Cell Libraries

Although the transistors from one or more basic cells could be connected together to implement practically any digital function, design engineers using gate arrays do not work at the transistor level. Instead, each ASIC vendor selects a set of logic functions such as primitive gates, multiplexers, and registers that they wish to make available to the engineers. The vendor also determines how each of these functions can be implemented using the transistors from one or more basic cells.

A primitive logic gate may only require the transistors from a single basic cell, while a more complex function like a D-type flip-flop or a multiplexer may require the use of several basic cells. Each of these "building block" functions is referred to as a *cell*—not to be confused with a *basic cell*—and the set of functions provided by the ASIC vendor are known collectively as the *cell library*. The number of functions in a typical cell library can range from 50 to 250 or more. The ASIC vendor also provides the design engineers with a cell library data book containing details about each function, including its truth table and timing attributes.

A High-level View of the Design Flow

Aided by a suite of design tools (see also the main *ASIC Design Flow* discussions later in this chapter), design engineers select the cells and the connections between them that are necessary for the gate array to perform its desired function. In the classical ASIC design flow, this information is passed to the ASIC vendor in a computer-readable form, and the vendor uses another suite of tools to complete the device. One of these tools—called the *placer*—assigns the cells selected by the engineers to basic cells on the silicon (Figure 17-5).

Another tool—called the *router*—determines the optimal way to connect the cells together (these tools operate hand-in-hand and are referred to collectively as *place-and-route*). Further tools are used to create the masks required to implement the final metalization layers. These layers are used to connect the transistors in the basic cells to form logic gates, and to connect the logic gates to form the complete device. The channels in channeled devices are used for the tracks that connect the logic gates together. By comparison, in channel-less devices, the connections between logic gates have to be deposited over the top of other basic cells. In the case of early processes based on two layers of

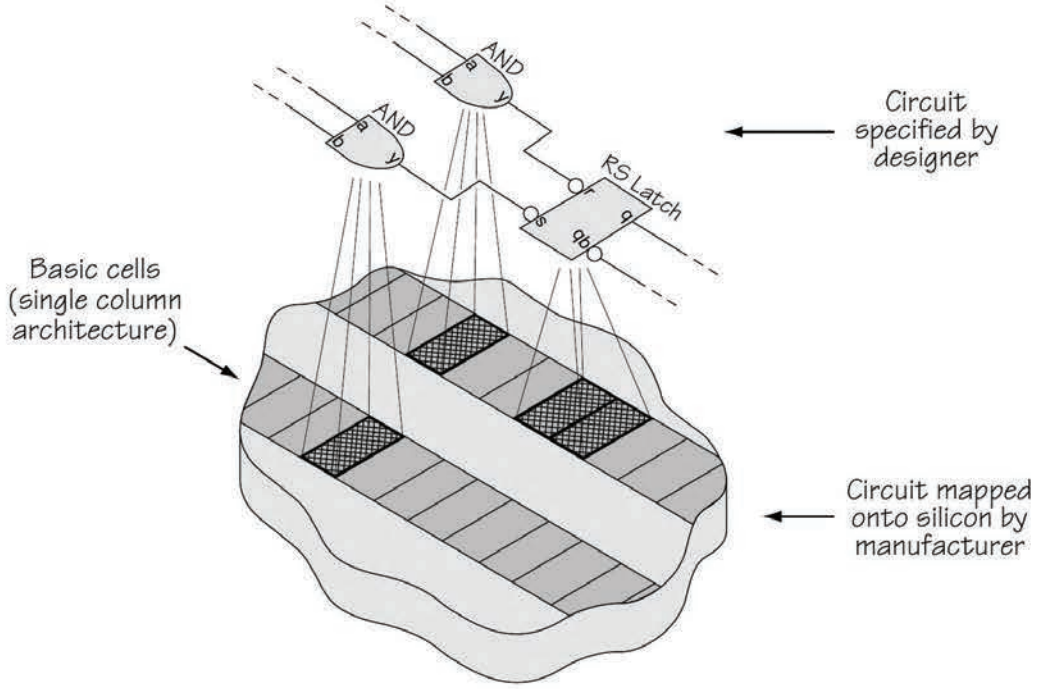


Figure 17-5. Assigning cells to basic cells

metalization, any basic cell overlaid by a track was no longer available to the user. More recent processes, which can potentially have up to eight metalization layers, overcome this problem.

Gate arrays can contain from thousands to millions of equivalent gates,[3] but there is a difference between the number of *available gates* and the number of *usable gates*. Due to limited space in the routing channels of channeled devices, the design engineers may find that they can only actually use between 70% and 90% of the total number of available gates. Sea-of-gates architectures provide significantly more available gates than those of channeled devices because they don't contain any dedicated wiring channels. However, in practice, only about 40% of the gates were usable in devices with two layers of metalization, rising to 60% or 70% in devices with three or four metalization layers, and even higher as the number of metalization layers increases.

Functions in gate array cell libraries are generally fairly simple, ranging from primitive logic gates to the level of registers. The ASIC vendor may also provide libraries of more complex logical elements called *hard-macros* (or *macro-cells*)

[3] The concept of *equivalent gates* was introduced in Chapter 14.

and *soft-macros* (or *macro-functions*).[4] In the case of gate arrays, the functions represented by the hard-macros and soft-macros are usually at the MSI and LSI level of complexity, such as comparators, shift-registers, and counters. Hard-macros and soft-macros are both constructed using cells from the cell library. In the case of a hard-macro, the ASIC vendor predetermines how the cells forming the macro will be assigned to the basic cells and how the connections between the basic cells will be realized. By comparison, in the case of a soft-macro, the assignment of cells to basic cells is performed at the same time, and by the same tool, as for the simple cells specified by the design engineers.

Gate arrays are classed as *semi-custom* devices. The definition of a semi-custom device is that it has one or more customizable mask layers, but not all the layers are customizable. Additionally, the design engineers can only utilize the predefined logical functions provided by the ASIC vendor in the cell and macro libraries.

Standard Cell Devices

Standard cell devices bear many similarities to gate arrays. Once again, each ASIC vendor decides which logic functions they will make available to the design engineers. Some vendors supply both gate array and standard cell devices, in which case the majority of the logic functions in the cell libraries will be identical and the main differences will be in their timing attributes.

Standard cell vendors also supply hard-macro and soft-macro libraries, which include elements at the LSI and VLSI level of complexity, such as processors, controllers, and communication functions. Additionally, these macro libraries typically include a selection of RAM and ROM functions, which cannot be implemented efficiently in gate array devices.[5] Last but not least, the design engineers may decide to reuse previously designed functions and/or to purchase blocks of *intellectual property (IP)*[6] (see sidebar).

---

[4] The term *macro* was inherited from the software guys. In software terms, a macro is like a subroutine. So a single computer instruction (to call the macro) initiates a series of additional instructions for the computer to perform. Similarly, in hardware terms, a single macro represents a large number of primitive logic gates.

[5] In the case of a design that contains large functions connected together by relatively small amounts of simple logic, this interfacing logic is referred to as "glue logic."

[6] In conversation, IP is pronounced by spelling it out as "I-P."

A High-level View of the Design Flow

Once again, aided by a suite of design tools, the design engineers determine which elements they wish to use from the cell and macro libraries, and how they require these elements to be connected together (see also the main *ASIC Design Flow* discussions later in this chapter). As before, in the classical ASIC design flow, the design engineers pass this information to the ASIC vendor in a computer-readable form, and the vendor uses another suite of tools to complete the device.

Unlike gate arrays, however, standard cell devices do not use the concept of a basic cell, and no components are prefabricated on the silicon chip. The ASIC vendor creates custom masks for every stage of the device's fabrication. This allows each logic function to be created using the minimum number of transistors required to implement that function with no redundant components. Additionally, the cells and macro functions can be located anywhere on the chip, there are no dedicated interconnection areas, and the functions are placed so as to facilitate any connections made between them. Standard cell devices therefore provide a closer-to-optimal utilization of the silicon than do gate arrays (Figure 17-6).

There are several different flavors of intellectual property (IP). In the generic sense, if you have a capriciously cunning idea, then that idea is your intellectual property and no one else can use it unless they pay you for it.

When a team of electronics engineers is tasked with designing a complex integrated circuit, rather than "reinvent the wheel," they may decide to purchase the plans for one or more functional blocks that have already been created by someone else. The plans for these functional blocks are known as *intellectual property (IP)*.

IP blocks can range all the way up to sophisticated communications functions and microprocessors. The more complex functions—like microprocessors—may be referred to as "cores."

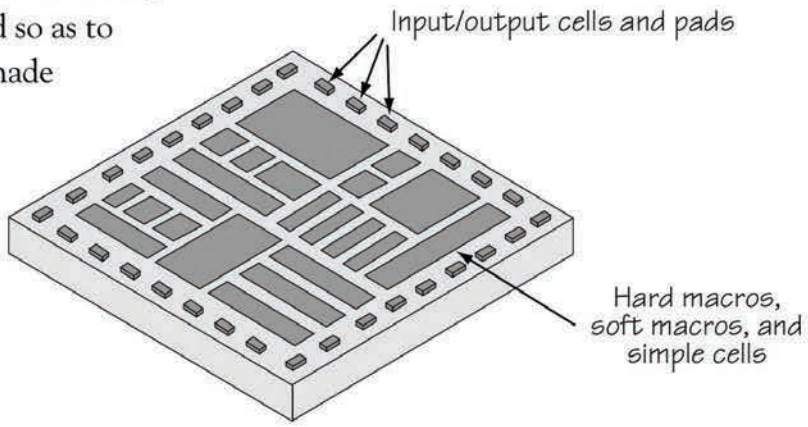


Figure 17-6. Standard cell architecture

Compiled Cell Technology

Design engineers working with standard cell devices may also make use of *compiled cell technology*. In this case, the engineers specify a desired function in such terms as Boolean equations or state tables, and then employ a tool called a *silicon compiler* to generate the masks required to implement the transistors and interconnections. These tools can generate highly efficient structures, because they have the ability to vary the size of individual transistors to achieve the optimum tradeoff between speed requirements and power consumption. Silicon compilers are also useful in the creation of selected macro functions such as highly structured RAMs and ROMs. The user simply specifies the required widths and depths of such functions, leaving the silicon compiler to determine an optimal implementation.

Compiled cell technology was originally intended to provide a mechanism for the creation of entire devices. This promise was never fully realized, and the technology is now considered to be an adjunct to standard cell devices, allowing the design engineers to create custom cells and macros to augment the libraries supplied by the ASIC vendor.

From the ASIC vendor's point of view, standard cell devices may be considered to be *full custom* on the basis that all the mask layers are customized. From the design engineer's point of view, however, these devices are still considered to be *semi-custom* because, with the exception of compiled cells, the engineers can only utilize the predefined functions provided by the ASIC vendor in the cell and macro libraries.

Full Custom Devices

In the case of *full custom* devices, the design engineers have complete control over every mask layer used to fabricate the silicon chip. The ASIC vendor does not prefabricate any components on the silicon and does not provide any cell libraries. With the aid of appropriate design tools, engineers can "handcraft" the dimensions of individual transistors, and then create higher-level functions based on these components. For example, if the engineers require a slightly faster logic gate, they can alter the dimensions of the transistors used to build that gate. The design tools used for full custom devices are often created in-house.

Full custom devices may employ compiled cell technology and may also include analog circuitry such as comparators, amplifiers, filters, and digital-to-analog and analog-to-digital converters. The design of full custom devices is highly complex and time consuming, but the resulting chips contain the maximum amount of logic with minimal waste of silicon real estate.

Input/Output Cells and Pads

Around the periphery of the silicon chip are power and signal pads used to interface the device to the outside world. The cell library data books for gate array and standard cell devices include a set of functions known as *input/output (I/O) cells*. The signal pads contain a selection of transistors, resistors, and diodes necessary to implement *input*, *output*, or *bi-directional* buffers, and the design engineers can decide how each pad will be configured. The masks and metalization used to interconnect the internal logic are also used to configure the components in the input/output cells and to connect the internal logic to these cells (Figure 17-7).

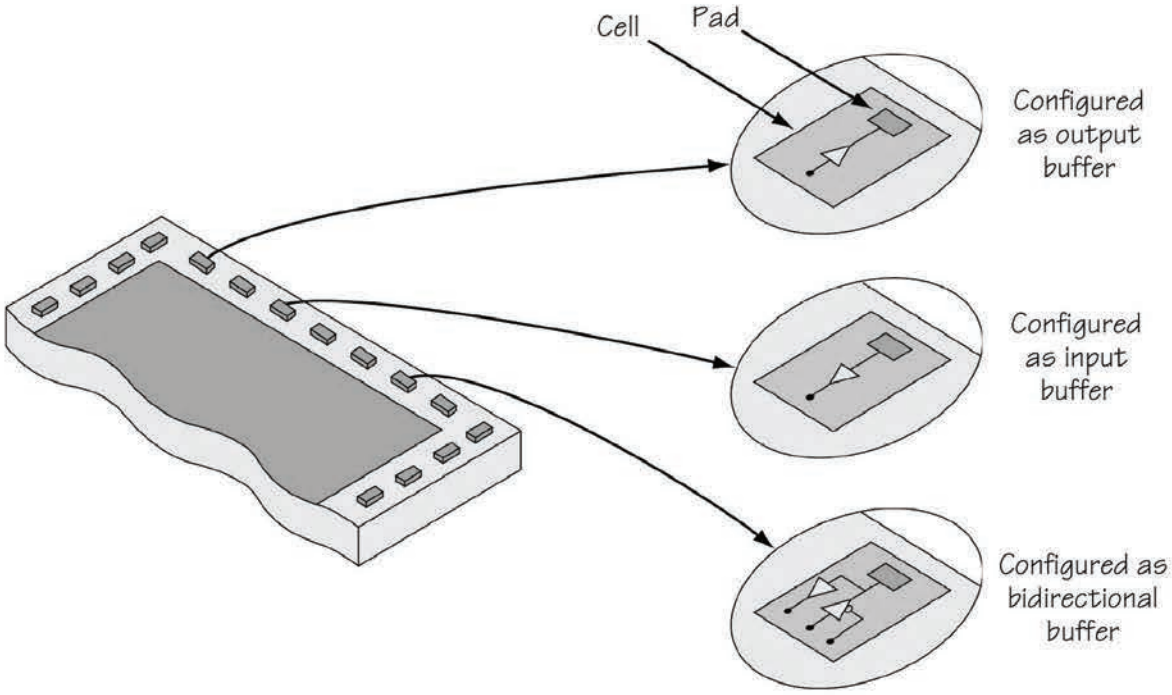


Figure 17-7. Input/output (I/O) cells and pads

The ASIC vendor may permit the design engineers to individually specify whether each input/output cell should present CMOS, TTL, or ECL characteristics to the outside world. The input/output cells also contain any circuitry required to provide protection against *electrostatic discharge* (ESD). After the metalization layers have been added, the chips are separated and packaged using the same techniques as for standard integrated circuits.

Who Are All the Players?

In a little while we're going to look at the ASIC design flow, but before that we need to have some understanding as to who all of the players are in this complicated game. As you will soon come to realize, understanding the technology is the easy part—it's when you try to work out who does what that things start to get hairy.

First of all we have the folks who create the tools (software programs) that the engineers use to design integrated circuits, circuit boards, and electronic systems. Prior to the 1970s, electronic circuits were handcrafted. Circuit diagrams (known as *schematics*) showing symbols for the components to be used and the connections between them were drawn using pen, paper, and stencils. Similarly, the copper tracks on a circuit board were drawn using red and blue pencils to represent the top and bottom of the board. Any form of analysis (for example, *"What frequency will this oscillator run at if I use this capacitor and this resistor?"*) was performed with pencil, paper, and a slide rule (or a mechanical calculator if you were lucky). Not surprisingly, this style of design was time-consuming, expensive, and prone to error.

Computer-aided Design (CAD)

As electronic designs and devices grew more complex, it became necessary to develop automated techniques to aid in the design process. In the early 1970s, companies like Calma, ComputerVision, and Applicon created special computer programs that helped personnel in the drafting department[7] capture hand-drawn designs in digital form using large-scale digitizing tables.

Over time, these early computer-aided drafting tools evolved into interactive programs that performed integrated circuit layout (that is, they could be

[7] The drafting department is referred to as the "drawing office" in the UK.

used to describe the locations of the transistors forming the integrated circuit and the connections between them). Other companies like Racal-Redac, SCI-Cards, and Telesis created equivalent layout programs for printed circuit boards. These integrated circuit and circuit board layout programs became known as *computer-aided design (CAD)* tools.[8,9]

Computer-aided Engineering (CAE)

Also in the late 1960s and early 1970s, a number of universities and commercial companies started to develop computer programs known as *simulators*. These programs allowed students and engineers to emulate the operation of an electronic circuit without actually having to build it first. Perhaps the most famous of the early simulators was the *simulation program with integrated circuit emphasis (SPICE)*.[10] This was developed by the University of California in Berkeley and was made available for widespread use around the beginning of the 1970s. SPICE was designed to simulate the behavior of analog circuits—other programs called logic simulators were developed to simulate the behavior of digital circuits.

Around the beginning of the 1980s, companies like Daisy, Mentor, and Valid spawned computer programs that allowed engineers to capture schematic (circuit) diagrams on the computer screen. These tools could then be used to generate textual representations of the circuits called *netlists* that described the components to be used and the connections between them. In turn, these netlists could be used to drive analog and digital simulators (and eventually layout tools).

The companies promoting front-end tools for schematic capture and simulation classed them as *computer-aided engineering (CAE)*.[11] This was based on the fact that these tools were targeted toward design engineers, and the CAE companies wished to distinguish their products from the CAD tools that were originally used by the drafting department.

---

[8] In conversation, CAD is pronounced as a single word to rhyme with "bad."

[9] The term CAD is also used to refer to computer-aided design tools intended for a variety of other engineering disciplines, such as mechanical and architectural design.

[10] In conversation, SPICE is pronounced like the seasoning to rhyme with "mice."

[11] In conversation, CAE is spelled out as "C-A-E".

Designers versus Engineers

If you say things the wrong way when talking to someone in the industry, you immediately brand yourself as an outsider (one of "them" instead of one of "us"). For historical reasons that are based on the origins of the terms CAD and CAE, the term *layout designer* or simply *designer* is typically used to refer to someone who lays out a circuit board or integrated circuit (determines the locations of the components and the routes of the tracks connecting them together). By comparison, the term *design engineer* or simply *engineer* is typically used to refer to someone who conceives and describes the functionality of an integrated circuit, printed circuit board, or electronic system (what it does and how it does it).

Electronic Design Automation (EDA)

Sometime during the 1980s, all of the CAE and CAD tools used to help design electronic components and systems came to be referred to by the "umbrella" name of *electronic design automation (EDA)*,[12] and everyone was happy (apart from the ones who weren't, but they don't count).

Some of the Key Players

This is where things start to get interesting. Let's start with the *system house*, A block in the middle of Figure 17-8. These are the folks who design and build the system-level products (from cell phones to televisions to computers) that eventually wend their way into our hands.

When these system house folks are creating a new product, they may decide it requires one or more ASICs, which they also design. Once they've created the design for an ASIC, they may pass it over to an ASIC vendor to be implemented and fabricated (see also the *ASIC Design Flow* discussions below). In this case, the ASIC vendor is in charge of creating the masks and constructing and packaging the final devices.

As opposed to a full-line ASIC vendor, a *fabless semiconductor company* is one that designs and sells integrated circuits, but does not have the ability to manufacture them. By comparison, a *foundry* is a company that manufactures integrated circuits, but doesn't actually do any designs of their own (these are also known as "fabs" because they fabricate the integrated circuits).

---

[12] In conversation, EDA is spelled out as "E-D-A".

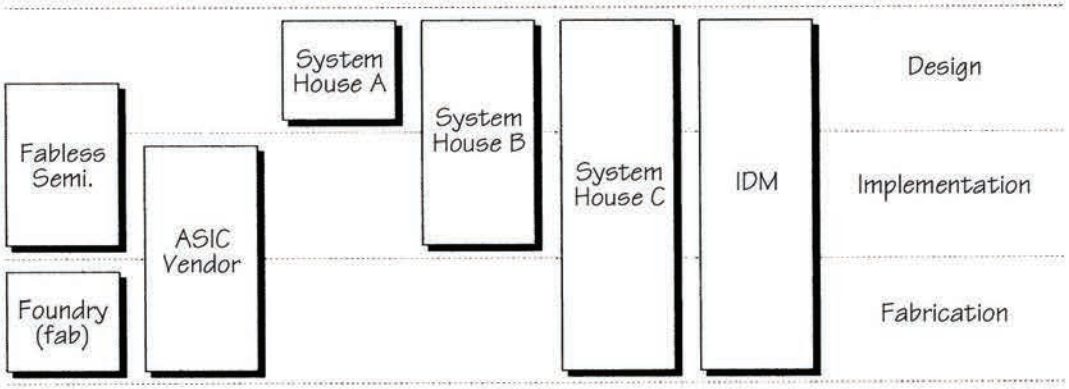


Figure 17-8. Some of the key players

Initially we described a system house as passing its ASIC designs over to an ASIC vendor for implementation and fabrication. However, some system houses (B) perform both the design and implementation, and then hand the masks over to a foundry for fabrication. Alternatively, a system house (C) may have its own fab capability, in which case it will perform the entire process—design, implementation, and fabrication—"in-house."

And just to confuse the issue even further, we have *integrated device manufacturers (IDMs)*.[13] These are companies that are very similar to system houses, except that IDMs focus on designing, manufacturing, and selling integrated circuits as opposed to complete electronic systems.

The ASIC Design Flow

Before we start this portion of our discussions, we should note that some integrated circuits contain only digital functions, others are analog in nature, and some contain both analog and digital elements (these latter components are referred to as *mixed-signal* devices). For the purpose of these discussions, we'll concentrate on a digital ASIC design flow. The full-up design flow is quite complex and is beyond the scope of this book, so we'll only consider a much-simplified version of the flow here (Figure 17-9).[14]

[13] In conversation, IDM is spelled out as "I-D-M".

[14] The book *EDA: Where Electronics Begins* (www.techbites.com/eda), which was co-authored by the author of this book, introduces design flows and tools (*capture*, *simulation*, *synthesis*, *verification*, *layout*, etc.), for integrated circuits, printed circuit boards, and electronic systems. Don't worry; *EDA: Where Electronics Begins* is presented at a nontechnical level that won't make your brains leak out of your ears.

Verilog started life as a commercial (proprietary) simulation language. However, in the mid 1990s it became an open standard under the auspices of the Open Verilog International (OVI) organization.

Similarly, VHDL, which came out of the American Department of Defense (DoD), has also evolved into an open standard. VHDL is an acronym for *VHSIC HDL* (where VHSIC is itself an acronym for *very high speed integrated circuit*).

VHDL was originally intended to provide a way of documenting designs in such a way that the design intent could be preserved for a long time. This was important for military projects with life spans in terms of decades. However, when VHDL was first conceived, little thought was given to using it to drive synthesis and simulation tools (which may explain why doing so can be such a painful experience).

