Semiconductor Technology from A to Z

Wafer Test

Wafer Test

Fundamentals of Wafer Testing

After all front-end process steps are complete, the wafer contains a large number of finished but still unseparated dies. Before the wafer is diced, every single die is electrically tested while still on the wafer – this step is known as wafer test or probe test.

The test serves two purposes: first, it identifies defective dies so they are not unnecessarily processed further in the subsequent assembly steps. Second, it provides early electrical characterization data, such as threshold voltages or leakage currents, which can be used for process control.

A test system essentially consists of three components:

- Wafer prober: a precision tool that positions the wafer and moves it die by die under the test needles
- Probe card: the actual contacting unit with fine needles that land on each die contact pads
- Tester: the electronic test equipment that generates test patterns and evaluates the circuit response

Since modern chips can have several hundred to several thousand contact pads, the probe card must feature a correspondingly large number of precisely positioned needles that reliably make contact without damaging the sensitive chip surface.

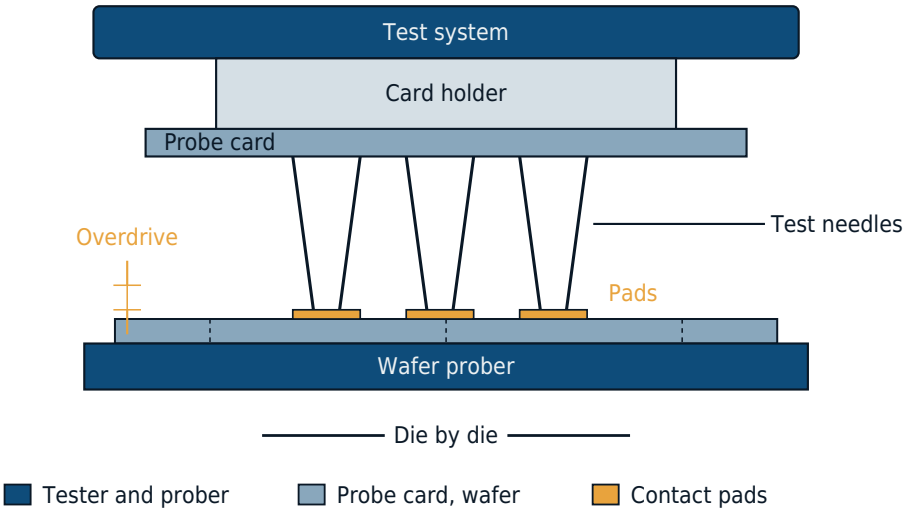
The Probe Card

The probe card forms the mechanical and electrical interface between tester and wafer. Older designs use free-standing needles made of tungsten or a beryllium-copper alloy, mounted radially on a ring. For very fine pad pitches, MEMS-based probe cards are increasingly used today, whose contact tips are manufactured using lithographic processes from silicon or metal – a method that allows significantly higher positioning accuracy.

When the needles land on the pads, a small but deliberate force (overdrive) is applied that breaks through the native oxide layer on the pad surface, ensuring a low-resistance electrical contact. However, excessive overdrive damages the pads and can harm underlying layers – tuning this force is therefore a critical process parameter.

Structure of the Wafer Test

The needles land on the pads before the wafer is diced



The needles are set down with a small force, the overdrive. It breaks through the native oxide layer on the pads and ensures contact. Too much overdrive damages the pads and the layers beneath them.

To increase throughput, modern test systems frequently contact multiple dies simultaneously (multi-site testing). This allows several chips to be tested in parallel, substantially reducing the test time per wafer.

Yield Mapping and Redundancy

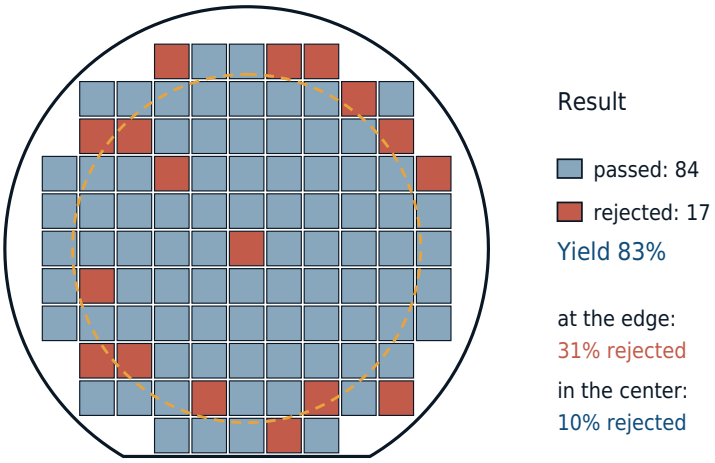
The result of each individual test is stored with its exact position in a so-called wafer map – a digital map that records, for every die on the wafer, whether it passed the test (pass) or not (fail). This map is referred to as a yield map and accompanies the wafer through all subsequent stages of production.

