none">• Converted from technical note to customer service note.	
Rev. A	3/11
<ul style="list-style-type: none">• Initial release.	