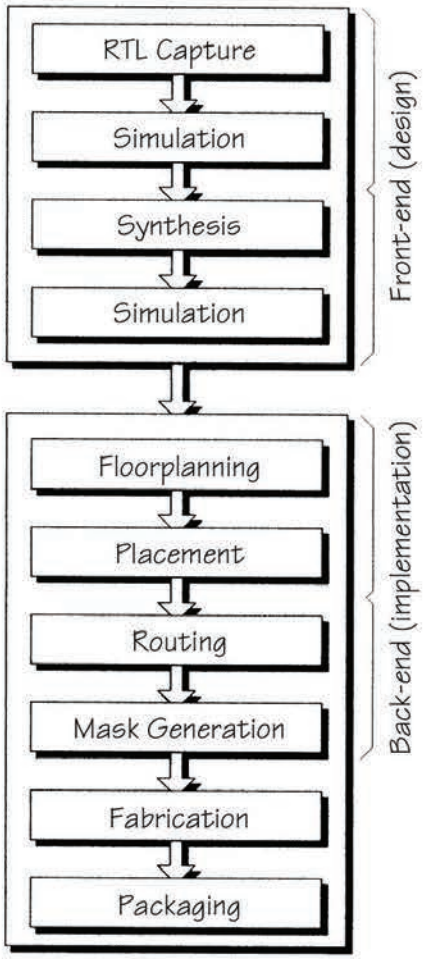


Figure 17-9. A simplified version of the ASIC design flow

Capturing the Design

First of all, the design engineers have to decide exactly what logic gates and functions they wish to use in their integrated circuit and how they are to be connected together. In the early days of integrated circuit design, schematic (circuit) diagrams were hand-drawn by the engineers using pencil and paper. Over time, schematic capture programs were developed that allowed engineers to place symbols representing components on a computer screen and then draw the wires connecting the symbols together.

However, today's digital integrated circuits can contain tens of millions of logic gates, and it simply isn't possible to capture and manage designs of this

complexity at the schematic level. Thus, as opposed to using schematics, the functionality of a high-end integrated circuit is now captured using a textual *hardware description language (HDL)*.[15] The two most popular HDLs for digital designs are Verilog and VHDL.[16,17]

Initially, the design engineers describe the device's functionality at a high level of abstraction called *register transfer level (RTL)*. For example, assuming that we had already declared an input signal called clock and three 8-bit registers called a[7:0], b[7:0], and y[7:0], an RTL statement might look something like the following:

```
WHEN clock RISES
    y[7:0] = a[7:0] & b[7:0];
END;
```

When capturing the design's functionality, the design engineers may decide to reuse RTL descriptions of functional blocks from previous designs. Also, they may decide to purchase *intellectual property (IP)* blocks from a third-party design house.

Simulation and Synthesis

Once the RTL has been captured, the engineers use a computer program called a *simulator*, which reads in the RTL and creates a virtual representation of the integrated circuit in the computer's memory. The engineers use the simulator to apply a sequence of test patterns to the simulation model's virtual inputs and check the responses on its virtual outputs to ensure that the device will function as planned (this sequence of patterns is called a *testbench*). The simulator models how signals propagate through—and are processed by—the high-level logic functions forming the integrated circuit.

Next, a *logic synthesis* program is used to automatically convert the high-level RTL representation into equivalent Boolean Equations (like the ones we discussed in Chapter 9). The synthesis tool automatically performs simplifications and minimizations, and eventually outputs a gate-level netlist (Figure 17-10).

---

15 In conversation, HDL pronounced by spelling it out as "H-D-L."

16 In conversation, VHDL pronounced by spelling it out as "V-H-D-L."

17 In the 1980s, engineers in Japan came up with their own HDL called UDL/I. This was a very useful language that was designed with synthesis and simulation in mind. However, Verilog and VHDL already held the high ground, and interest in UDL/I is now almost nonexistent even in Japan.

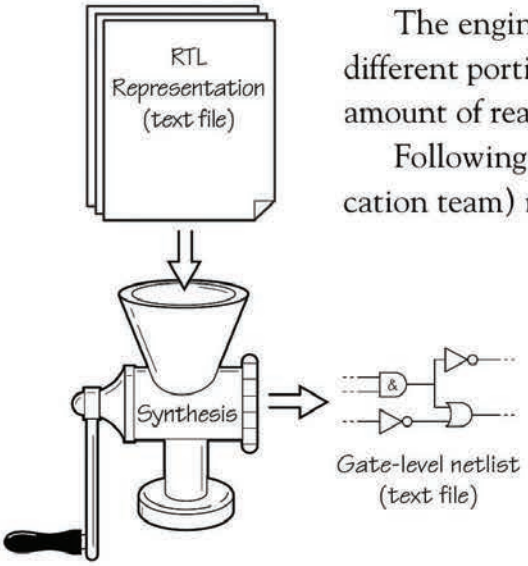


Figure 17-10. A synthesis program converts an RTL description into a gate-level netlist

The engineers can direct the synthesis tool to optimize different portions of designs for area (to use the smallest amount of real-estate on the silicon) or for speed.

Following synthesis, the engineers (or a separate verification team) may re-simulate the design at the gate level using the same testbench as before. This simulation is performed to ensure that the synthesis tool didn't inadvertently change the functionality of the design. As opposed to simulation—or in addition to it—the engineers may also run a *formal verification* program that compares the RTL and gate-level views of the design to ensure that they are functionally equivalent (synthesis tools have been known to make mistakes).

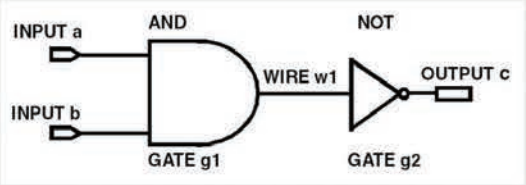
Floorplanning

Traditionally, once the design engineers had generated a gate-level netlist, they handed it over to the layout designers to perform *floorplanning*, *placement*, and *routing*. This may still be the case for some design flows. In the case of the majority of today's high-performance digital integrated circuits, however, the design engineers typically perform these functions also. For the purposes of these discussions, we'll assume that the design engineers undertake all of the tasks associated with designing the integrated circuit.

Floorplanning[18] involves creating a high-level representation showing where the major functional blocks will be placed on the surface of the silicon chip. Depending on the particular design flow, the floorplan may be created by hand, interactively, or automatically. The creator of the floorplan tries to ensure that blocks that communicate with each other are positioned as close to each other as possible. Also, it may be preferable to place certain blocks near to the periphery of the chip, while other blocks may be better situated toward the center.

[18] The term "floorplan" comes from the architectural and construction industries, because the graphical floorplan for an integrated circuit looks similar to a floorplan showing the rooms and corridors in a building.

A gate-level netlist is a textual representation that describes the logic gates to be used and the connections between them. For example, consider a very simple circuit called "test" comprising just two logic gates:



A simple textual netlist representation of this circuit could be as follows:

```
START CIRCUIT NAME=test;
INPUTS = a, b;
OUTPUTS = c;
INTERNAL WIRES = w1;
GATE=g1, TYPE=AND, INPUT=a, INPUT=b, OUTPUT=w1;
GATE=g2, TYPE=NOT, INPUT=w1, OUTPUT=c;
END CIRCUIT NAME=test;
```

Such a netlist can be used to drive simulation and layout tools. There are a variety of different netlist formats, such as the industry-standard *electronic design interchange format (EDIF)*. However, integrated circuit designers often work with netlists in *VHDL* or *Verilog*.

Placement and Routing

Next, the design engineer uses a placement program to determine the optimum locations for the individual logic gates forming the major functions. This program tries to ensure that gates that communicate with each other are placed as close to each other as possible. Also, the program may decide that it's preferable to place certain gates near to the periphery of a block, while other gates may be better situated toward its center.

Once all of the logic gates have been placed, the engineer uses a routing program to determine the optimum paths for the tracks linking the logic gates together. Somewhere in the flow, the engineer will also use special programs to design and route the chip's clock signals and power supplies.

Mask Generation

Throughout the entire process, the engineers will use a variety of complex verification and timing analysis programs to ensure that the device will work as planned. Eventually, when everyone is happy, the computer will take the placed and routed design and generate all of the masks that will be required to actually fabricate the device. These masks, which are represented in a machine-readable format known as GDSII,[19,20] will be used by the foundry to fabricate the device.

ASIC, ASSP, and COT

This is where things start to get really "fluffy" around the edges when you're trying to explain them to someone, but we'll take a stab at trying to put everything into some sort of context. Throughout this chapter we've referred to ASICs as being all-encompassing. In reality, an ASIC is generally considered to be a component that is designed by and/or used by a single company in a specific system (for example, a telecommunications system house may design an ASIC for use in their new cell phone). By comparison, an *application-specific standard part (ASSP)*[21] is a more general-purpose device targeted at a specific market (say the communications market) and intended for use by multiple design houses.

It's important to note that there are a number of points at which the team designing the integrated circuit may decide to hand over control to the ASIC vendor or foundry. In some (rare) cases, the design team may create only the high-level RTL description and then pass this to the ASIC vendor to take the design through synthesis, floorplanning, and layout (place-and-route) to a physical implementation. In this case, the ASIC vendor will charge a lot of money for its services.

A traditional ASIC design flow has the design engineers performing RTL capture, synthesis, and simulation (this may be referred to as the *front-end*

---

19 In conversation, GDSII is pronounced by spelling it out as "G-D-S-2."

20 Officially, GDSII doesn't actually stand for anything in particular. Unofficially, it's rumored to be an acronym for *Graphics Design Station Two*, because GDSII was the output format from the second generation of design stations from Calma.

21 In conversation, ASSP is pronounced by spelling it out as "A-S-S-P."

portion of the flow). The design engineers then hand over the resulting gate-level netlist to the ASIC vendor's layout designers, who perform floorplanning and layout and progress the design to its final implementation (this may be referred to as the *back-end* portion of the flow).

In contrast, in the case of today's high-end ASIC and ASSP devices, the design engineers are typically in charge of capturing the RTL and performing simulation, synthesis, floorplanning, layout, etc. The resulting placed-and-routed design is then passed to the ASIC vendor, whose task is largely reduced to generating the masks used to create the chips and then manufacturing the chips themselves.

There is also the *customer-owned tooling (COT)*[22] business model, in which the design engineers may take the design all the way through to generating the masks used to create the chips. In this case, the only task remaining is for a foundry to actually fabricate the chips.

Non-recurring engineering (NRE)

One important concept that people in the electronics industry often talk about is that of *non-recurring engineering (NRE)*,[23] which refers to the costs associated with developing an ASIC or ASSP. The NRE depends on a number of factors, including the complexity of the design, the style of packaging, and who does what in the design flow (that is, how the various tasks are divided between the design house and the semiconductor vendor).

Summary

Electronic designs typically contain a selection of highly complex functions—such as processors, controllers, and memory—interfaced to each other by a plethora of simple functions such as primitive logic gates. The simple interfacing functions are often referred to as the *glue logic*. Both complex and simple functions are available as standard integrated circuits. Devices containing small numbers of simple logic functions (for example, four 2-input AND gates) are sometimes referred to as "jellybean" parts. When glue logic is implemented

[22] In conversation, COT is pronounced by spelling it out as "C-O-T" (people in the know do NOT pronounce COT to rhyme with "hot").

[23] In conversation, NRE is pronounced by spelling it out as "N-R-E."

using jellybean parts, these components may require a disproportionate amount of the circuit board's real estate. A few complex devices containing the vast majority of the design's logic gates may occupy a small area, while a large number of jellybean devices containing relatively few logic gates can occupy the bulk of the board's surface.

Early ASICs contained only a few hundred logic gates. These were mainly used to implement glue logic, and it was possible to replace fifty or more jellybean devices with a single ASIC. This greatly reduced the size of circuit boards, increased speed and reliability, and reduced power consumption. As time progressed, application-specific devices with tens of millions of gates and thousands of pins became available. Thus, today's ASICs can be used to implement the most complex functions.

Both gate array and standard cell devices essentially consist of building blocks designed and characterized by the manufacturer and connected together by the designer. Gate arrays are mask programmable with a predefined number of transistors, and different designs are essentially just changes in the interconnect. This means that gate arrays require the customization of fewer layers than standard cell devices. Gate arrays are therefore faster to implement than their standard cell equivalents, but the latter can contain significantly more logic gates.

In certain respects gate arrays are moving towards having similar capabilities to those of standard cells, and some support complex functions such as memory and processor cores. The design engineers generally know the memory and processing requirements in advance of the rest of the logic, and the gate array manufacturer may supply devices with pre-built processor and memory functions surrounded by arrays of basic cells. Some gate array devices support simple analog cells in addition to the digital cells, and standard cell devices can be constructed with complex analog functions if required.

Chapter 18

Circuit Boards

Electronic components are rarely useful in isolation, and it is usually necessary to connect a number of them together in order to achieve a desired effect. Early electronic circuits were constructed using discrete (individually packaged) components such as transistors, resistors, capacitors, and diodes. These were mounted on a non-conducting board and connected using individual pieces of insulated copper wire. The thankless task of wiring the boards by hand was time-consuming, boring, prone to errors, and expensive.

The First Circuit Boards

The great American inventor Thomas Alva Edison had some ideas about connecting electronic circuits together. In a note to Frank Sprague, founder of Sprague Electric, Edison outlined several concepts for printing additive traces on an insulating base. He even talked about the possibility of using conductive inks (it was many decades before this technology—which is introduced later in this chapter—came to fruition).

In 1903, Albert Hanson (a Berliner living in London) obtained a British patent for a number of processes for forming electrical conductors on an insulating base material. One of these described a technique for cutting or stamping traces out of copper foil and then sticking them to the base. Hanson also came up with the idea of double-sided boards and through-holes (which were selectively connected by wires).

In 1913, Arthur Berry filed a British patent for covering a substrate with a layer of copper and selectively etching parts of it away to leave tracks. In another British patent issued in 1925, Charles Ducas described etching, plating up, and even multi-layer circuit boards (including the means of interconnecting the layers). For the next few decades, however, it was easier and cheaper to wire boards manually. The real push into circuit boards only came with the invention of the transistor and later the integrated circuit.

PCBs and PWBs

By the 1950s, the interconnection technology now known as *printed wire boards (PWBs)* or *printed circuit boards (PCBs)* had gained commercial acceptance. Both terms are synonymous, but the former is more commonly used in America, while the latter is predominantly used in Europe and Asia. These circuit boards are often referred to as *laminates* because they are constructed from thin layers or sheets. In the case of the simpler boards, an insulating base layer has conducting tracks formed on one or both sides. The base layer may technically be referred to as the *substrate*, but this term is rarely used in the circuit board world.[1]

The original board material was *Bakelite*, but modern boards are predominantly made from woven glass fibers which are bonded together with an epoxy. The board is cured using a combination of temperature and pressure, which causes the glass fibers to melt and bond together, thereby giving the board strength and rigidity. These materials may also be referred to as *organic substrates*, because epoxies are based on carbon compounds as are living creatures. The most commonly used board material of this type is known as FR4, where the first two characters stand for flame retardant, and you can count the number of people who know what the "4" stands for on the fingers of one hand.

To provide a sense of scale, a fairly representative board might be 15 cm × 20 cm in area and in the region of 1.5 mm to 2.0 mm thick, but they can range in size from 2 cm × 2 cm or smaller (and thinner) to 50 cm × 50 cm or larger (and thicker).[2]

Subtractive Processes

In a subtractive process, a thin layer of copper foil in the order of 0.02 mm thick is bonded to the surface of the board. The copper's surface is coated with an *organic resist*, which is cured in an oven to form an impervious layer (Figure 18-1).

---

[1] See also the glossary definition of "substrate."

[2] These dimensions are not intended to imply that circuit boards must be square or even rectangular. In fact, circuit boards may be constructed with whatever outline is necessary to meet the requirements of a particular enclosure: for example, the shape of a car dashboard.

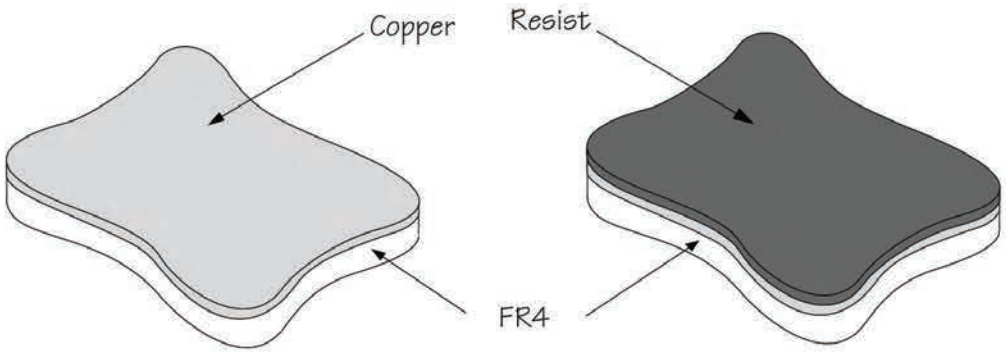


Figure 18-1. Subtractive process: applying resist to copper-clad board

Next, an optical mask is created with areas that are either transparent or opaque to ultraviolet light. The mask is usually the same size as the board, and can contain hundreds of thousands of fine lines and geometric shapes.

The mask is placed over the surface of the board, which is then exposed to ultraviolet (UV) light. This ionizing radiation passes through the transparent areas of the mask to break down the molecular structure of the resist. After the board has been exposed, it is bathed in an organic solvent, which dissolves the degraded resist. Thus, the pattern on the mask has been transferred to a corresponding pattern in the resist (Figure 18-2).

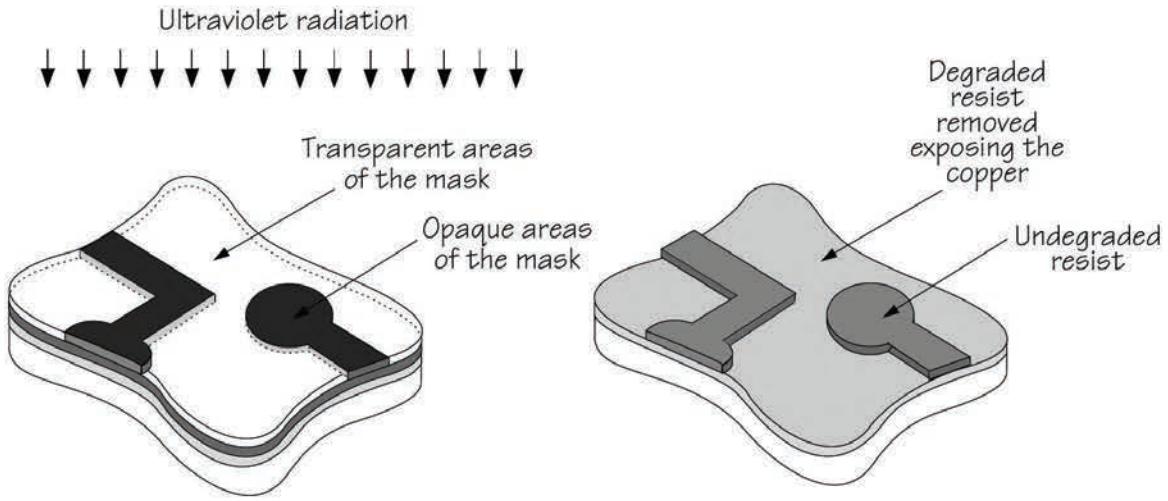


Figure 18-2. Subtractive process: degrading the resist and exposing the copper

A process in which ultraviolet light passing through the transparent areas of the mask causes the resist to be degraded is known as a *positive-resist* process; *negative-resist* processes are also available.[3] The following discussions assume positive-resist processes unless otherwise noted.

After the unwanted resist has been removed, the board is placed in a bath containing a cocktail based on sulfuric acid, which is agitated and aerated to make it more active. The sulfuric acid dissolves any exposed copper not protected by the resist in a process known as *etching*. The board is then washed to remove the remaining resist. Thus, the pattern in the mask has now been transferred to a corresponding pattern in the copper (Figure 18-3).

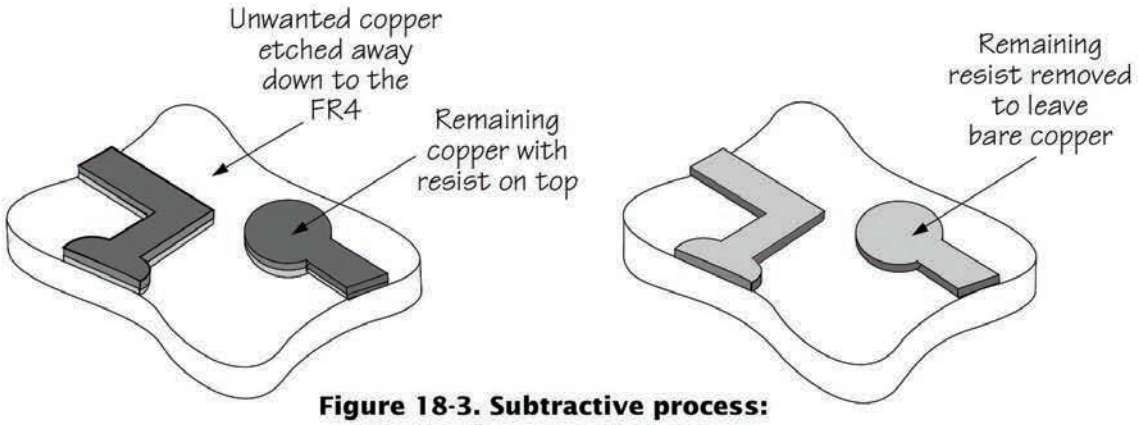


Figure 18-3. Subtractive process: removing the unwanted copper

This type of process is classed as *subtractive* because the board is first covered with the conductor and any unwanted material is then removed, or subtracted. As a point of interest, much of this core technology predates the modern electronics industry. The process of copper etching was well known by the printing industry in the 1800s, and opto-lithographic techniques involving organic resists were used to create printing plates as early as the 1920s. These existing processes were readily adopted by the fledgling electronics industry. As process technologies improved, it became possible to achieve ever-finer features. By 2002, a large proportion of boards were using *lines* and *spaces* of

[3] In a *negative-resist* process the ultraviolet radiation passing through the transparent areas of the mask is used to cure the resist. The remaining uncured areas are then removed using an appropriate solvent. Thus, a mask used in a negative-resist process is the photographic negative of one used in a positive-resist process to achieve the same effect; that is, the transparent areas are now opaque and vice versa.

5 mils or 4 mils, with some as small as 3 mils (one mil is one thousandth of an inch). In this context, the term "lines" refers to the widths of the tracks, while "spaces" refers to the gaps between adjacent tracks.

Additive Processes

An additive process does not involve any copper foil being bonded to the board. Instead, the coating of organic resist is applied directly to the board's surface (Figure 18-4).

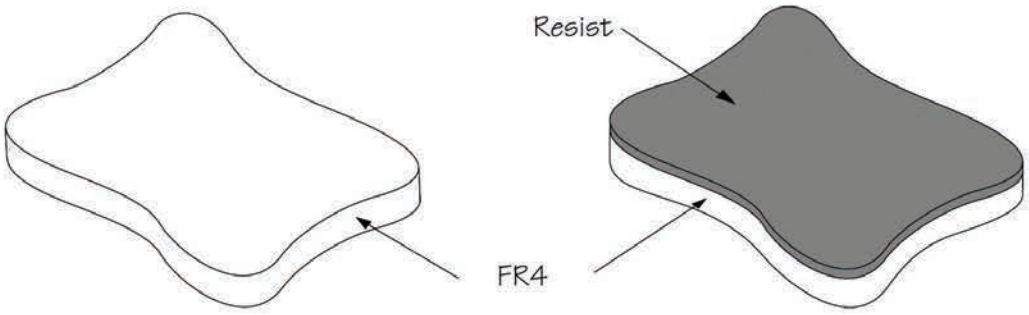


Figure 18-4. Additive process: applying resist to bare board

Once again the resist is cured, an optical mask is created, ultraviolet light is passed through the mask to break down the resist, and the board is bathed in an organic solvent to dissolve the degraded resist (Figure 18-5).[4]

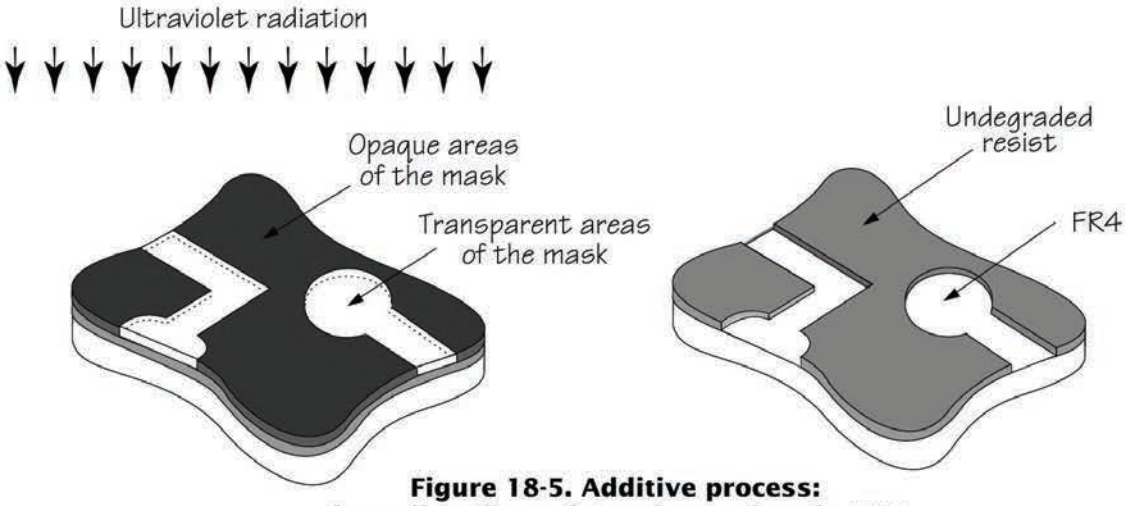


Figure 18-5. Additive process: degrading the resist and exposing the FR4

[4] Note that a mask used in an additive process is the photographic negative of one used in a subtractive process to achieve the same effect; that is, the transparent areas are now opaque and vice versa.