The yield map serves several important functions:

- During the subsequent dicing and assembly processes, only functional dies are further processed based on the map, while defective ones are discarded
- Spatial patterns in the yield map (such as clustered failures near the wafer edge) provide clues to systematic process issues
- For memory devices (DRAM, NAND flash), the map enables the targeted use of redundant circuit blocks: if a die contains individual defective memory cells, these can be electrically replaced by built-in spare circuits (redundancy), allowing the die to remain usable overall

Yield Map of a Wafer

Every die carries its test result – failures accumulate at the edge



The map accompanies the wafer through further processing: during dicing, only the passing dies are processed further. Spatial patterns like this edge loss point to systematic process problems.

The average yield of a wafer – that is, the proportion of functional dies relative to the total number – is one of the most important metrics in semiconductor manufacturing, as it directly determines the production cost per chip.

Final Test and Binning

The Final Functional Test

Once the die has received its finished package through assembly and encapsulation, it undergoes one last, comprehensive functional test – the final test. Unlike the wafer test, which typically offers only limited test coverage due to a restricted number of needles and contact quality, the finished, packaged chip can be tested far more comprehensively and under more realistic electrical conditions through its regular pins.

Tests performed include:

- Complete functional verification of all logic and analog circuit blocks
- Electrical parameters such as operating voltage, current consumption and signal timing
- Behavior under varying temperature and voltage conditions

- For memory devices: a complete cell test across the entire address range

Only after passing the final test is a chip considered ready for sale. Defective components are sorted out and typically scrapped, since repair at this stage of production is no longer economically viable.

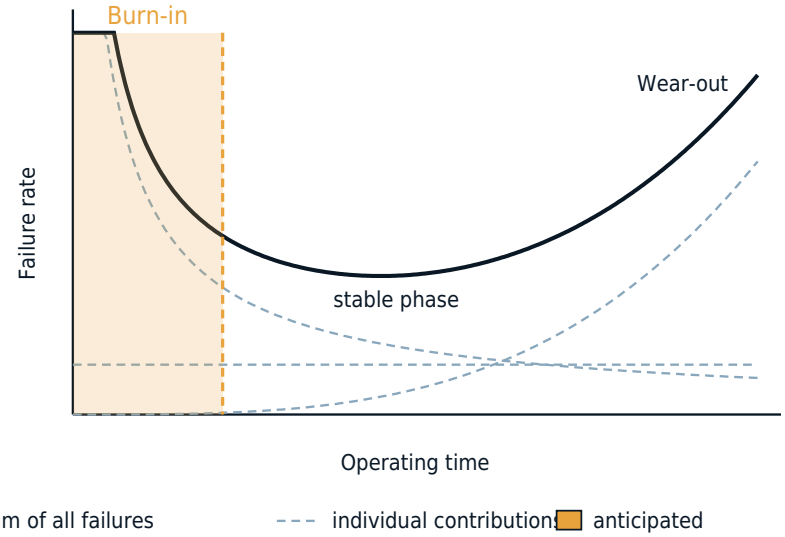
Burn-in

For applications with particularly strict reliability requirements (such as automotive or aerospace electronics), a so-called burn-in process is often performed before the final test itself. During burn-in, chips are stressed for a defined period – ranging from a few hours to several days – under elevated temperature and, in some cases, elevated operating voltage.

The rationale lies in the so-called bathtub curve of electronic component failure rates: an unusually high proportion of early failures typically occurs within the first operating hours of a component, most often caused by hidden manufacturing defects such as material flaws or contamination that the final test alone would not catch. After this early-failure phase, the failure rate drops to a low, stable level before rising again toward the end of the component lifetime due to wear-related effects.

Bathtub Curve of the Failure Rate

Burn-in anticipates the early failures



Hidden manufacturing defects show up in the first hours of operation. Under elevated temperature and voltage, these components already fail at the manufacturer – only what survives the steep initial region gets shipped.

The burn-in process anticipates these early failures: components with hidden

defects fail during burn-in itself and are sorted out before ever reaching a customer. The surviving portion of chips consequently exhibits significantly higher reliability during subsequent field use.

Binning

Even within a single wafer, and indeed within the very same die design, unavoidable process variations lead to slight differences in electrical characteristics – for instance in the maximum achievable clock frequency of a processor or in power consumption. Rather than discarding chips with slightly inferior characteristics, they are instead sorted according to their actual performance and marketed as different product variants. This practice is known as binning.

A well-known example is processor families in which technically identical chips are sold under different model names and at different prices depending on the maximum clock frequency they achieve. Chips that fail to meet the highest requirements are placed into a lower performance segment (bin) rather than being discarded as scrap.

Binning allows manufacturers to economically capture the natural variation inherent in manufacturing, rather than treating it purely as yield loss. Following binning, chips receive their final marking – typically via laser engraving of the package – and are packaged for shipment to device manufacturers.