After the unwanted resist has been removed, the board is placed in a bath containing a cocktail based on copper sulfate where it undergoes a process known as *electroless plating*. Tiny crystals of copper grow on the exposed areas of the board to form copper tracks. The board is then washed in an appropriate solvent to remove the remaining resist (Figure 18-6).

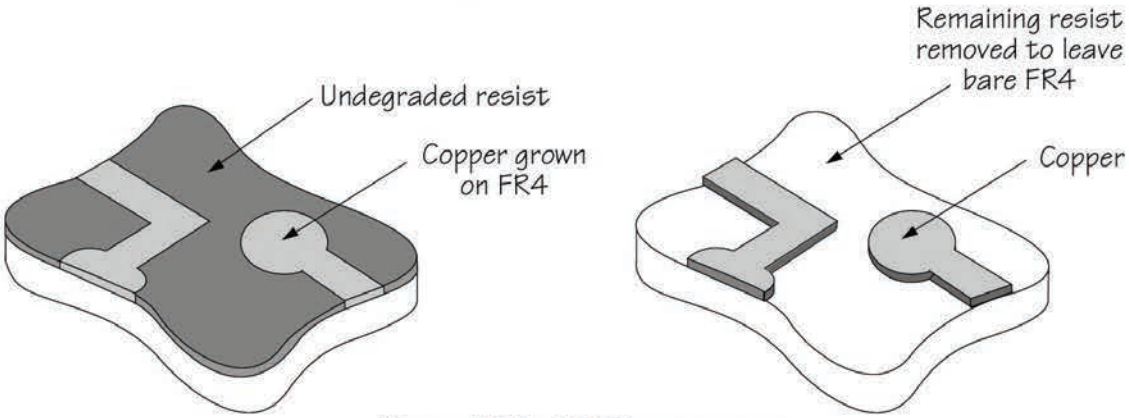


Figure 18-6. Additive process: adding the desired copper

A process of this type is classed as *additive* because the conducting material is only grown on, or added to, specific areas of the board. Additive processes are increasing in popularity because they require less processing and result in less wasted material. Additionally, fine tracks can be grown more accurately in additive processes than they can be etched in their subtractive counterparts. These processes are of particular interest for high-speed designs and microwave applications, in which conductor thicknesses and controlled impedances are critical. Groups of tracks, individual tracks, or portions of tracks can be built up to precise thicknesses by iterating the process multiple times with selective masking.

Single-sided Boards

It probably comes as no great surprise to find that *single-sided* boards have tracks on only one side. These tracks, which may be created using either subtractive or additive processes, are terminated with areas of copper known as *pads*. The shape of the pads and other features are, to some extent, dictated by the method used to attach components to the board. Initially, these discussions will assume that the components are to be attached using a technique known as

through-hole, which is described in more detail below. The pads associated with the through-hole technique are typically circular, and holes are drilled through both the pads and the board using a computer-controlled drilling machine (Figure 18-7).[5]

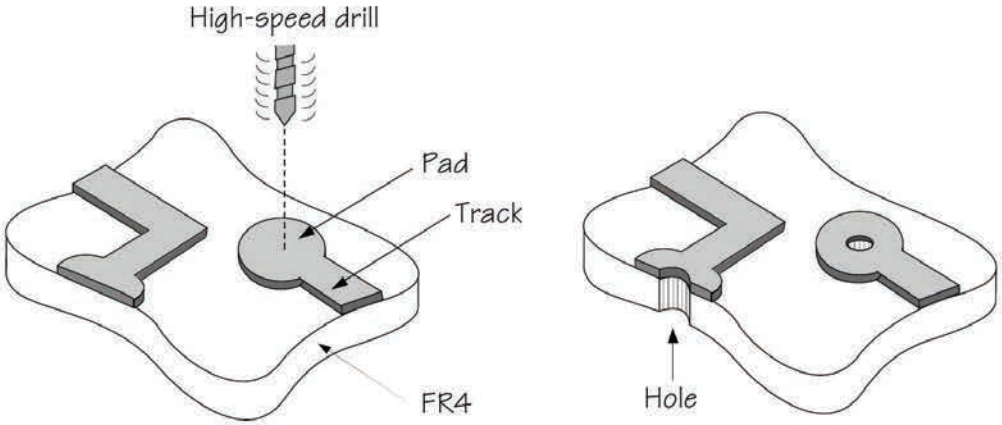


Figure 18-7. Single-sided boards: drilling the holes

Once the holes have been drilled, an electroless plating process referred to as *tinning* is used to coat the tracks and pads with a layer of *tin-lead* alloy. This alloy is used to prevent the copper from oxidizing and provides protection against contamination. The deposited alloy has a rough surface composed of vertical crystals called *spicules*, which, when viewed under a microscope, resemble a bed of nails. To prevent oxygen from reaching the copper through pinholes in the alloy, the board is placed in a *reflow oven* where it is heated by either *infrared (IR)* radiation or hot air. The reflow oven causes the alloy to melt and form a smooth surface (Figure 18-8).

After the board has cooled, a layer known as the *solder mask* is applied to the surface carrying the tracks (the purpose of this layer is discussed in the next section). One common technique is for the solder mask to be screen printed onto the board. In this case, the screen used to print the mask has patterns that leave the areas around the pads exposed, and the mask is then cured in an

[5] This is usually referred to as an *NC drilling machine*, where NC stands for "Numerically Controlled."

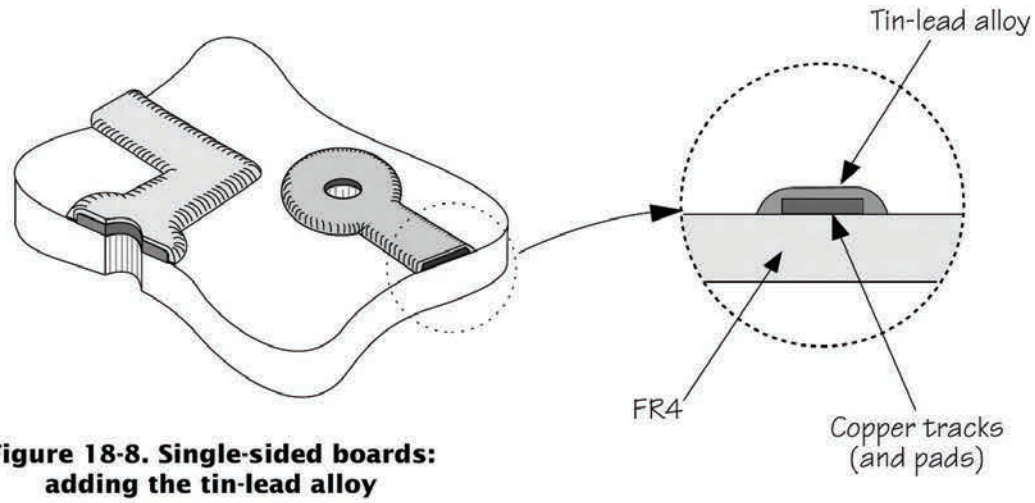


Figure 18-8. Single-sided boards: adding the tin-lead alloy

oven. In an alternative technique, the solder mask is applied across the entire surface of the board as a film with an adhesive on one side. In this case, a further opto-lithographic stage is used to cure the film with ultraviolet light. The optical mask used in this process contains opaque patterns, which prevent the areas around the pads from being cured; these areas are then removed using an appropriate solvent (Figure 18-9).

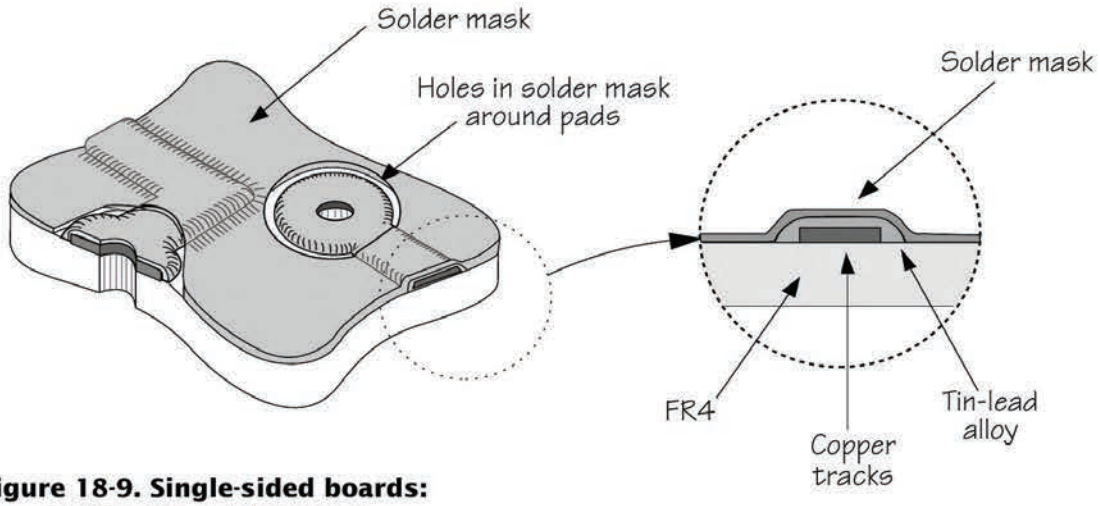


Figure 18-9. Single-sided boards: adding the solder mask

Beware! Although there are relatively few core processes used to manufacture circuit boards, there are almost endless variations and techniques. For example, the tracks and pads may be created before the holes are drilled or vice versa. Similarly, the tin-lead alloy may be applied before the solder mask or vice versa. This latter case, known as *solder mask over bare copper (SMOBC)*, prevents solder from leaking under the mask when the tin-lead alloy melts during the process of attaching components to the board. Thus, the tin-lead alloy is applied only to any areas of copper that are left exposed by the solder mask, such as the pads at the end of the tracks. As there are so many variations and techniques, these discussions can only hope to offer an overview of the main concepts.

Lead Through-Hole (LTH)

Prior to the early 1980s, the majority of integrated circuits were supplied in packages that were attached to a circuit board by inserting their leads through holes drilled in the board. This technique, which is still widely used, is known as *lead through-hole (LTH)*, *plated through-hole (PTH)*, or, more concisely, *through-hole*. In the case of a single-sided board, any components that are attached to the board in this fashion are mounted on the opposite side to the tracks. This means that any masks used to form the tracks are actually created as mirror-images of the required patterns (Figure 18-10).

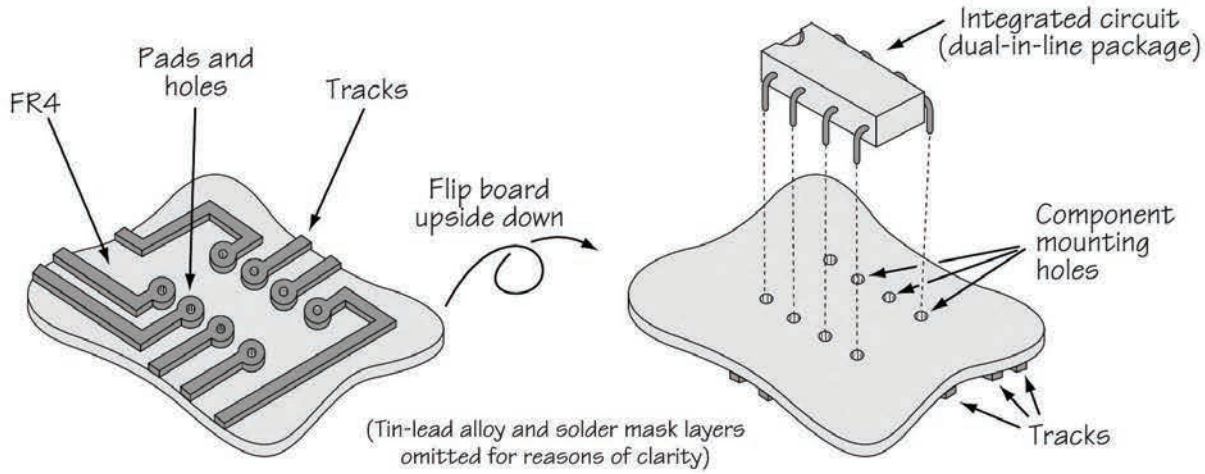


Figure 18-10. Lead through-hole (LTH)

The act of attaching components is known as *populating* the board, and the area of the board occupied by a component is known as its *footprint*. Early manufacturing processes required the boards to be populated by hand, but modern processes make use of automatic insertion machines. Similarly, component leads used to be hand-soldered to the pads, but most modern processes employ automatic *wave-soldering* machines (Figure 18-11).

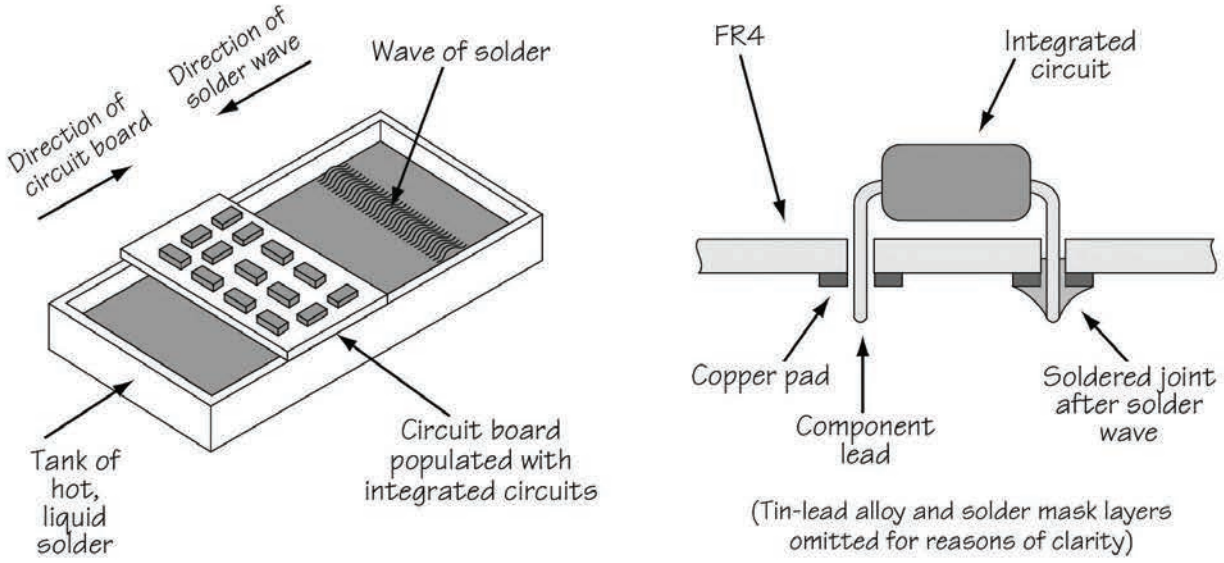


Figure 18-11. Wave-soldering

A wave-soldering machine is based on a tank containing hot, liquid *solder*.[6] The machine creates a wave (actually a large ripple) of solder which travels across the surface of the tank. Circuit boards populated using the through-hole technique are passed over the machine on a conveyer belt. The system is carefully controlled and synchronized such that the solder wave brushes across the bottom of the board only once.

The solder mask introduced in the previous section prevents the solder from sticking to anything except the exposed pads and component leads. Because the solder is restricted to the area of the pads, surface tension causes it to form

[6] An alloy of tin and lead with a relatively low melting point.

good joints between the pads and the component leads. Additionally, capillary action causes the solder to be drawn up the hole, thereby forming reliable, low-resistance connections. If the solder mask were omitted, the solder would run down the tracks away from the component leads. In addition to forming bad joints, the amount of heat absorbed by the tracks would cause them to separate from the board (this is not considered to be a good thing to happen).

Surface Mount Technology (SMT)

In the early 1980s, new techniques for packaging integrated circuits and populating boards began to appear. One of the more popular is *surface mount technology (SMT)*, in which component leads are attached directly to pads on the surface of the board. Components with packages and lead shapes suitable for this technology are known as *surface mount devices (SMDs)*. One example of a package that achieves a high lead count in a small area is the *quad flat pack (QFP)*, in which leads are present on all four sides of a thin square package.[7]

Boards populated with surface mount devices are fabricated in much the same way as their through-hole equivalents (except that the pads used for attaching the components are typically square or rectangular and do not have any holes drilled through them).[8] However, the processes begin to diverge after the solder mask and tin-lead plating layers have been applied. A layer of solder paste is screen-printed onto the pads, and the board is populated by an automatic *pick-and-place machine*, which pushes the component leads into the paste (Figure 18-12).

Thus, in the case of a single-sided board, the components are mounted on the same side as the tracks. When all of the components have been attached, the solder paste is melted to form good conducting bonds between their leads and the board's pads. The solder paste can be melted by placing the board in a *reflow oven* where it is heated by infrared (IR) radiation or hot air. Alternatively, the solder may be melted using *vapor-phase soldering*, in which the board is lowered into the vapor-cloud above a tank containing boiling hydrocarbons.

---

[7] Other packaging styles such as *pad grid arrays (PGAs)* and *ball grid arrays (BGAs)*—which are also suitable for surface mount technology—are introduced in more detail in Chapter 20.

[8] Actually, the pads may have holes in the case of the microvia technologies discussed later in this chapter.

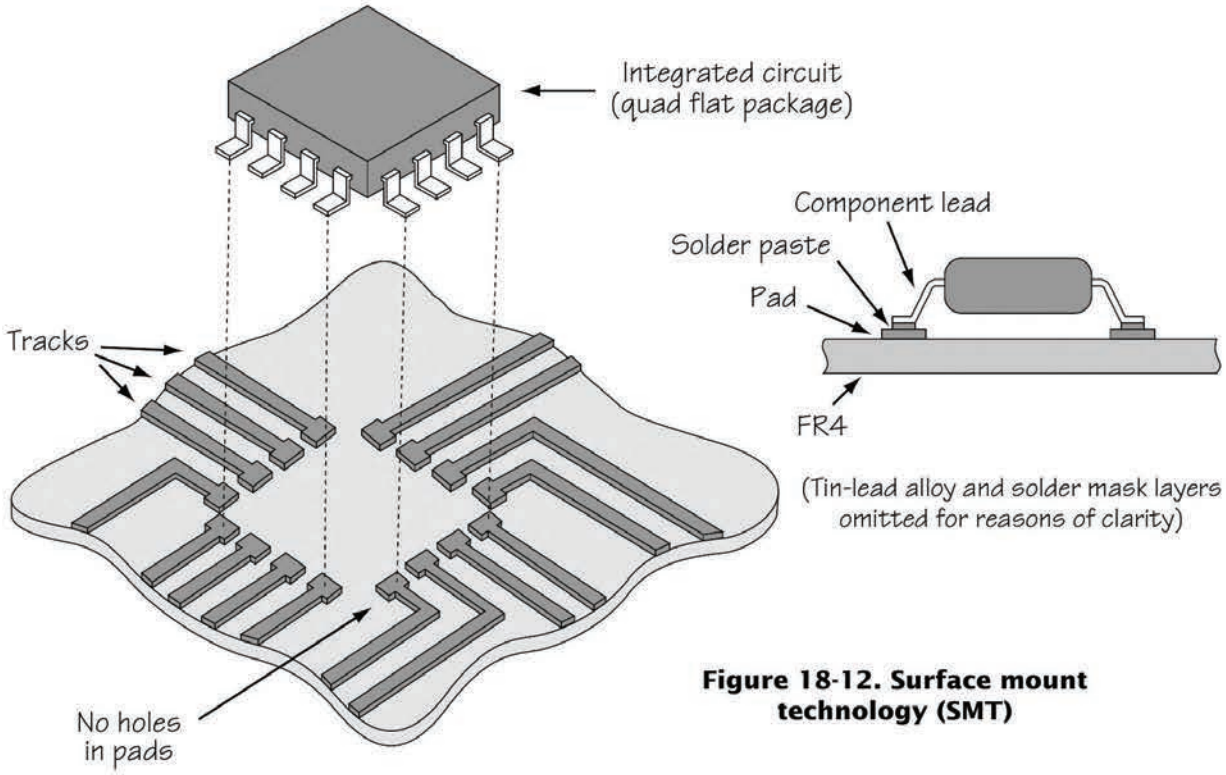


Figure 18-12. Surface mount technology (SMT)

However, vapor-phase soldering is becoming increasingly less popular due to environmental concerns.

Surface mount technology is well suited to automated processes. Due to the fact that the components are attached directly to the surface of the board, they can be constructed with leads that are finer and more closely spaced than their through-hole equivalents. The result is smaller and lighter packages which can be mounted closer together and occupy less of the board's surface area, which is referred to as *real estate*. This, in turn, results in smaller, lighter, and faster circuit boards. The fact that the components do not require holes for their leads is also advantageous, because drilling holes is a time-consuming and expensive process. Additionally, if any holes are required to make connections through the board (as discussed below), they can be made much smaller because they do not have to accommodate component leads.[9]

[9] Typical pad and hole diameters are presented in the *Holes versus Vias* section later in this chapter.

Double-sided Boards

There is a simple game played by children all over the world. The game commences by drawing three circles and three squares on a piece of paper, and then trying to connect each circle to each square without any of the connecting lines crossing each other (Figure 18-13).

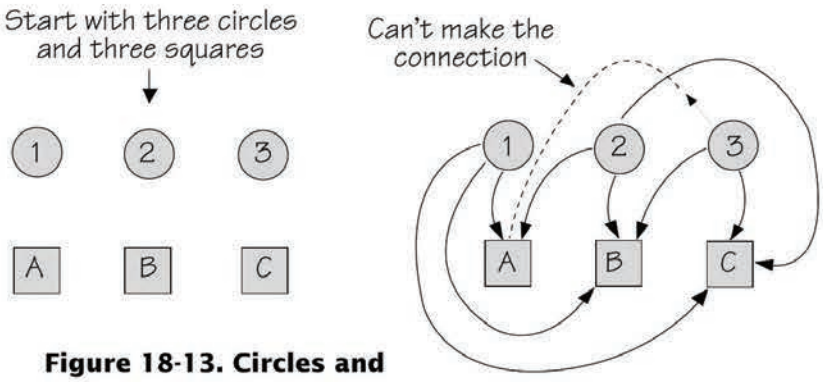


Figure 18-13. Circles and squares game

Children can devote hours to this game, much to the delight of their parents. Unfortunately, there is no solution, and one circle-square pair will always remain unconnected. This simple example illustrates a major problem with single-sided boards, which may have to support large numbers of component leads and tracks. If any of the tracks cross, an undesired electrical connection will be made and the circuit will not function as desired. One solution to this dilemma is to use wire links called *jumpers* (Figure 18-14).

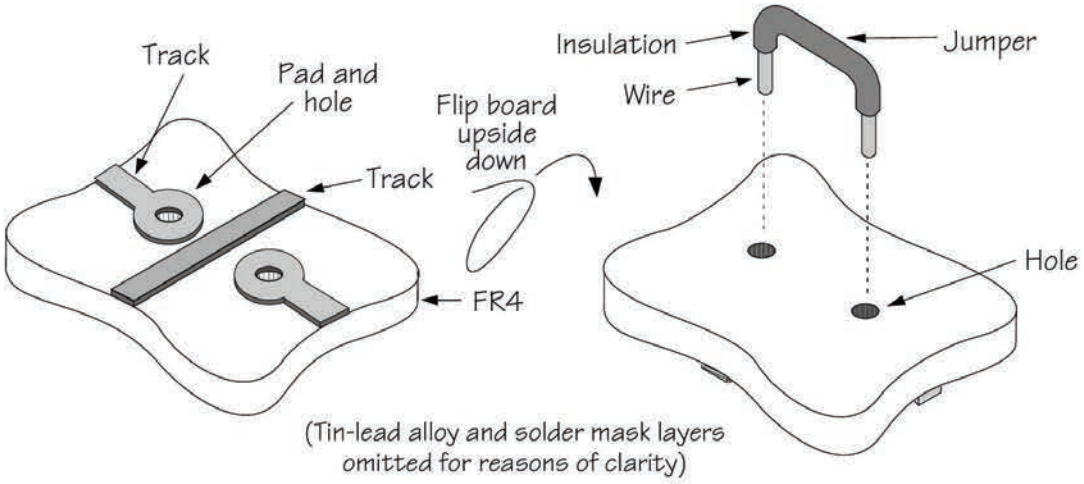


Figure 18-14. Single-sided boards: using jumpers

Unfortunately, the act of inserting a jumper is as expensive as for any other component. At some point it becomes more advantageous to employ a *double-sided* board, which has tracks on both sides.

Initially, the construction of a double-sided board is similar to that for a single-sided board. Assuming a subtractive process, copper foil is bonded to both sides of the board, and then organic resist is applied to both surfaces and cured. Separate masks are created for each side of the board, and ultraviolet light is applied to both sides. The ultraviolet radiation that is allowed to pass through the masks degrades the resist, which is then removed using an organic solvent. Any exposed copper that is not protected by resist is etched, the remaining resist is removed, and holes are drilled. However, a double-sided board now requires an additional step. After the holes have been drilled, a plating process is used to line them with copper (Figure 18-15).

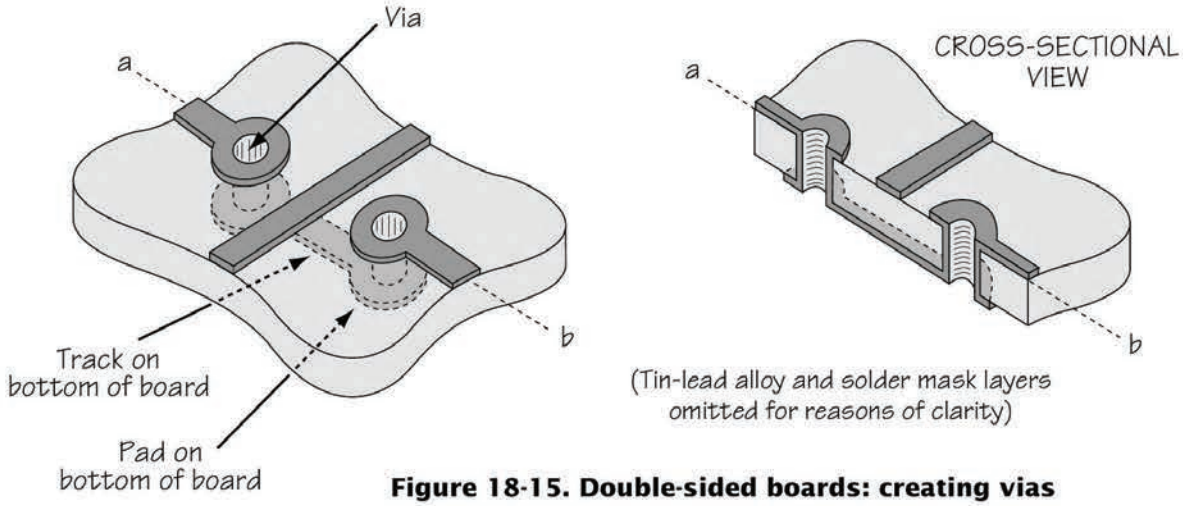


Figure 18-15. Double-sided boards: creating vias

Instead of relying on jumpers, a track can now pass from one side of the board to the other by means of these copper-plated holes, which are known as *vias*.[10,11] The tracks on one side of the board usually favor the y-axis (North-South), while the tracks on the other side favor the x-axis (East-West).

---

10 The term *via* is taken to mean a conducting path linking two or more conducting layers, but does not include a hole accommodating a component lead (see also the following section entitled *Holes versus Vias*).

11 There are a number of alternative techniques that may be used to create circuit board vias. By default, however, the term is typically understood to refer to holes plated with copper as described here.

The inside of the vias and the tracks on both sides of the board are plated with tin-lead alloy, and solder masks are applied to both surfaces (or vice versa in the case of the SMOBC-based processes, which were introduced earlier).

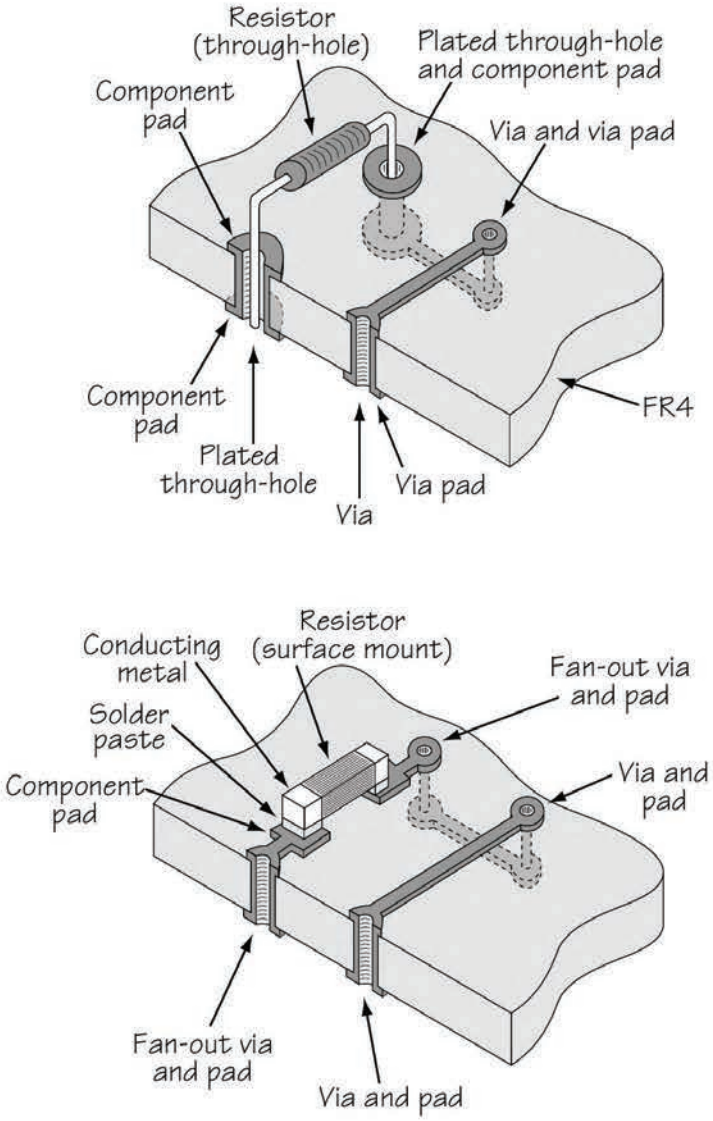
Some double-sided boards are populated with through-hole or surface mount devices on only one side. Some boards may be populated with through-hole devices on one side and surface mount devices on the other. And some boards may have surface mount devices attached to both sides. This latter case is of particular interest, because surface mount devices do not require holes to be drilled through the pads used to attach their leads (they have separate fan-out vias as discussed below). Thus, in surface mount technology, it is possible to place two devices directly facing each other on opposite sides of the board without making any connections between them.

Having said this, in certain circumstances it may be advantageous to form connections between surface mount devices directly facing each other on opposite sides of the board. The reason for this is that a through-board connection can be substantially shorter than an equivalent connection between adjacent devices on the same side of the board. Thus, this technique may be of use for applications such as high-speed data buses, because shorter connections result in faster signals.

Holes versus Vias

Manufacturers of circuit boards are very particular about the terminology they use, and woe betide anyone caught mistakenly referring to a *hole* as a *via*, or vice versa. Figure 18-16 should serve to alleviate some of the confusion.

In the case of a single-sided board (as illustrated in Figures 18-7 and 18-14), a hole that is used to accommodate a through-hole component lead is simply referred to as a *hole*. By comparison, in the case of double-sided boards (or multilayer boards as discussed below), a hole that is used to accommodate a through-hole component lead is plated with copper and is referred to as a *plated through-hole*. Additionally, a hole that is only used to link two or more conducting layers, but does not accommodate a component lead, is referred to as a *via* (or, for those purists among us, an *interstitial via*). The qualification attached to this latter case is important, because even if a hole that is used to accommodate a component lead is also used to link two or more conducting layers, it is still referred to as a plated through-hole and not a via (phew!).



(Tin-lead alloy and solder mask layers omitted for reasons of clarity)

Figure 18-16. Holes versus vias

Due to the fact that vias do not have to accommodate component leads, they can be created with smaller diameters than plated through-holes, thereby occupying less of the board's real estate. To provide a sense of scale, the diameters of plated through-holes and their associated pads are usually in the order of 24 mils (0.6 mm) and 48 mils (1.2 mm), respectively, while the diameters of vias and via pads are typically 12 mils (0.3 mm) and 24 mils (0.6 mm), respectively.

Finally, in the case of surface mount devices attached to double-sided boards (or multilayer boards as discussed below), each component pad is usually connected by a short length of track to a via, known as a *fan-out via*,[12] which forms a link to other conducting layers (an example of a fan-out via is shown in Figure 18-16). However, if this track exceeds a certain length (it could meander all the way around the board), an otherwise identical via at the end would be simply referred to as a via. Unfortunately,

[12] Some engineers attempt to differentiate vias that fall inside the device's footprint (under the body of the device) from vias that fall outside the device's footprint by referring to the former as *fan-in vias*, but this is not a widely used term.

there is no standard length of track that differentiates a fan-out via from a standard via, and any such classification depends solely on the in-house design rules employed by the designer and board manufacturer.

Multilayer Boards

It is not unheard of for a circuit board to support thousands of components and tens of thousands of tracks and vias. Double-sided boards can support a higher population density than single-sided boards, but there quickly comes a point when even a double-sided board reaches its limits. A common limiting factor is lack of space for the necessary number of vias. In order to overcome this limitation, designers may move onwards and upwards to *multilayer* boards.

A multilayer board is constructed from a number of single-sided or double-sided sub-boards.[13] The individual sub-boards can be very thin, and multilayer boards with four or six conducting layers are usually around the same thickness as a standard double-sided board. Multilayer boards may be constructed using a double-sided sub-board at the center with single-sided sub-boards added to each side (Figure 18-17a).[14] Alternatively, they may be constructed using only double-sided sub-boards separated by non-conducting layers of semi-cured FR4 known as *prepreg* (Figure 18-17b).

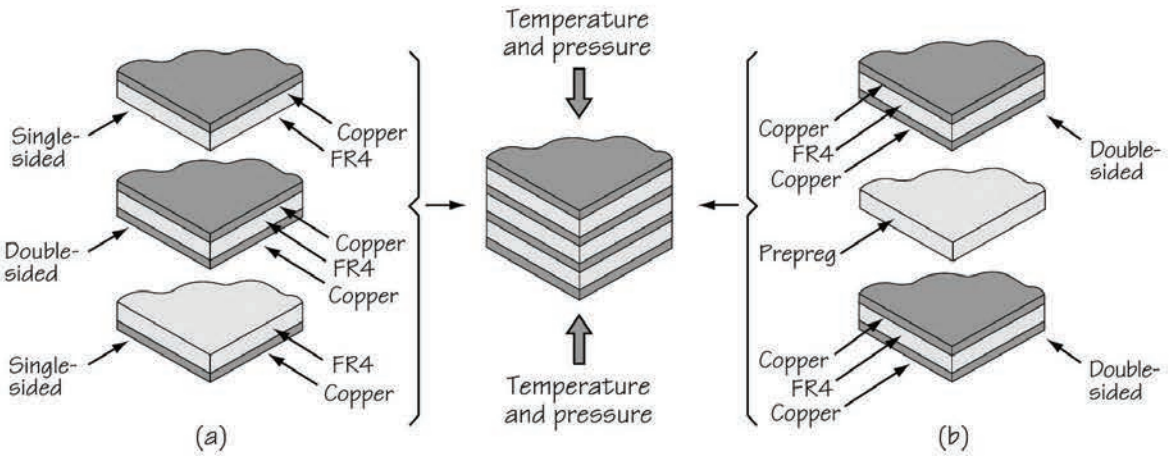


Figure 18-17. Multilayer boards: alternative structures

[13] The term "sub-board" is not an industry standard, and it is used in these discussions only to distinguish the individual layers from the completed board.

[14] This technique is usually reserved for boards that carry only four conducting layers.

After all of the layers have been etched to form tracks and pads, the sub-boards and prepreg are bonded together using a combination of temperature and pressure. This process also serves to fully cure the prepreg. Boards with four conducting layers are typical for designs intended for large production runs. The majority of multilayer boards have less than ten conducting layers, but boards with twenty-four conducting layers or more are not outrageously uncommon, and some specialized boards like *backplanes* (as discussed later in this chapter) may have 60 layers or more!

Through-Hole, Blind, and Buried Vias

To overcome the problem of limited space, multilayer boards may make use of *through-hole*, *blind*, and *buried* vias. A through-hole via passes all the way through the board, a blind via is only visible from one side of the board, and a buried via is used to link internal layers and is not visible from either side of the board (Figure 18-18).

Unfortunately, although they help to overcome the problem of limited space, blind and buried vias significantly increase the complexity of the manufacturing process. When these vias are only used to link both sides of a sub-board, that board must be drilled individually and a plating process used to line its vias with copper.

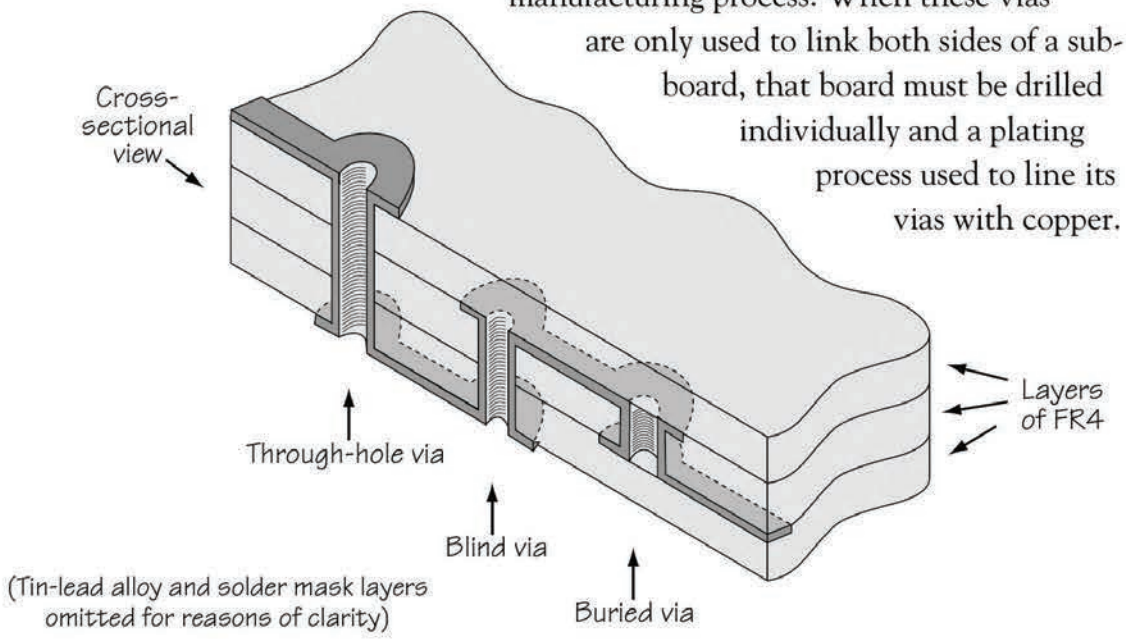


Figure 18-18. Multilayer boards: through-hole, blind, and buried vias

Similarly, when these vias only pass through a number of sub-boards, those boards must be bonded together, drilled, and plated as a group. Finally, after all of the sub-boards have been bonded together, any holes that are required to form plated through-holes and vias are drilled and plated. Blind and buried vias can greatly increase the number of tracks that a board can support but, in addition to increasing costs and fabrication times, they can also make it an absolute swine to test.

Power and Ground Planes

The layers carrying tracks are known as the *signal layers*. In a multilayer board, the signal layers are typically organized so that each pair of adjacent layers favors the y-axis (North-South) and the x-axis (East-West), respectively. Additionally, two or more conducting layers are typically set aside to be used as *power* and *ground* planes. The power and ground planes usually occupy the central layers, but certain applications have them on the board's outer surfaces. This latter technique introduces a number of problems, but it also increases the board's protection from external sources of noise such as electromagnetic radiation.

Unlike the signal layers, the bulk of the copper on the power and ground planes remains untouched. The copper on these layers is etched away only in those areas where it is not required to make a connection. For example, consider a through-hole device with eight leads. Assume that leads 4 and 8 connect to the ground and power planes respectively, while the remaining leads are connected into various signal layers (Figure 18-19).

For the sake of simplicity, the exploded view in Figure 18-19 only shows the central sub-board carrying the power planes; the sub-boards carrying the signal layers would be bonded to either side. Also note that the holes shown in the prepreg in the exploded view would not be drilled and plated until all of the sub-boards had been bonded together.

In the case of component leads 1, 2, 3, 5, 6, and 7, both the power and ground planes have copper removed around the holes. These etched-away areas, which are referred to as *anti-pads*, are used to prevent connections to the planes when the holes are plated. Similarly, the power plane has an anti-pad associated with lead 4 (the ground lead), and the ground plane has an anti-pad associated with lead 8 (the power lead).

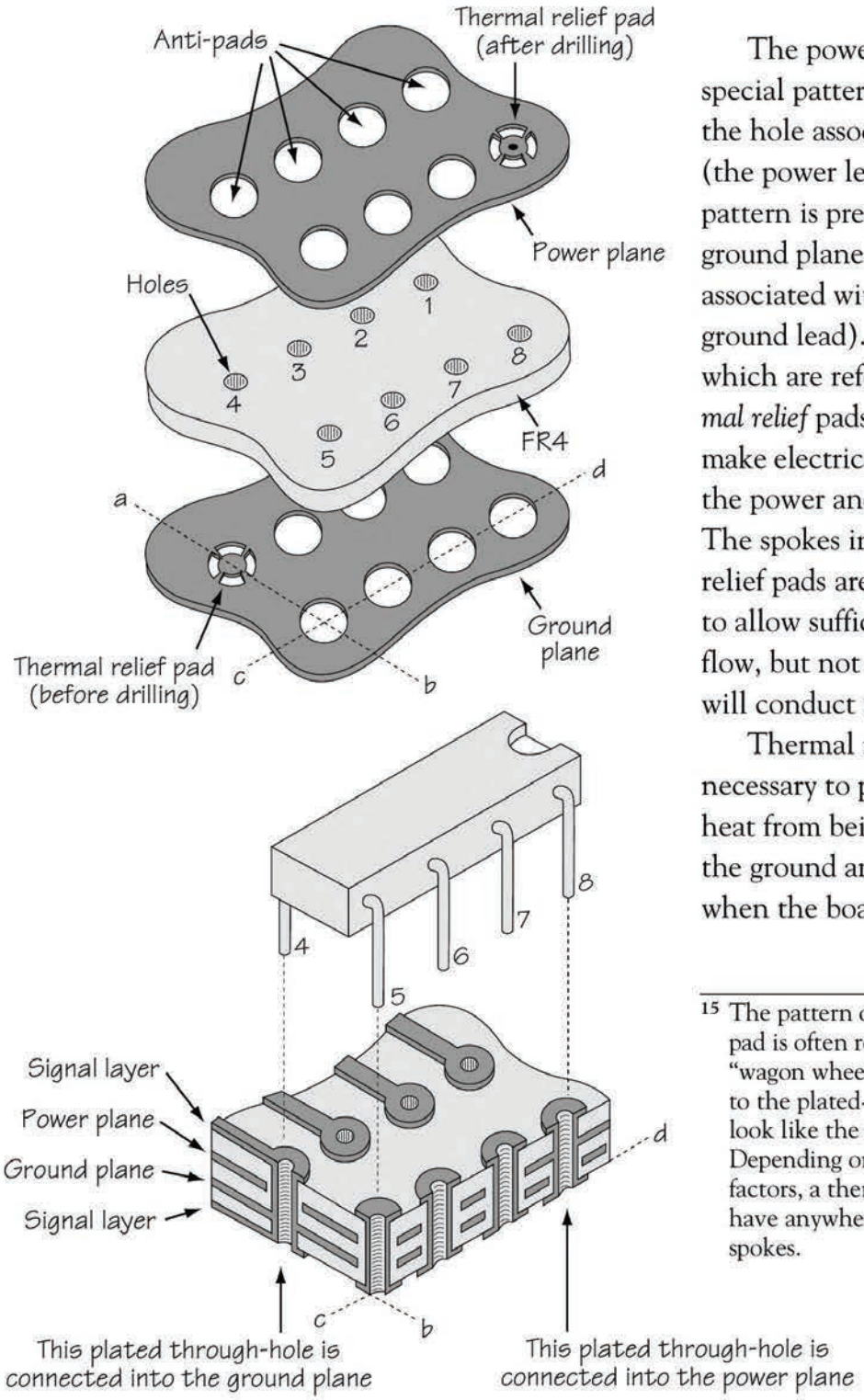


Figure 18-19. Multilayer boards: power and ground planes, and thermal relief and anti-pads

The power plane has a special pattern etched around the hole associated with lead 8 (the power lead), and a similar pattern is present on the ground plane around the hole associated with lead 4 (the ground lead). These patterns, which are referred to as *thermal relief* pads,[15] are used to make electrical connections to the power and ground planes. The spokes in the thermal relief pads are large enough to allow sufficient current to flow, but not so large that they will conduct too much heat.

Thermal relief pads are necessary to prevent excessive heat from being absorbed into the ground and power planes when the board is being

[15] The pattern of a thermal relief pad is often referred to as a "wagon wheel," because the links to the plated-through hole or via look like the spokes of a wheel. Depending on a number of factors, a thermal relief pad may have anywhere from one to four spokes.

soldered.[16] When the solder is applied, a surface-tension effect known as capillary action sucks it up the vias and plated through-holes. The solder must be drawn all the way through to form reliable, low-resistance connections. The amount of copper contained in the power and ground planes can cause problems because it causes them to act as *thermal heat sinks*. The use of thermal relief pads ensures good electrical connections, while greatly reducing heat absorption. If the thermal relief pads were not present, the power and ground planes would absorb too much heat too quickly. This would cause the solder to cool and form plugs in the vias resulting in unreliable, high-resistance connections. Additionally, in the case of wave soldering, so much heat would be absorbed by the power and ground planes that all of the layers forming the board could separate in a process known as *delamination*.

A special flavor of multilayer boards known as *Padcap* (or "Pads-Only-Outer-Layers") are sometimes used for high-reliability military applications. Padcap boards are distinguished by the fact that the outer surfaces of the board only carry pads, while any tracks are exclusively created on inner layers and connected to the pads by vias. Padcap technology offers a high degree of protection in hostile environments because all of the tracks are inside the board.

Microvia, HDI, and Build-up Technologies

One exciting recent circuit board development is that of *microvia* technology, which officially refers to any vias and via pads with diameters of 6 mils (0.15 mm) and 12 mils (0.3 mm)—or smaller—respectively. By 2002, boards using 4 mil, 3 mil, and 2 mil diameter microvias were reasonably common, and some folks were even using 1 mil microvias.

In a typical implementation, one or two microvia layers are added to the outer faces of a standard multiplayer board, which is why the term *buildup technology* is also commonly used when referring to microvia boards. Just to make things even more confusing, the term *high density interconnect (HDI)* is also commonly used. In reality, the terms *microvia*, *HDI*, and *buildup technology* are synonymous.

---

[16] For future reference, the term *pad-stack* refers to any pads, anti-pads, and thermal relief pads associated with a particular via or plated-through hole as it passes through the board.

Microvias—which are actually blind vias that just pass through one or more of the buildup layers on the outer faces of the main board—may be created using a variety of techniques. One common method is to use a laser, which can "drill" 20,000+ microvias per second.

The reason microvias are so necessary is largely tied to recent advances in device packaging technologies. It's now possible to get devices with 1,000 pins (or pads or leads or connections or whatever you want to call them), and packages with 2,000 and 4,000 pins are on the way. These pins are presented as an array across the bottom of the device. The pin pitch (the distance between adjacent pins) has shrunk to the extent that it simply isn't possible to connect the package to a board using conventional via diameters and line widths (there just isn't enough space to squeeze in all of the fan-out vias and route all of the tracks). The use of microvia technology alleviates this problem, making it possible to place a microvia in the center of a component pad, thereby eliminating the need for fan-out vias.

As a simple example, consider a simple 8-pin TSOP type package. Using conventional technologies, each of the component pads would be connected to a fan-out via (Figure 18-20a). (This is obviously a much-magnified view, and this simple component would actually be only a few millimeters in size.) Compared to the footprint of the device itself, having these fan-out vias means

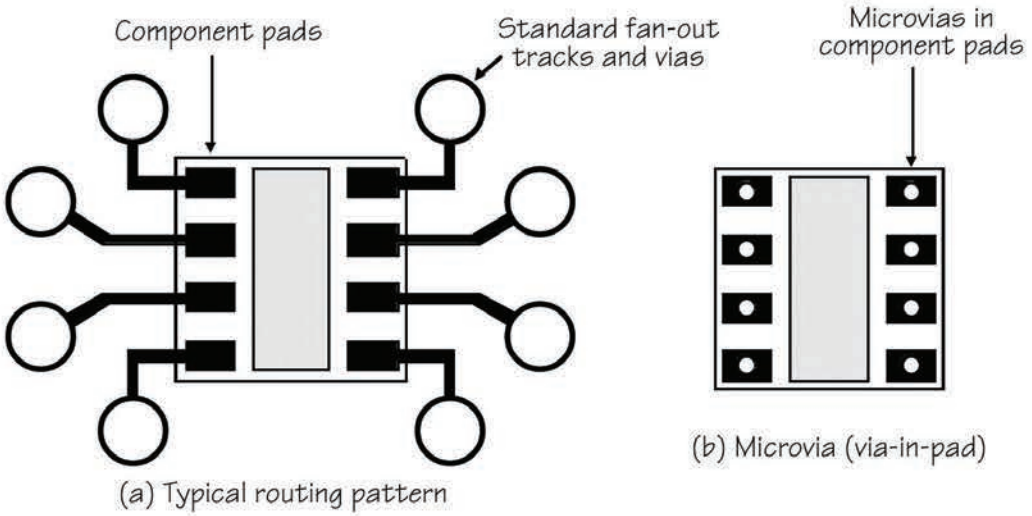


Figure 18-20. Microvias can save a lot of board real estate

that this component now occupies a substantial amount of the board's real estate. By comparison, placing vias in the component's pads using microvia technology means that the component occupies much less space on the board (Figure 18-20b).

Although using microvia technology isn't cheap per se, it can actually end up being very cost effective. In one example of which the author is aware, the use of microvias enabled an 18-layer board to be reduced to 10 layers, made the board smaller, and halved the total production cost.

Discrete Wire Technology

Discrete wire technology is an interesting discipline that has enjoyed only limited recognition, but its proponents continue to claim that it's poised to emerge as the technology-of-choice for designs that demand the highest signal speeds and component densities. Circuit boards created using this technology are known as *discrete wired boards (DWBs)*.

Multiwire Boards

The earliest form of discrete wire technology, commonly known as *multiwire*,[17,18] was developed in the late 1960s, and had gained both military and commercial acceptance by the late 1970s. The discrete wire process commences with a conventional FR4 base layer with copper foil bonded to both sides to form the power and ground planes.[19] After any thermal relief pads and anti-pads have been etched into the copper using conventional opto-lithographic processes (Figure 18-21), a layer of insulating prepreg is bonded to each side and cured. This is followed by a layer of *adhesive*,[20] or *wiring film*, which is applied using a hot roll laminating machine.

---

[17] Multiwire is a registered trademark of Advanced Interconnection Technology (AIT), Islip, New York, USA.

[18] Many thanks for the wealth of information on discrete wire technology, which was provided by Hitachi Chemical Electro-Products Division (Atlanta, Georgia, USA), I-CON Industries (Euless, Texas, USA), and MW Technologies (Aldershot, Hampshire, England).

[19] These discussions concentrate on a reasonably simple implementation: more complex boards with multiple and/or split power and ground planes are also available.

[20] Manufacturers of discrete wire boards are trying to discard the term "adhesive" on the basis that this layer is not actually "sticky."

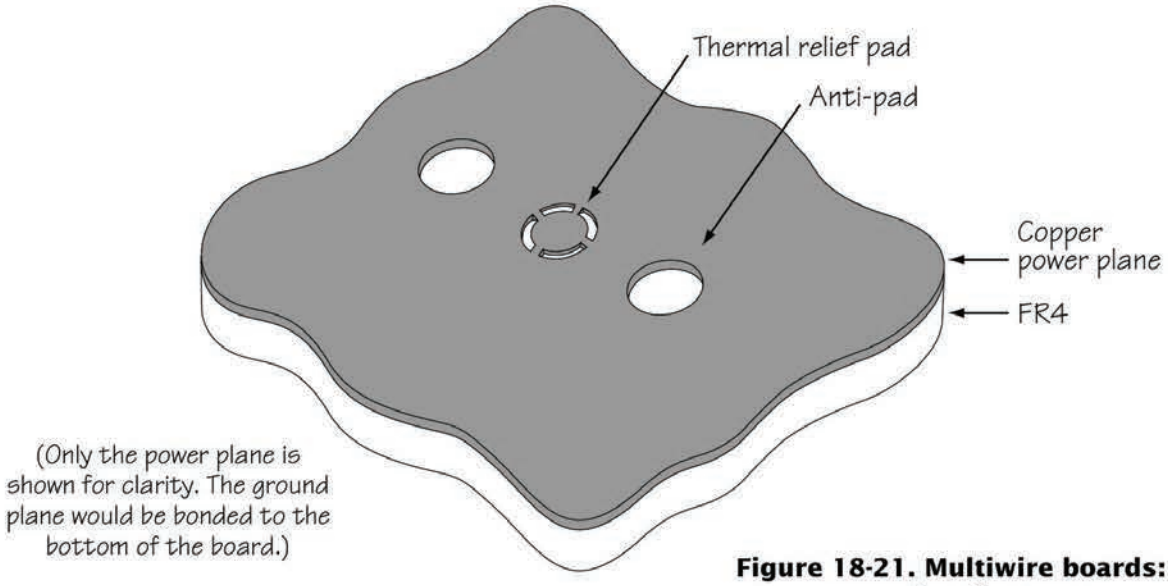


Figure 18-21. Multiwire boards: preparing the core

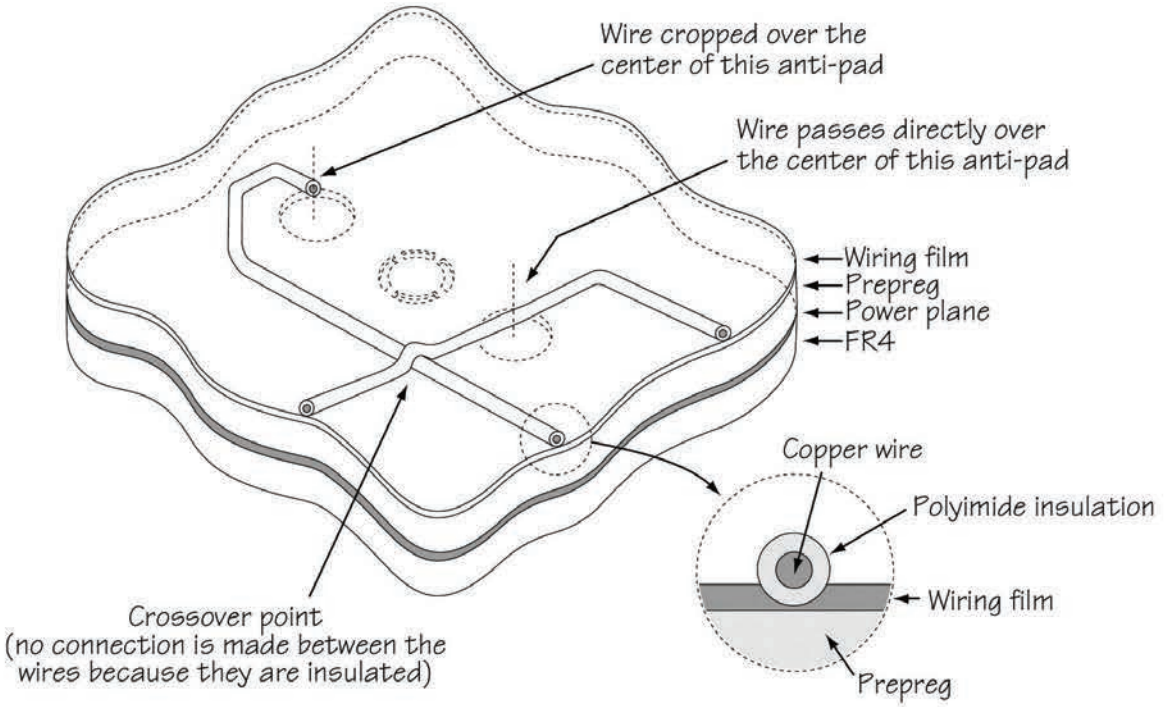


Figure 18-22. Multiwire boards: ultrasonically bonding the wire

A special computer-controlled wiring machine[21] is used to ultrasonically bond extremely fine insulated wires[22] into the wiring film. The wire is routed through all the points to which it will eventually be connected. When the last point of a particular net is reached, the wiring machine cuts the wire and positions the wiring head at the next starting point (Figure 18-22).

The wire has an insulating coat of *polyimide* and can be wired *orthogonally* (North-South and East-West), *diagonally*, or as a combination of both. Due to the fact that the wires are insulated, they can cross over each other without making any unwanted electrical connections. After all the wires have been applied to one side of the board, the board is inverted and more wires can be applied to the other side. The majority of such boards have just these two wiring layers, one on each side. If necessary, however, additional layers of prepreg and wiring film can be bonded to the outer surfaces of the board and more wiring layers can be applied. It is not unusual for a very dense multiwire board to support four wiring layers, two on each side. Supporters of this technology claim that an equivalent multilayer board might require twenty or more signal layers to achieve the same effect.

After all of the wiring layers have been created, a final layer of prepreg and copper foil are laminated onto the outer surfaces of each side of the board, and any necessary holes are drilled through the board (Figure 18-23).

The drilling process leaves the wire ends exposed in the holes. The board is now exposed to a polyimide-selective etchant, which etches away the insulation at the end of wires. Although the insulation is only etched away to a depth of approximately 0.05 mm, this is sufficient for the subsequent plating process to form a wrap-around joint, which provides mechanical reliability as opposed to a simple electrical connection.

The holes are plated with copper to form plated through-holes and vias, and then the outer surfaces of the board are etched to form any via pads and component mounting pads (Figure 18-24). (Note that the pads shown in this illustration indicate that the board is to be populated with surface mount devices.)

---

[21] This is usually referred to as an NC wiring machine, where NC stands for "Numerically Controlled."

[22] The wires used in the original process were 0.16 mm in diameter. Later processes made use of additional wire diameters of 0.10 mm and 0.06 mm.

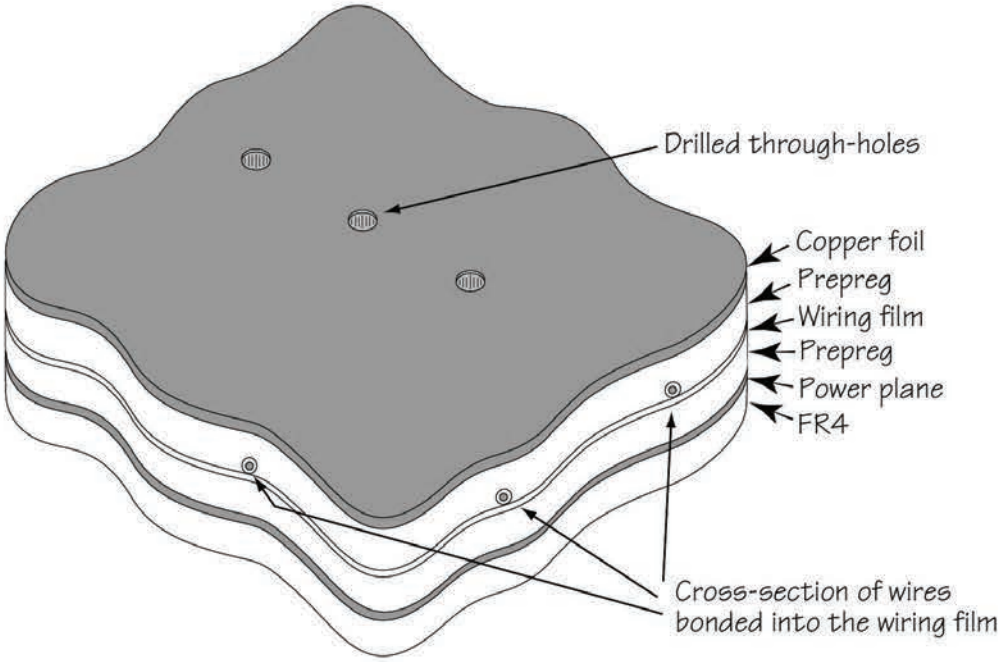


Figure 18-23. Multiwire boards: drilling the holes

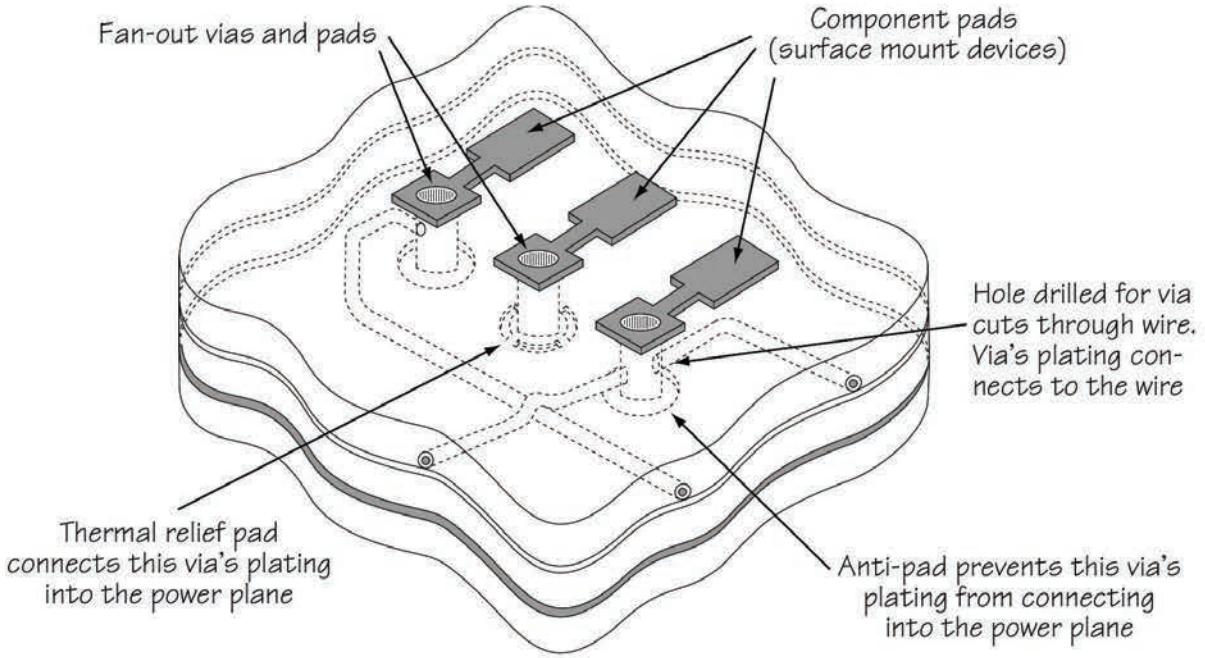


Figure 18-24. Multiwire boards: plating and etching

Finally, the board is tinned, solder masks are applied, and components are attached in the same way as for standard printed circuit boards. In fact, a completed multiwire board appears identical to a standard printed circuit board.

Microwire Boards

If multiwire is a niche market, then its younger brother, *microwire*,[23] forms a niche within a niche. Microwire augments the main attributes of multiwire with laser-drilled blind vias, allowing these boards to support the maximum number of wires and components.[24] Due to the large numbers of components they support, microwire boards typically have to handle and dissipate a great deal of heat, so their cores may be formed from sandwiches of copper-Invar-copper[25] or similar materials. This sandwich structure is used because the resulting coefficient of thermal expansion of the core combined with the other materials used in the board is almost equal to that of any components with ceramic packages that may be attached to the board. The coefficient of thermal expansion defines the amount a material expands and contracts due to changes in temperature. If materials with different coefficients are bonded together, changes in temperature will cause shear forces at the interface between them. Engineering the board to have a similar coefficient to any ceramic component packages helps to ensure that a change in temperature will not result in those components leaping off the board.

To provide a sense of scale, in one of the more common microwire implementations each copper-Invar-copper sandwich is 0.15 mm thick. The board's central core is formed from two of these sandwiches separated by a layer of insulating prepreg, where the upper sandwich forms the power plane and the lower sandwich forms the ground plane (Figure 18-25).[26]

---

[23] Microwire is a registered trademark of Advanced Interconnection Technology (AIT), Islip, New York, USA.

[24] Actually, as multiwire boards can also be created with blind vias, and as the diameter of these vias can be as small as their microwire equivalents, someone less trusting than the author might wonder whether the main justification for microwire was the fact that some people simply wanted to play with lasers.

[25] Invar is an alloy similar to bronze.

[26] More complex structures are also possible, including multiple and/or split power and ground planes.

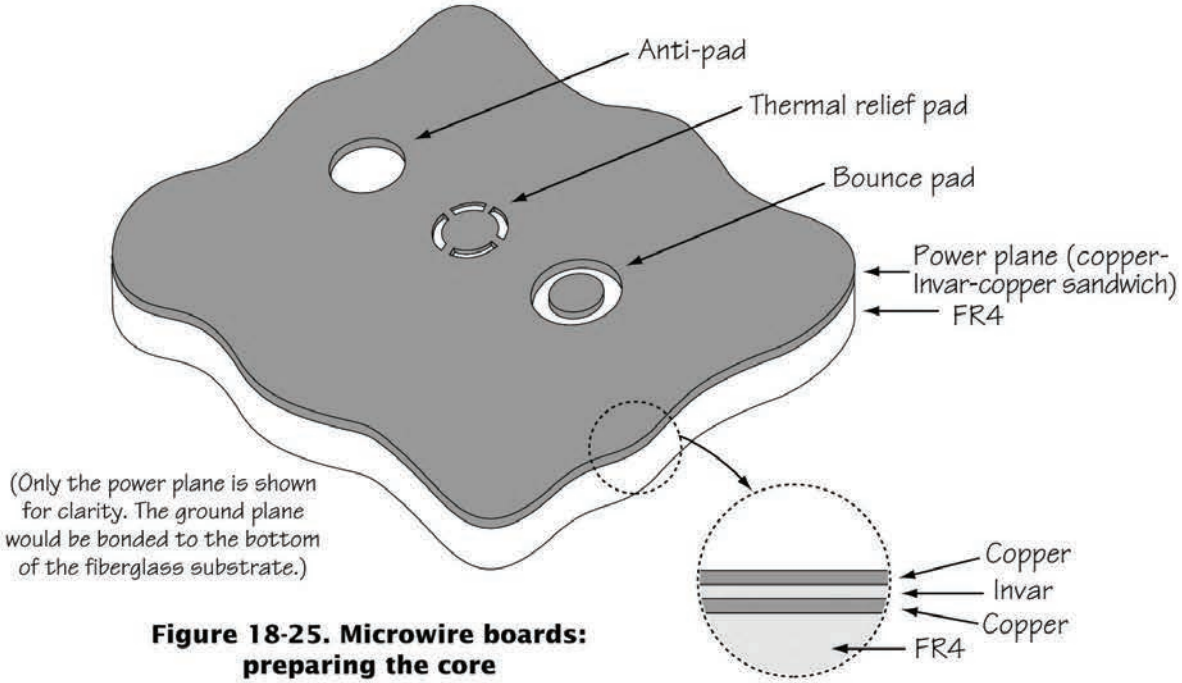


Figure 18-25. Microwire boards: preparing the core

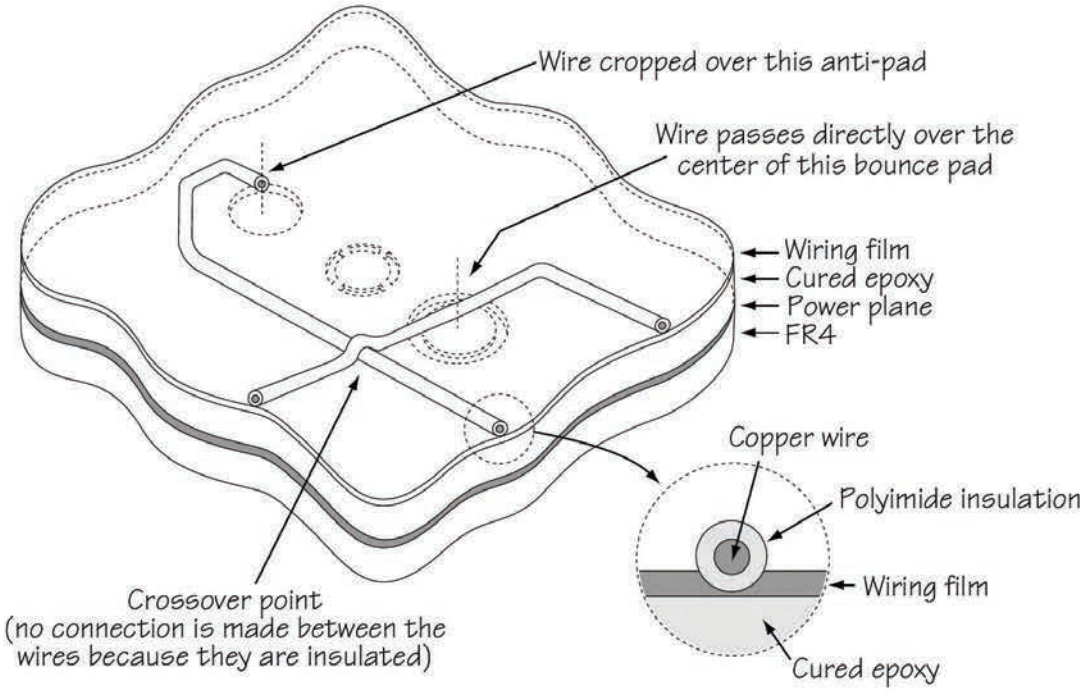


Figure 18-26. Microwire boards: ultrasonically bonding the wire

Thermal relief pads and anti-pads are etched into the upper and lower copper-Invar-copper sandwiches in the same manner as for multiwire boards. The etch is performed through all of the layers forming the sandwich, all the way down to the surface of the prepreg in the center. In addition to the standard thermal relief and anti-pads, microwire boards also contain special pads known as *bounce pads* for use during the laser drilling process discussed below.

Following the etching process, a layer of liquid epoxy is applied to the outer surfaces.[27] After this epoxy has been cured, a layer of wiring film is applied into which wires are ultrasonically bonded as for multiwire boards. However, in the case of microwire, the wires are always 0.06 mm in diameter (Figure 18-26).

As with multiwire, the majority of microwire boards have just two wiring layers, one on each side. However, if necessary, additional layers of liquid epoxy and wiring film can be bonded to the outer surfaces of the board and more wiring layers can be applied.

After all of the wiring layers have been created, a final encapsulation layer of liquid epoxy is applied to the outer surfaces of the board and cured. Now comes the fun part of the process, in which blind vias are drilled using a carbon-dioxide laser (Figure 18-27).

The laser evaporates the organic materials and strips the polyimide coating off the top of wire, but does not affect the wire itself. The laser beam continues down to the bounce pad, which reflects, or bounces, it back up, thereby stripping the polyimide coating off the bottom of the wire. To complete the board, holes for through-hole vias and, possibly, plated through-holes are created using a conventional drilling process. The through-hole vias, laser-drilled blind vias, and plated through-holes are plated with copper, along with any via pads and component pads which are plated onto the board's outer surfaces (Figure 18-28), and standard tinning and solder mask processes are then applied.

---

[27] Liquid epoxy is used to cover the wires because the glass fibers in prepreg would be impervious to the subsequent laser-drilling operations.

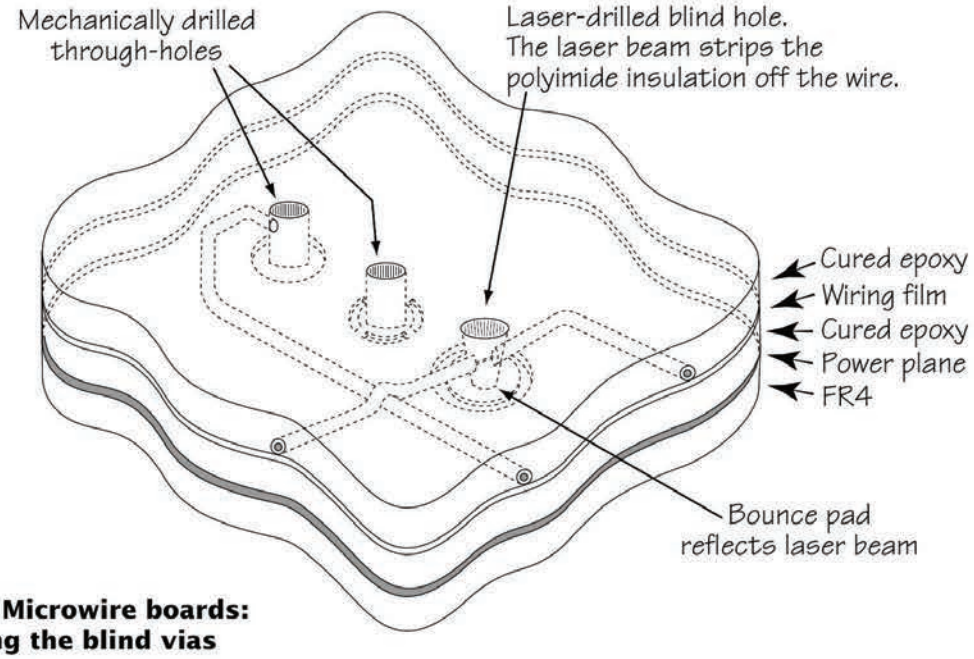


Figure 18-27. Microwire boards: laser-drilling the blind vias

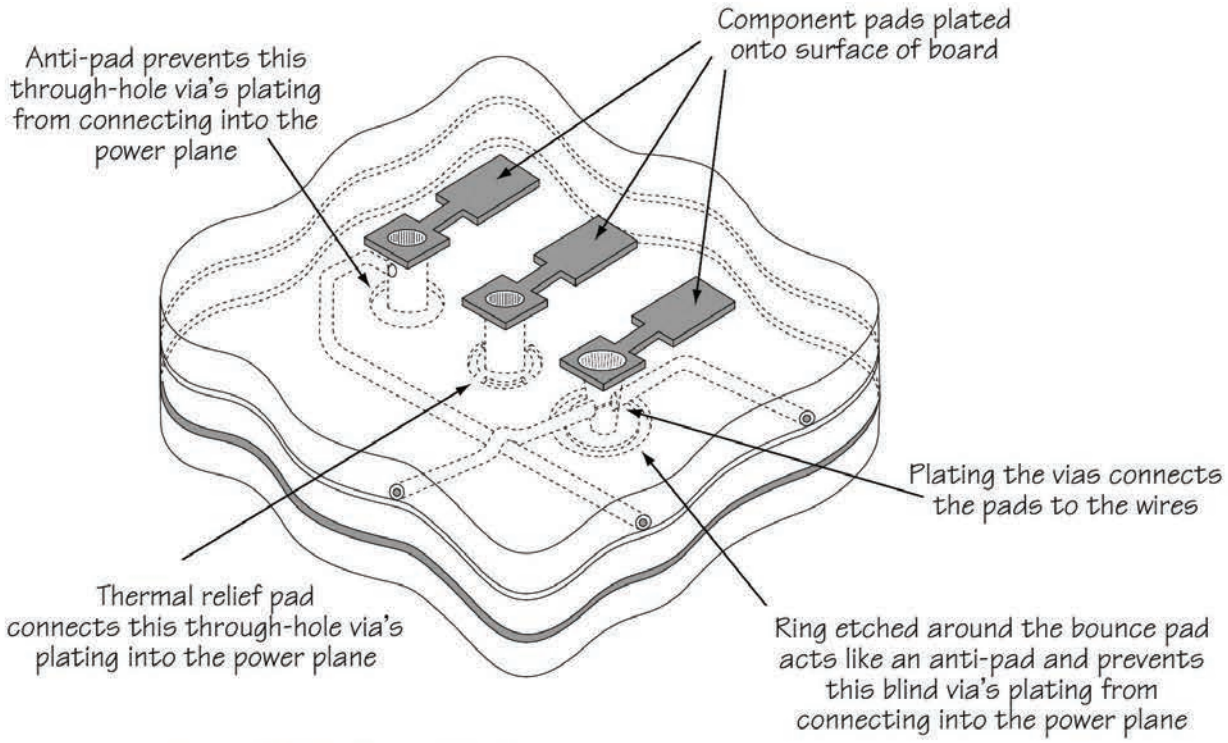


Figure 18-28. Microwire boards: plating the pads and vias

The Advantages of Discrete Wire Technology

Since its inception, discrete wire technology has attracted a small number of dedicated users, but has never achieved widespread recognition. Both multiwire and microwire processes are more expensive than their conventional printed wire counterparts, and therefore they are normally reserved for the most dense and electrically sophisticated applications. In fact, this discipline was originally regarded by many designers as simply being a useful prototyping technology, which offered fast turnaround times on design modifications. However, as many of its devotees have known for a long time, discrete wire technology offers a number of advantages over traditional printed wire boards. One obvious advantage is that discrete wired boards with only two wiring layers (one on each side) can provide an equivalent capability to ten or more signal layers in their multilayer counterparts. Thus, discrete wire technology can offer substantially thinner and lighter circuit boards.

Additionally, discrete wire technology is particularly advantageous in the case of high-speed applications. One major requirement of high-speed designs is to control the impedance (capacitance and inductance) of the interconnect, where the impedance is a function of the distance of the tracks from the power and ground planes. In the case of multilayer boards, the tracks keep on changing their distance from the planes as they pass through vias from one layer to another. To compensate for this, additional ground planes must be added between each pair of signal layers, which increases the thickness, weight, and complexity of the boards. By comparison, the wires on discrete wire boards maintain a constant distance from the planes.

Another consideration is the vias themselves. Transmitting a signal down a conducting path may be compared to shouting down a corridor, where any sharp turns in the corridor cause reflections and echoes. Similarly, the vias in a multilayer board have significant capacitive and inductive parasitic effects associated with them, and these effects cause deterioration and corruption of the signals being passed through them. In the case of discrete wire technology, the points at which wires cross over each other have similar effects, but they are orders of magnitude smaller.

Yet another consideration is that, as the frequency of signals increases, electrons start to be conducted only on the outer surface, or skin, of a conductor. This phenomenon is known as the *skin effect*. By the nature of their construction,

the etched tracks on multilayer boards have small imperfections at their edges. These uneven surfaces slow the propagation of signals and cause noise. By comparison, the wire used in discrete wire technology is uniform to ± 0.0025 mm, which is smoother than etched tracks by a factor of 10:1.

Last but not least, the optimal cross-section for a high-speed conductor is circular rather than square or rectangular, and a constant cross-section should be maintained throughout the conductor's length. Both of these criteria are met by discrete wire technology, which, in fact, offers interconnections that are close to ideal transmission lines. The end result is that discrete wire boards can operate at frequencies that are simply not possible with their traditional multilayer counterparts.[28] It is for this reason that discrete wire technology may yet be poised to emerge as the technology-of-choice for designers of high-speed circuit boards.

Backplanes and Motherboards

Another flavor of multilayer boards, known as *backplanes*, have their own unique design constraints and form a subject in their own right. Backplanes usually have a number of connectors into which standard circuit boards are plugged (Figure 18-29).

Backplanes typically do not carry any active components such as integrated circuits, but they often carry passive components such as resistors and capacitors. If a backplane does contain active components, then it is usually referred to as a *motherboard*. In this case, the boards plugged into it are referred to as *daughter boards* or *daughter cards*.

Because of the weight that they have to support, backplanes for large systems can be 1 cm thick or more. Their conducting layers have thicker copper plating than standard boards and the spaces between adjacent tracks are wider to reduce noise caused by inductive effects.

Backplanes may also support a lot of hardware in the form of bolts, earth straps, and power cables. It is not unusual for a backplane to have multiple power planes such as +5 volts, –5 volts, +12 volts, –12 volts, +24 volts, and –24 volts. Each power plane typically has an independent ground plane associated

[28] Discrete wire boards can be constructed using microcoaxial cable, and such boards can operate at frequencies that would make your eyes water!

with it to increase noise immunity. So, our hypothetical backplane would require 12 conducting layers for the power and ground planes alone. Additionally, systems containing both analog and digital circuits often require independent power and ground planes for purposes of noise immunity. Backplanes also require excellent thermal tolerance, because some of the more heroic systems can consume upwards of 200 amps and generate more heat than a rampaging herd of large electric radiators.

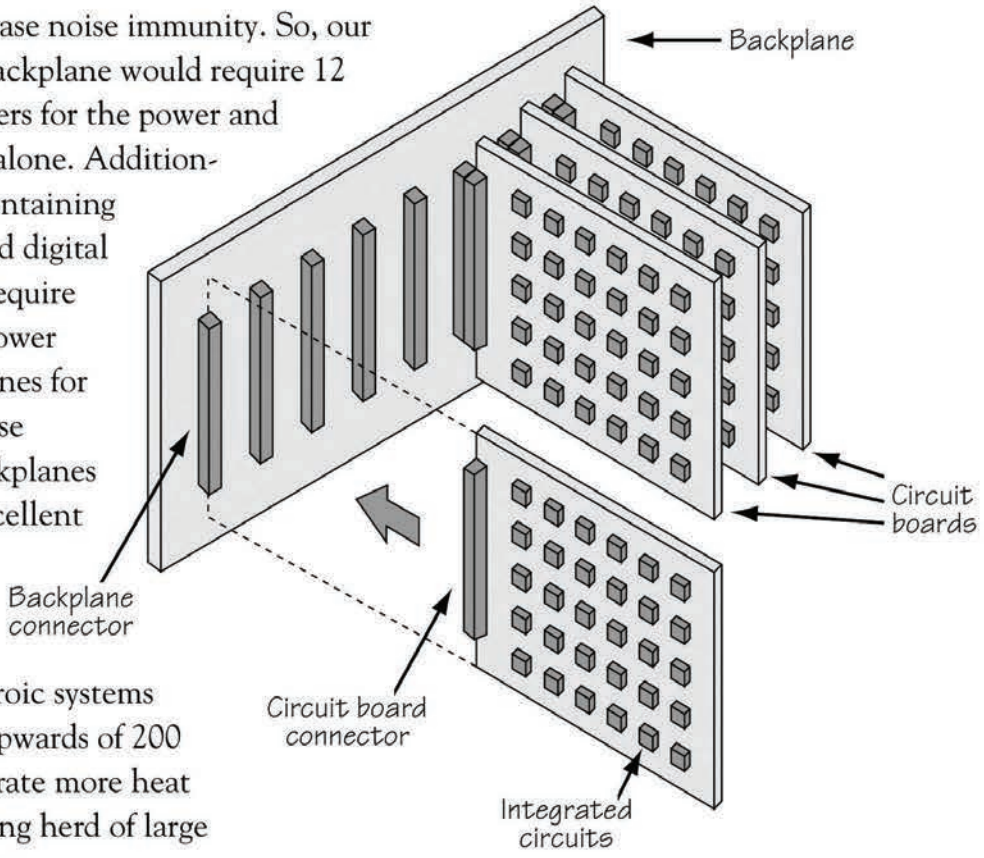


Figure 18-29. Backplanes

Backplanes may be constructed using multilayer techniques or discrete wire technology. In fact, discrete wire technology is starting to see increased use in backplane applications because of its high performance, and also due to the resulting reduction in layers, thickness, and weight.

Conductive Ink Technology

The underlying concept of conductive ink technology is relatively simple. Tracks are screen-printed onto a bare board using a conducting ink, which is then cured in an oven. Next, a dielectric, or insulating, layer is screen-printed over the top of the tracks. The screen used to print the dielectric layer is patterned so as to leave holes over selected pads on the signal layer. After the dielectric layer has been cured, the cycle is repeated to build a number of signal layers separated by dielectric layers. The holes patterned into the dielectric layers are used to form vias between the signal layers (Figure 18-30).

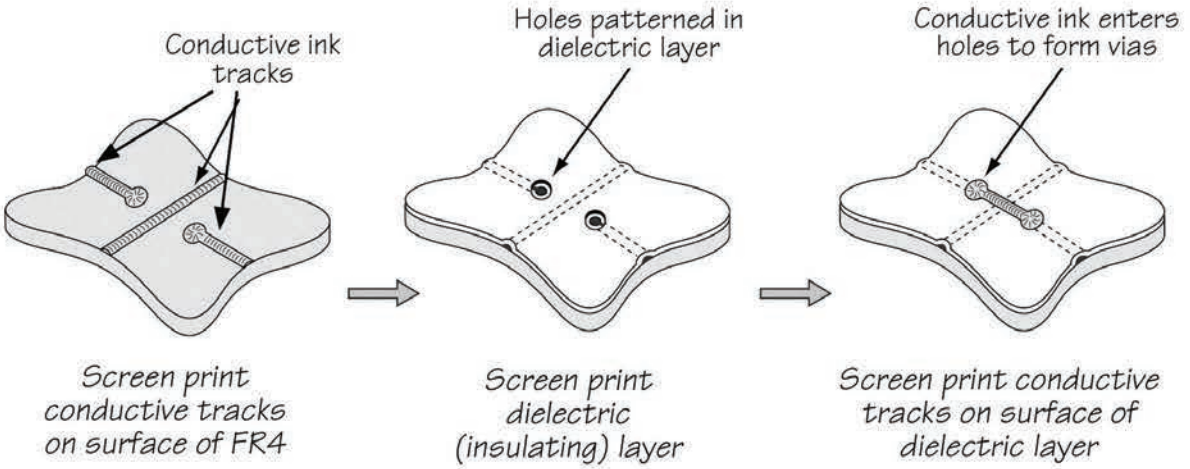


Figure 18-30. Conductive ink technology

Finally, plated through-holes and vias can be created, and components can be mounted, using the standard processes described previously. The apparent simplicity of the conductive ink technique hides an underlying sophistication in materials technology. Early inks were formed from resin pastes loaded with silver or copper powder. These inks required high firing temperatures to boil off the paste and melt the powder to form conducting tracks. Additionally, the end product was not comparable to copper foil for adhesion, conductivity, or solderability.

In the early 1990s, new inks were developed based on pastes containing a mixture of two metal or alloy powders. One powder has a relatively low melting point, while the other has a relatively high melting point. When the board is cured in a reflow oven at temperatures as low as 200°C, a process called *sintering* occurs between the two powders, resulting in an alloy with a high melting point and good conductivity.

Conductive ink technology has not yet achieved track widths as fine as traditional circuit board processes, but it does have a number of attractive features, not the least of which is that it uses commonly available screen-printing equipment. Modern inks have electrical conductivity comparable to copper and they work well with both wave soldering and reflow soldering techniques. Additionally, these processes generate less waste and are more cost-effective and efficient than the plating and etching of copper tracks.

Chip-On-Board (COB)

Chip-on-board (COB) is a relatively modern process that only began to gain widespread recognition in the early 1990s, but which is now accepted as a common and cost-effective die attachment technique. As the name implies, unpackaged integrated circuits are mounted directly onto the surface of the board. The integrated circuits are mechanically and electrically connected using similar *wire bonding*, *tape-automated bonding*, and *flip-chip* techniques to those used for hybrids and multichip modules.[29] The final step is *encapsulation*, in which the integrated circuits and their connections are covered with "globs" of epoxy resin or plastic, which are then cured to form hard protective covers (Figure 18-31).[30]

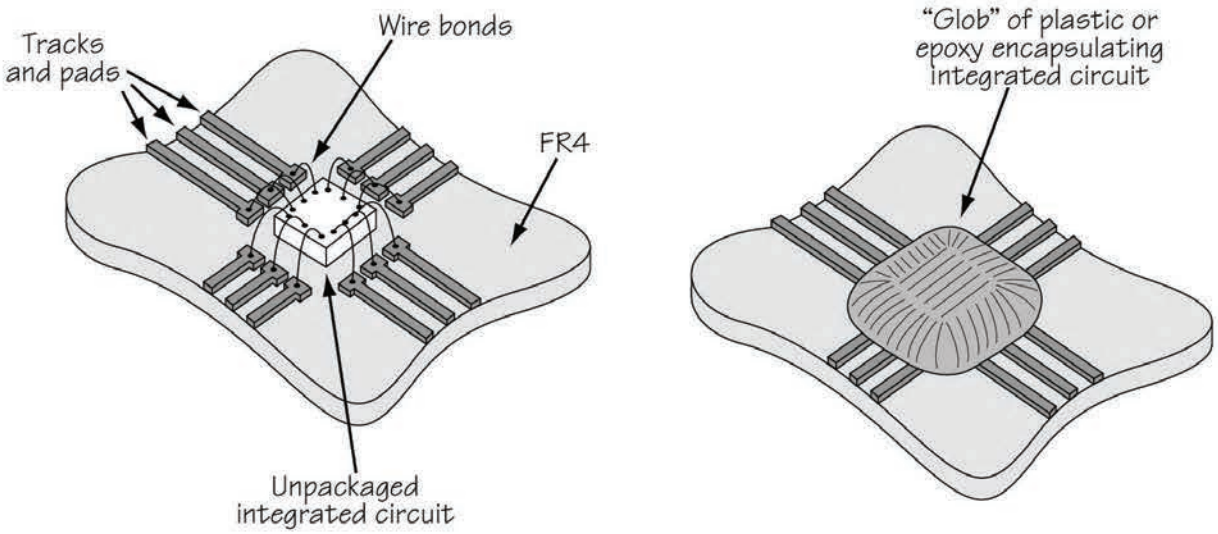


Figure 18-31. Chip-on-board (COB)

There are a number of variations of chip-on-board. For example, the designer may wish to maintain an extremely low profile for applications such as intelligent credit cards. One way to achieve this is to form cavities in the board into which the integrated circuits are inserted. Compared to surface mount technology, and especially to through-hole technology, chip-on-board

[29] Hybrids and multichip modules are introduced in Chapters 19 and 20, respectively.

[30] In addition to mechanical and environmental protection, the encapsulating material is also used to block out light.

offers significant reductions in size, area, and weight. Additionally, this technique boosts performance because the chips can be mounted closer together, resulting in shorter tracks and faster signals.

Flexible Printed Circuits (FPCs)

Last, but not least, are *flexible printed circuits (FPCs)*, often abbreviated to *flex*, in which patterns of conducting tracks are printed onto flexible materials. Surprisingly, flexible circuits are not a recent innovation: they can trace their ancestry back to 1904 when conductive inks were printed on linen paper. However, modern flexible circuits are made predominantly from organic materials such as *polyesters* and *polyimides*. These base layers can be thinner than a human hair, yet still withstand temperatures up to approximately 700°C without decomposing.

There are many variants of flexible circuits, not the least being *flexing*, or *dynamic flex*, and *non-flexing*, or *static flex*. Dynamic flex is used in applications that are required to undergo constant flexing such as ribbon cables in printers, while static flex can be manipulated into permanent three-dimensional shapes for applications such as calculators and high-tech cameras requiring efficient use of volume and not just area (Figure 18-32).

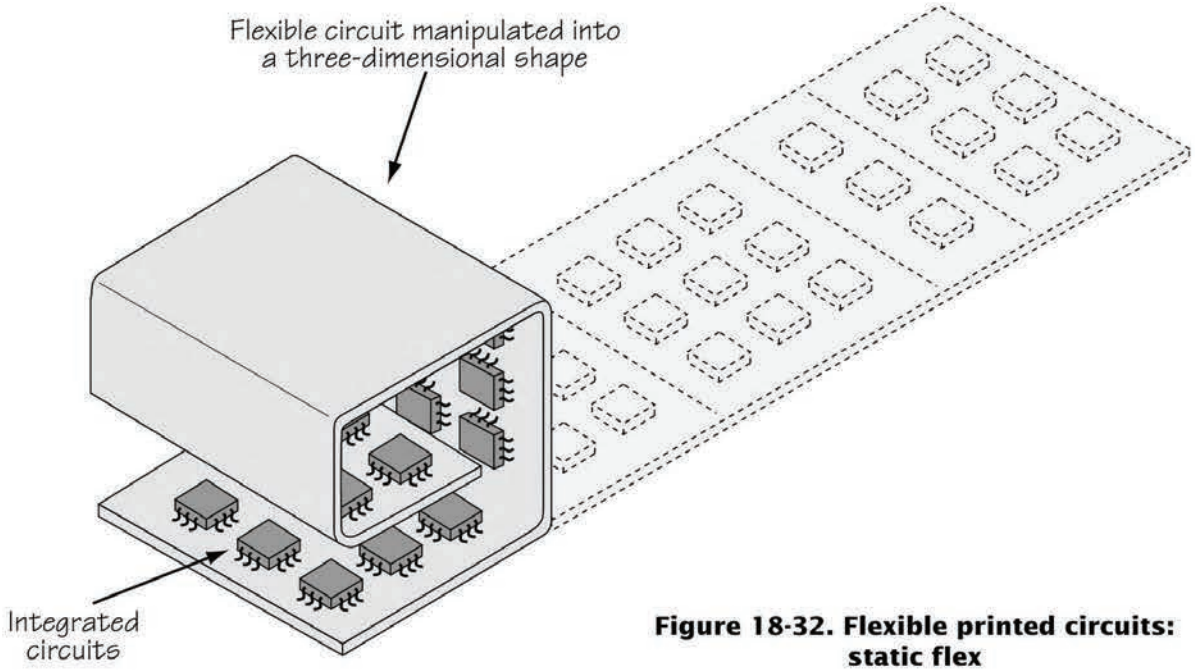


Figure 18-32. Flexible printed circuits: static flex

As well as single-sided flex, there are also double-sided and multilayer variants. Additionally, unpackaged integrated circuits can be mounted directly onto the surface of the flexible circuits in a similar manner to chip-on-board discussed above. However, in this case, the process is referred to as *chip-on-flex (COF)*.

A common manifestation of flex technology is found in hybrid constructions known as *rigid flex*, which combine standard rigid circuit boards with flexible printed circuits (Figure 18-33).

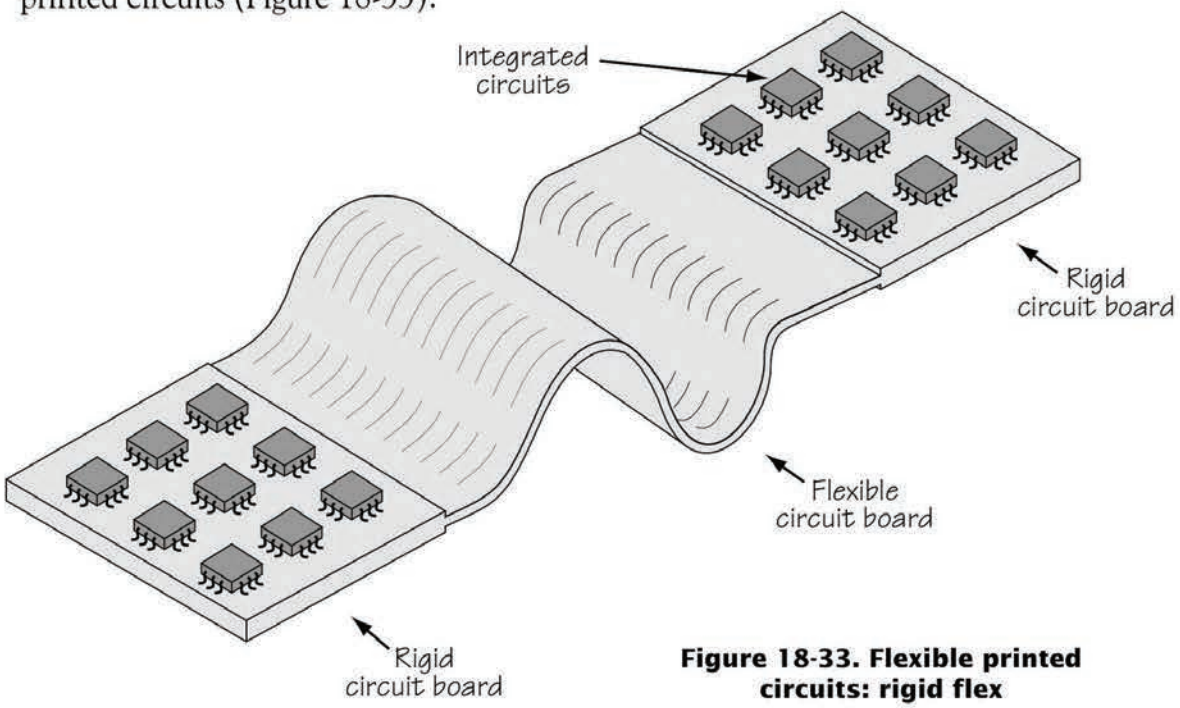


Figure 18-33. Flexible printed circuits: rigid flex

In this example, the flexible printed circuit linking two standard (rigid) boards eliminates the need for connectors on the boards (which would have to be linked by cables), thereby reducing the component count, weight, and susceptibility to vibration of the circuit, and greatly increasing its reliability.

While the use of flexible circuits is relatively low, it is beginning to increase for a number of reasons. These include the ongoing development of miniaturized, lightweight, portable electronic systems such as cellular phones, and the maturing of surface mount technology, which has been described as the ideal packaging technology for flexible circuits. Additionally, flexible circuits are amenable to being produced in the form of a continuous roll, which can offer significant manufacturing advantages for large production runs.

Chapter 19

Hybrids

The word hybrid is defined as *"the offspring resulting from crossbreeding."* Many would agree that this is an apt description for the species of electronic entities known as *hybrids*, which combine esoteric mixtures of interconnection and packaging technologies. In electronic terms, a hybrid consists of a collection of components mounted on a single insulating base layer called the *substrate*. A typical hybrid may contain a number of packaged or unpackaged integrated circuits and a variety of discrete components such as resistors, capacitors, and inductors, all attached directly to the substrate. Connections between the components are formed on the surface of the substrate; also, some components such as resistors and inductors may be fabricated directly onto the surface of the substrate.

Hybrid Substrates

Hybrid substrates are predominantly formed from alumina (aluminum oxide) or similar ceramic materials. Ceramics have many valuable properties, which have been recognized since the Chinese first created their superb porcelains during the Ming dynasty. In addition to being cheap, light, rugged, and well understood, ceramics have a variety of characteristics that make them particularly well suited to electronic applications. They are nonporous and do not absorb moisture, they can be extremely tough,[1] they have very good lateral thermal conductivity, and their coefficient of thermal expansion is close to that of silicon.

Good lateral thermal conductivity means that heat generated by the components can be conducted horizontally across the substrate and out

[1] As examples of their toughness, ceramics can be used to create artificial bone joints and to line the faces of golf clubs. In fact, during the cold war, Glock developed a handgun for espionage purposes that was fabricated almost completely out of ceramics (so it wouldn't trigger metal detectors at airports).

through its leads. The coefficient of thermal expansion defines the amount a material expands and contracts due to changes in temperature. If materials with different coefficients are bonded together, changes in temperature will cause shear forces at the interface between them. Because silicon and ceramic have similar coefficients of thermal expansion, they expand and contract at the same rate. This is particularly relevant when unpackaged silicon chips are bonded directly to the hybrid's substrate, because it helps to ensure that a change in temperature will not result in the chips leaping off the substrate.

Hybrid substrates are usually created by placing ceramic powder in a mold and firing it at high temperatures. The resulting substrates have very smooth surfaces and are flat without any significant curvature. Hybrid substrates typically range in size from 2.5 cm × 2.5 cm to 10 cm × 15 cm and are in the order of 0.8 mm thick.

Some hybrids require numbers of small holes between the top and bottom surfaces of the substrate. These holes will eventually be plated to form conducting paths between the two surfaces, at which time they start to be called *vias*. Additionally, depending on the packaging technology being used, slightly larger holes may be required to accommodate the hybrid's leads. One technique for forming these vias and lead-holes is to introduce tiny pillar structures into the mold. The ceramic powder flows around the pillars and, when the mold (including the pillars) is removed after firing, the substrate contains holes corresponding to the pillars. Until recent times there was no cost-effective method for drilling vias through a ceramic substrate after it had been fired.[2] However, developments in laser technology have made it possible to punch holes through fired substrates using laser beams.

While the majority of hybrid substrates are ceramic, a wide variety of other materials may also be employed. These include glass, small FR4 circuit boards (laminated substrates), and even cardboard. The latter may have appeared among your Christmas presents embedded in a pair of socks that play an annoying tune when you squeeze them.[3]

---

[2] A process not dissimilar to attempting to bore holes through a dinner plate.

[3] Thank you, Auntie Barbara, I wear them all the time.

The Thick-Film Process

The two most common techniques used to create tracks and components on the surface of hybrid substrates are known as the *thick-film* and *thin-film* processes. The thick-film process is based on screen-printing, an ancient art whose invention is usually attributed to the Chinese around 3,000 BC.

Creating Tracks

An optical mask is created carrying a pattern formed by areas that are either transparent or opaque to ultraviolet frequencies. The simple patterns shown in the following diagrams are used for reasons of clarity; in practice, such a mask may contain hundreds or thousands of fine lines and geometric shapes. Next, an extremely fine steel mesh the same size as the hybrid substrate is coated with a layer of photo-resistive emulsion (Figure 19-1).

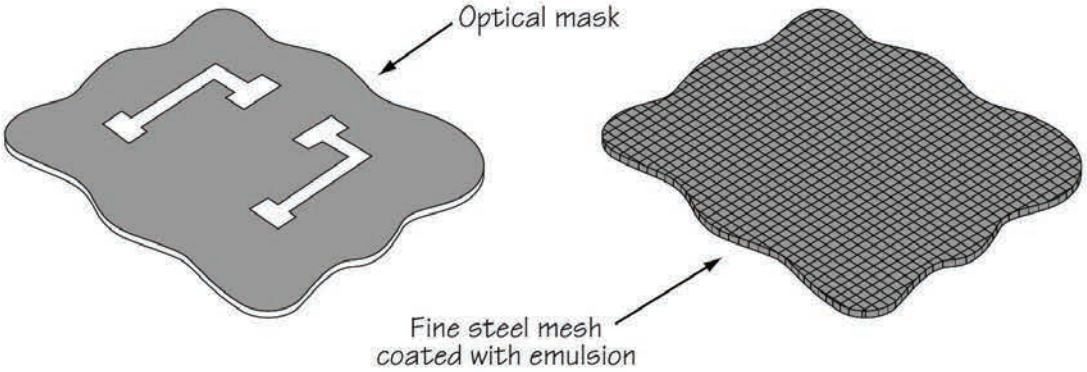


Figure 19-1. Thick-film: optical mask and emulsion-coated fine steel mesh

The emulsion-coated mesh is first dried, then baked, and then exposed to ultraviolet radiation passed through the optical mask. The ionizing radiation passes through the transparent areas of the optical mask into the emulsion where it breaks down the molecular structure of the resist. The mesh is then bathed in an appropriate solvent to dissolve the degraded resist. Thus, the pattern on the optical mask has been transferred to a corresponding pattern in the resist. The steel mesh with the patterned resist forms a screen-print mask, through which a paste containing metal and glass particles suspended in a solvent is applied onto the surface of the substrate (Figure 19-2). (Note that these figures show a magnified view of a very small portion of the entire substrate.)

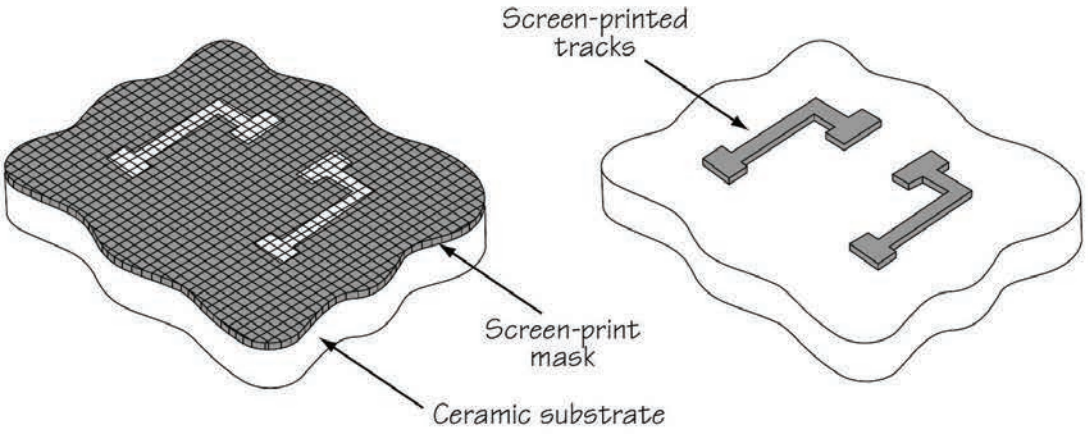


Figure 19-2. Thick-film: tracks screen-printed onto the substrate

The metal particles suspended in the paste are usually those of a noble metal such as gold, silver, or platinum, or an alloy of such metals as pladium-silver (platinum and silver). When the substrate is dried, the solvent evaporates, leaving the particles of glass and metal forming tracks on the surface of the substrate. Thick-film tracks are in the order of 0.01 mm thick. The widths of the tracks and the spaces between adjacent tracks are normally in the order of 0.25 mm, but can be as low as 0.1 mm or even finer in a leading-edge process.

Multiple layers of tracks can be printed onto the surface, each requiring the creation of a unique screen-print mask. A pattern of insulating material called a *dielectric layer* must be inserted between each pair of tracking layers to keep them separated. The dielectric patterns are formed from a paste containing only glass particles suspended in a solvent. Each dielectric layer requires its own screen-print mask and is applied using an identical process to that used for the tracking layers. Holes are included in the dielectric patterns where it is required for tracks from adjacent tracking layers to be connected to each other. Typical hybrids employ four tracking layers, commercial applications are usually limited to between seven and nine tracking layers, and a practical limit for current process technologies is around fourteen tracking layers.[4] More layers can be used, but

[4] The surface of the substrate becomes increasingly irregular with every layer that is applied. Eventually, the screen-print mask does not make sufficiently good contact across the substrate's surface, and paste "leaks out" under the edges of the patterns.

there is a crossover point where falling yields make the addition of successive layers cost-prohibitive. When all the tracking and insulating layers have been laid down and dried, the substrate is re-fired to approximately 1,000°C.

Creating Resistors

Resistors can be formed from a paste containing carbon compounds suspended in a solvent; the mixture of carbon compounds determines the resistivity of the final component. The resistors require their own screen-print mask and are applied to the substrate using an identical process to that described above (Figure 19-3).

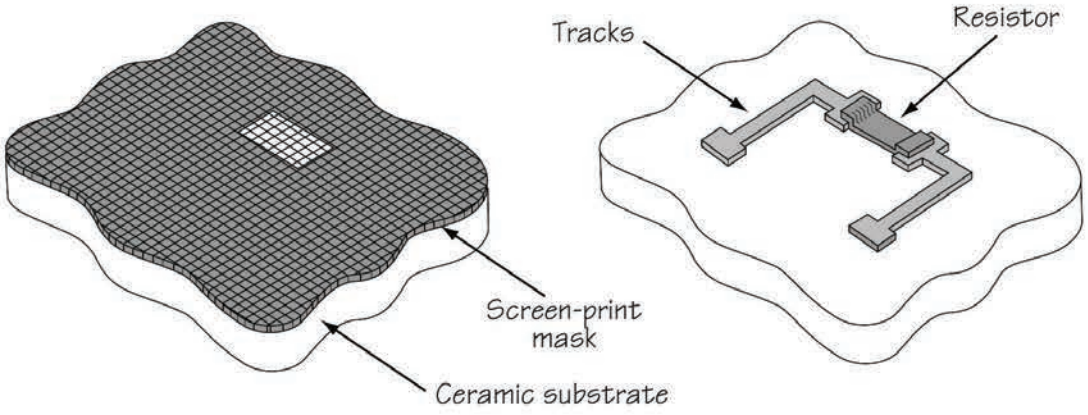


Figure 19-3. Thick-film: resistors screen-printed onto the substrate

When the substrate is dried, the solvent evaporates, leaving the carbon compounds to form resistors on the surface of the substrate. Assuming a constant thickness, each resistor has a resistance defined by its length divided by its width and multiplied by the resistivity of the paste. Thus, multiple resistors with different values can be created in a single screen-print operation by controlling the length and width of each component. However, if a low resistivity paste is used, resistors with large values will occupy too great an area. Similarly, if a high resistivity paste is used, resistors with small values will be difficult to achieve within the required tolerances. To overcome these limitations, the process may be repeated with a series of screen-print masks combined with pastes of different resistivity.

When all of the resistors have been created, the substrate is refired to approximately 600°C. An additional screen-print operation is employed to lay a protective overglaze over the resistors. This protective layer is formed from a paste containing glass particles in a solvent and is fired at approximately 450°C.

Laser Trimming

Unfortunately, creating resistors as described above is not as exact a process as one could wish. In order to compensate for process tolerances and to achieve precise values, the resistors have to be trimmed using a laser beam. There are two types of laser trimming known as *passive trimming* and *active trimming*. Passive trimming is performed before any of the integrated circuits or discrete components are mounted on the substrate. Probes are placed at each end of a resistor to monitor its value while a laser beam is used to cut parts of the resistive material away. There are a variety of different cuts which may be used to modify the resistor, including *plunge cuts*, *double plunge cuts*, or *L-shaped* cuts (Figure 19-4).

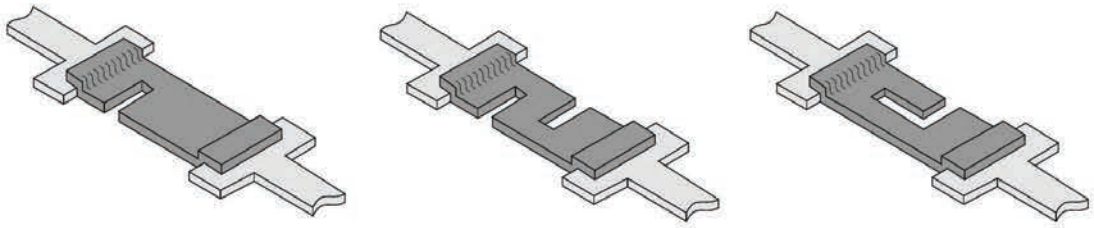


Figure 19-4. Thick-film: laser trimming of resistors

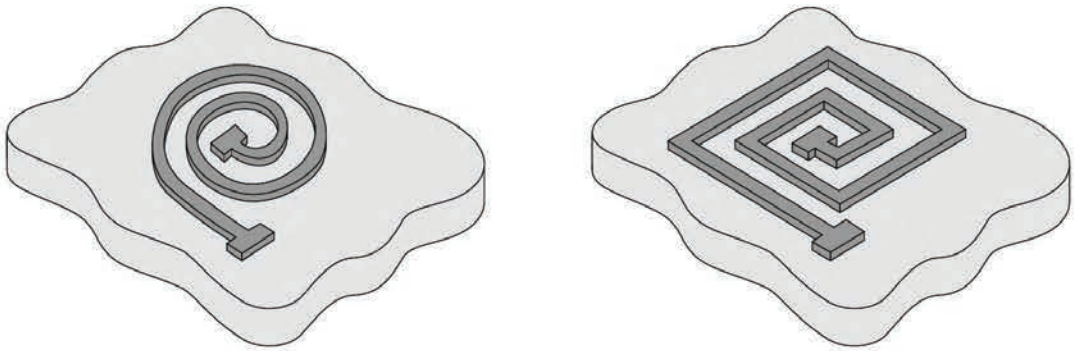
The L-shaped cut combines a plunge cut with a second cut at 90°. In this case the plunge cut provides a coarse alteration and the second cut supplies finer modifications. After each resistor has been trimmed, the probes are automatically moved to the next resistor and the process is repeated.

By comparison, active trimming is used to fine-tune analog circuits (such as active filters) and requires the integrated circuits and discrete components to be mounted on the substrate. The whole circuit is powered up and the relevant portion of the circuit is stimulated with suitable signals. A probe is placed at the output of the circuit to monitor characteristics such as amplification and

frequency response. A laser is then used to trim the appropriate resistors to achieve the required characteristics while the output of the circuit is being monitored.

Creating Capacitors and Inductors

Capacitors and inductors can also be fabricated directly onto the substrate. However, capacitors created in this way are usually not very accurate, and discrete components are typically used. If inductors are included on the substrate, they are created at the same time and using the same paste as one of the tracking layers. There are two main variations of such inductors: spiral and square spiral (Figure 19-5).



(a) Spiral (b) Square spiral

Figure 19-5. Thick-film: inductors screen-printed onto the substrate

The connection to the center-tap of the inductor can be made in several ways. A wire link can be connected to the center-tap, arched over the paths forming the spiral, and connected to a pad on the substrate outside the spiral. A somewhat similar solution is to use a track on another tracking layer to connect the center-tap to a point outside the spiral. As usual, an insulating dielectric layer would be used to separate the layer forming the inductor from the tracking layer. In yet another alternative, the center-tap can be connected to tracks on the bottom side of the substrate by means of a via placed at the center of the spiral.

Double-sided Thick-film Hybrids

Thick-film hybrids can support tracking layers on both sides of the substrate, with vias being used to make any necessary connections between the two sides. Components of all types can be mounted on both sides of the substrate as required, but active components such as integrated circuits are usually mounted only on the upper side.

Subtractive Thick-film Technology

As with any branch of electronics, new developments are always appearing on the scene. For example, a company called Silonex (www.silonex.com) has developed a process called *subtractive thick film (STF)*, which they claim is the next generation in substrate technology. STF is the integration of different technologies. In addition to standard thick film conductors, resistors, capacitors, and inductors, STF features photolithography and chemical etching steps used in conjunction with novel materials to produce unprecedented line width density and repeatability. Silonex say that STF expands the capabilities and performance of thick film modules for wireless telecommunication, instrumentation, telemetry and medical devices such as hearing aids—without sacrificing reliability or cost.

The Thin-Film Process

Thin-film processes typically employ either ceramic or glass substrates. The substrate is prepared by spluttering a layer of nichrome (nickel and chromium) alloy across the whole of its upper surface, then electroplating a layer of gold on top of the nichrome. The nichrome sticks to the substrate and the gold sticks to the nichrome. The nichrome and gold layers are each in the order of 5 μm (five-millionths of a meter) thick.

The thin-film process is similar to the opto-lithographic processes used to create integrated circuits. An optical mask is created carrying a pattern formed by areas that are either transparent or opaque to ultraviolet frequencies (Figure 19-6). As usual, the simple patterns shown in the following diagrams are used for reasons of clarity; in practice, such a mask may contain hundreds of thousands of fine lines and geometric shapes.

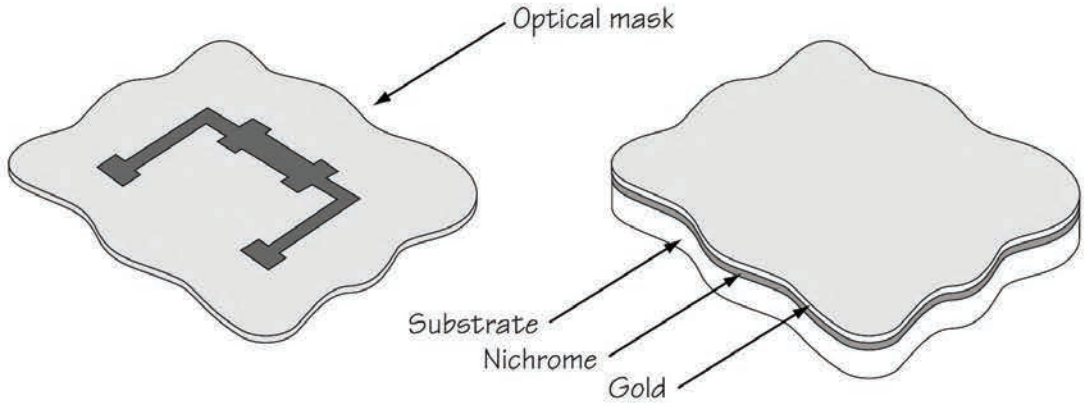


Figure 19-6. Thin-film: optical mask and substrate

The surface of the gold is coated with a layer of photo-resistive emulsion, which is first dried, then baked, and then exposed to ultraviolet radiation passed through the optical mask. The ionizing radiation passes through the transparent areas of the mask to break down the molecular structure of the emulsion. The substrate is bathed in a solvent which dissolves the degraded resist, then etched with a solvent that dissolves both the gold and nichrome from any areas left unprotected. The nichrome and gold remaining after the first mask-and-etch sequence represent a combination of tracks and resistors (Figure 19-7).

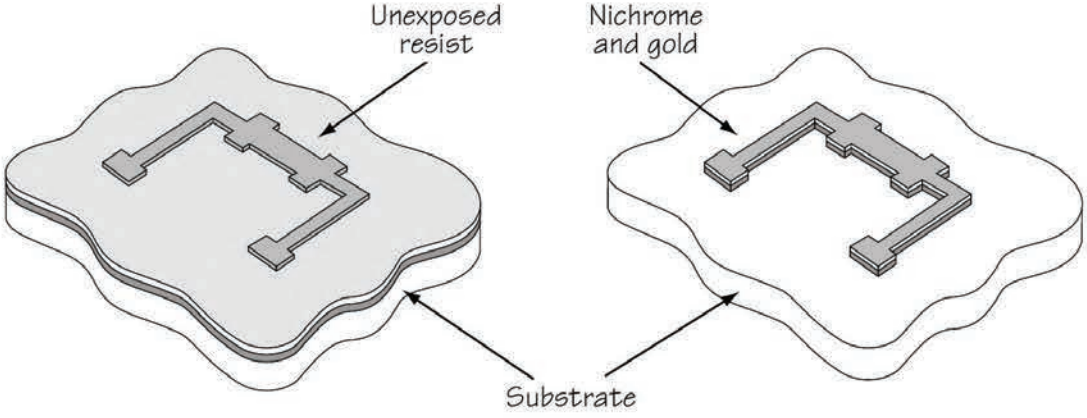


Figure 19-7. Thin-film: combined tracks and resistors

The thin-film tracks are typically in the order of 0.025 mm in width, but can be as narrow as 0.001 mm in a leading-edge process. The substrate is now recoated with a second layer of emulsion and the process is repeated with a different mask. The solvent used in this iteration only dissolves any exposed gold, but does not affect the underlying nichrome. The gold is removed from specific sites to expose the nichrome underneath, and it is these exposed areas of nichrome that form the resistors (Figure 19-8).

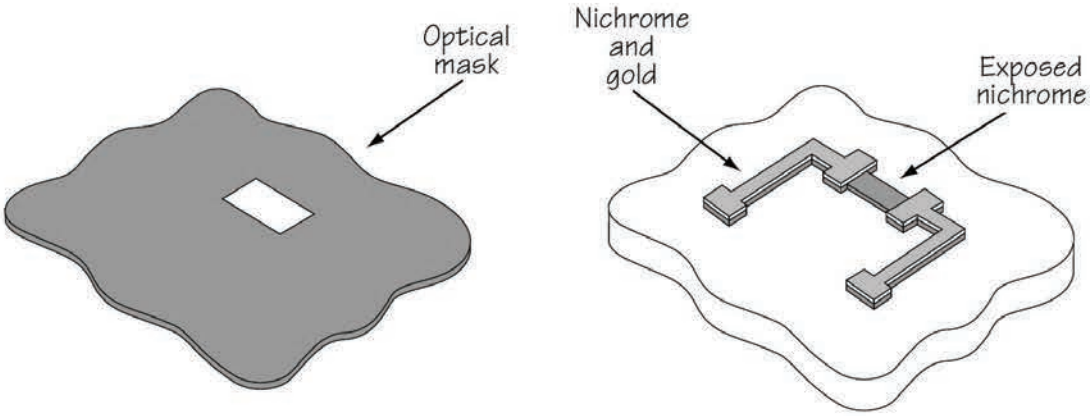


Figure 19-8. Thin-film: separated tracks and resistors

Laser Trimming

In certain respects, thin-film designers have less freedom in their control of resistance values than do their thick-film counterparts. Although the resistivity of the nichrome layer can be varied to some extent from hybrid to hybrid, the resistivity for a single hybrid is constant across the whole surface. Thus, the only way to select the value of an individual resistor is by controlling its length and width. In addition to simple rectangles, thin-film resistors are often constructed in complex concertina shapes with associated trimming blocks. The resistor values can be subsequently modified by laser trimming (Figure 19-9).

Resistors may also be created in ladder structures and their values modified by selectively cutting out rungs from the ladder using the laser. Thin-film hybrids normally employ only one tracking layer on a single side of the substrate; however, multilayer and double-sided variations are available.

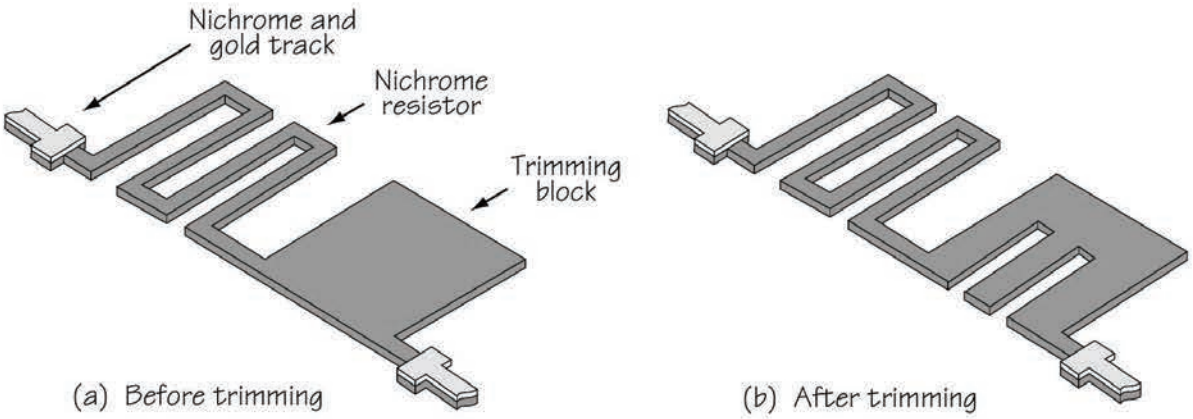


Figure 19-9. Thin-film: resistors may have complex shapes

The Assembly Process

Some hybrids contain only passive components such as resistor-capacitor networks; others contain active devices such as integrated circuits. If integrated circuits are present, they may be individually packaged surface mount devices or unpackaged bare die. Individually packaged integrated circuits are mounted onto the substrate using similar techniques to those used for surface mount technology printed circuit boards.[5] Unpackaged die are mounted directly onto the substrate using the techniques discussed below.

Attaching the Die

The primary method of removing heat from an unpackaged integrated circuit is by conduction through the back of the die into the hybrid's substrate. One technique for attaching die to the substrate is to use an adhesive such as *silver-loaded epoxy*. The epoxy is screen-printed onto the sites where the die are to be located, the die are pushed down into the epoxy by an automated pick-and-place machine, and the epoxy is then cured in an oven at approximately 180°C (Figure 19-10).

The diagrams shown here have been highly simplified for reasons of clarity; in practice, a hybrid may contain many die, each of which could have numerous pads for power and ground connections and hundreds of pads for input and output signals.

[5] These techniques were introduced in Chapter 18.

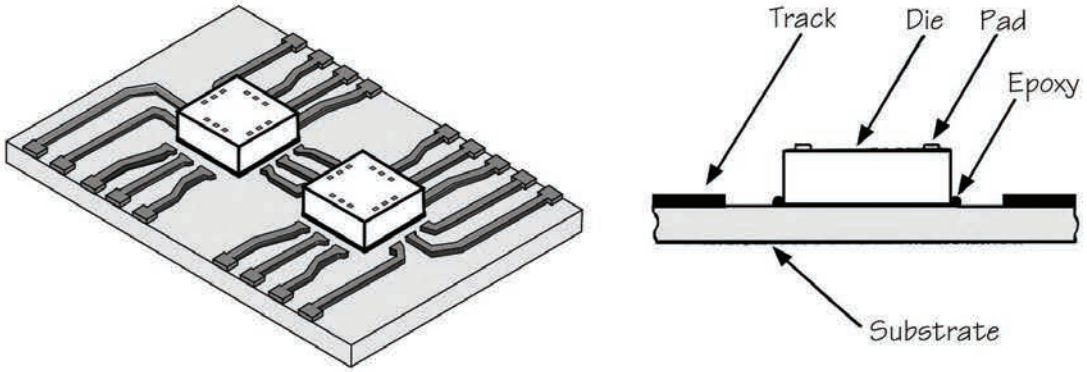


Figure 19-10. Bare die attached to substrate

An alternative method of die attachment is that of a *eutectic bond*, which requires *gold flash* (a layer of gold with a thickness measured on the molecular level) on the back of the die and also at the target site on the substrate. The substrate is heated to approximately 300ºC, and then an automatic machine presses the die against the substrate and vibrates it at ultrasonic frequencies to create a friction weld. The process of vibrating the die ultrasonically is called *scrubbing*. (These eutectic bonds are used mainly for military applications that are intended to withstand high accelerations.)

Wire Bonds

After the die have been physically attached to the substrate, electrical connections are made between the pads on each die and corresponding pads on the substrate. The most commonly used technique is *wire bonding*, in which an automatic machine makes the connections using gold or aluminum wires (Figure 19-11).

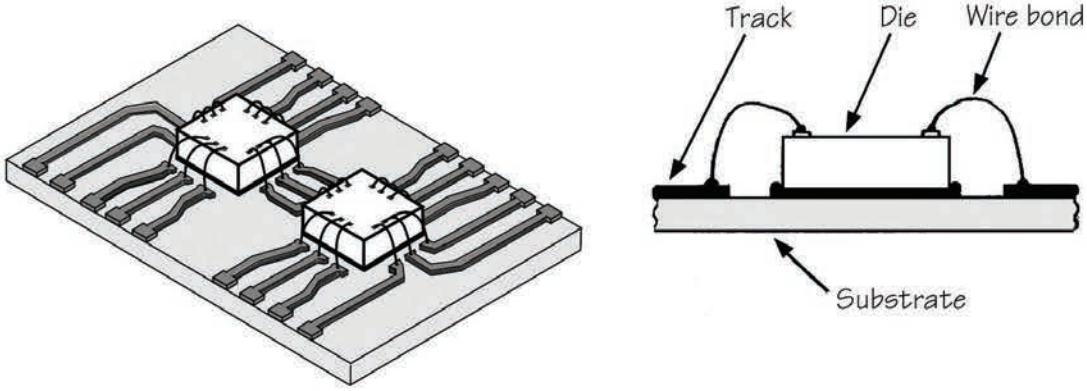


Figure 19-11. Wire bonding

These bonding wires are finer than a human hair, the gold and aluminum wires typically having diameters in the order of 0.05 mm and 0.025 mm, respectively. In one technique known as a *ball bond*, the end of the wire is heated with a hydrogen flame until it melts and forms a ball, which is then brought into contact with the target. As an alternative to ball bonds, aluminum wires are usually attached by pressing the wire against the target and scrubbing it at ultrasonic frequencies to form a friction weld. Similar wire bonds can also be used to form links between tracks on the substrate and to form connections to the hybrid's pins.

Tape-Automated Bonding

Another die attachment process is that of *tape automated bonding (TAB)*, which is mainly used for high-volume production runs or for devices with a large number of pads. In this process, a transparent flexible tape is coated with a thin layer of metal into which lead frames are patterned using standard opto-lithographic techniques. The pads on the bare die are attached to corresponding pads on the tape, which is then stored in a reel (Figure 19-12).

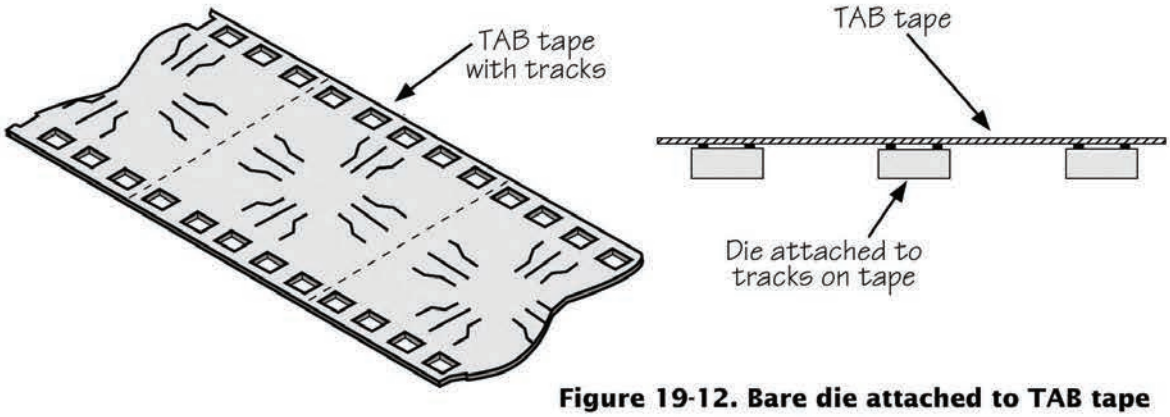


Figure 19-12. Bare die attached to TAB tape

In practice, the lead frame tracks on the tape can be so fine and so close together that they are difficult to distinguish with the human eye. When the time comes to attach the die to the hybrid's substrate, silver-loaded epoxy is screen-printed onto the substrate at the sites where the devices are to be located. This silver-loaded epoxy is also printed onto the substrate's pads to which the TAB leads are to be attached. The reel of TAB tape with attached

die is fed through an automatic machine, which pushes the die and the TAB leads down into the epoxy at their target sites. When the silver-loaded epoxy is subsequently cured, it forms electrical connections between the TAB leads and the pads on the substrate (Figure 19-13).

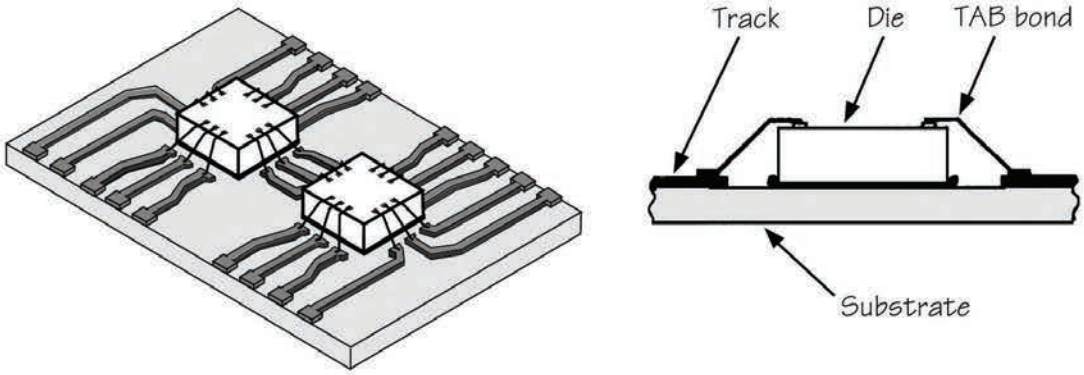


Figure 19-13. Tape-automated bonding (TAB)

Flipped-Chip Techniques

As an alternative to all of the above, there are several attachment techniques grouped under the heading of *flipped chip*. In addition to having the lowest inductance for the connections between the die and the substrate, flipped-chip techniques provide the highest packing densities because each die has a small *footprint* (the amount of area occupied by the die). Thus, flipped-chip techniques may be used where the substrate's real estate is at a premium and the die have to be mounted very closely together.

One flipped chip technique called *solder bumping* requires perfect spheres of solder to be formed on the pads of the die. The die are then "flipped over" to bring their solder bumps into contact with corresponding pads on the substrate (Figure 19-14).

One advantage of this solder bumping technique is that the pads are not obliged to be at the periphery of a die; connections may be made to pads located anywhere on a die's surface. When all of the die have been mounted on the substrate, the solder bumps can be melted in a *reflow oven* using infrared radiation or hot air. Alternatively, the solder may be melted using *vapor-phase soldering*, in which the board is lowered into the vapor-cloud above a tank containing boiling hydrocarbons. However, as for printed circuit boards, vapor-phase soldering is becoming increasingly less popular due to environmental concerns.

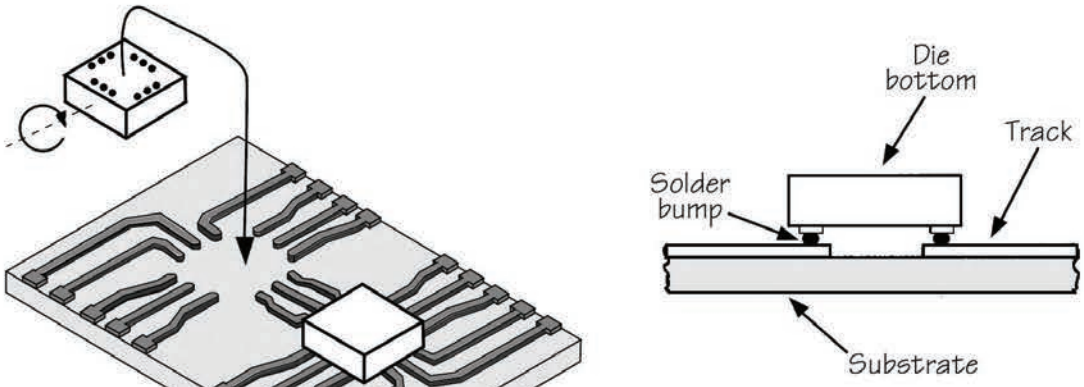


Figure 19-14. Flipped-chip – solder bumping

Another flipped chip alternative is *flipped TAB*, which, like standard TAB, may be of use for high volume production runs or for devices with a large number of pads (Figure 19-15).

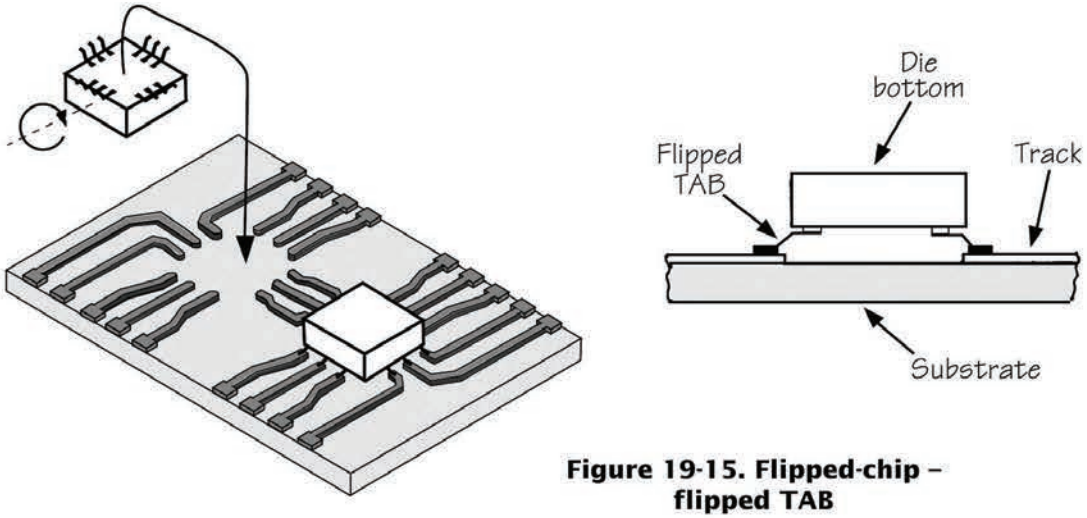


Figure 19-15. Flipped-chip – flipped TAB

Advantages of Using Bare Die

Hybrids in which bare die are attached directly to the substrate are smaller, lighter, and inherently more reliable than circuits using individually packaged integrated circuits. Every solder joint and other form of connection in an assembly is a potential failure mechanism. Individually packaged integrated circuits already have a connection from each pad on the die to the corresponding package lead and require a second connection from this lead to the substrate. By comparison, in the case of a bare die, only a single connection is required between each pad on the die and its corresponding pad on the substrate. Thus, using bare die results in one less level of interconnect, which can be extremely significant when using die with hundreds of pads.

The Packaging Process

Hybrids come in a tremendous variety of shapes, sizes, and packages. Many hybrids are not packaged at all, but simply have their external leads attached directly to the substrate. In this case, any bare die are protected by covering them with blobs of plastic polymer in a solvent; the solvent is evaporated, leaving the plastic to protect the devices (Figure 19-16).

This is very similar to the chip-on-board (COB) techniques introduced in Chapter 18. In fact, if the hybrid's substrate were a small printed circuit, this would fall into the category of chip-on-board.

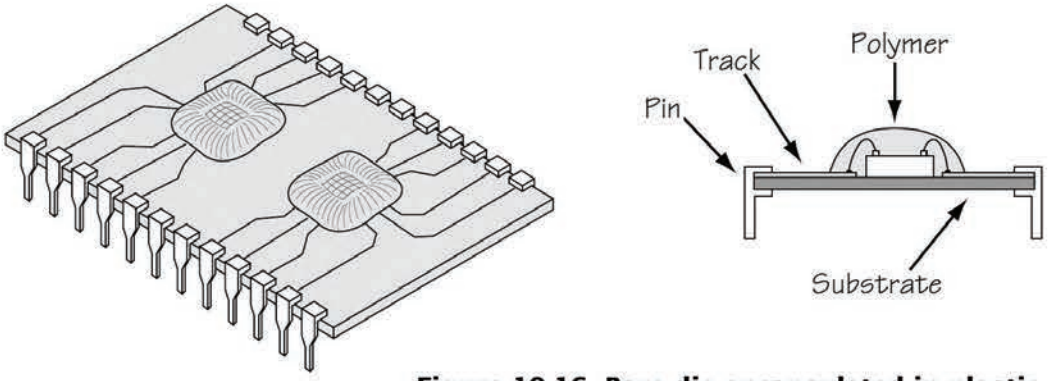


Figure 19-16. Bare die encapsulated in plastic

If the hybrid is targeted for a harsh environment, the entire substrate may be coated with polymer. If necessary, additional physical protection can be provided by enclosing the substrate with a ceramic cap filled with a dry nitrogen atmosphere (Figure 19-17).

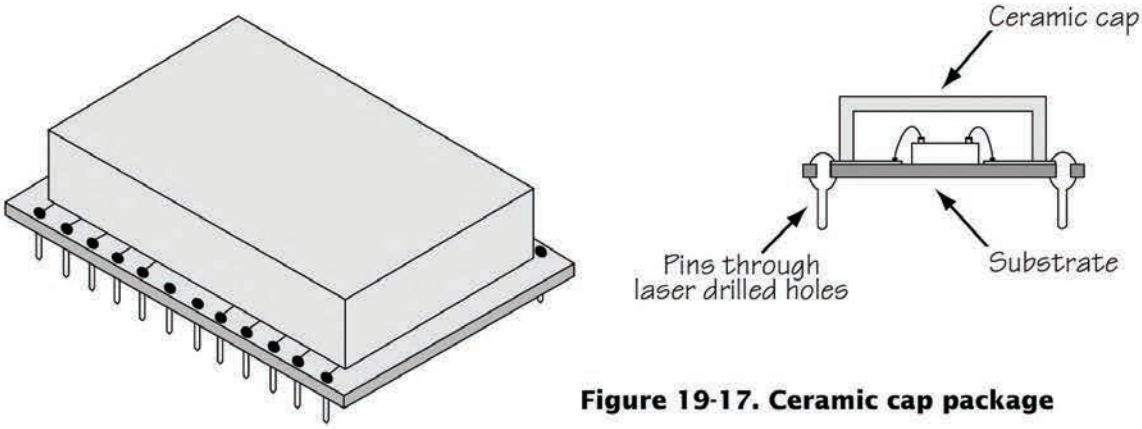


Figure 19-17. Ceramic cap package

In this case, any bare die may remain uncovered because the dry nitrogen environment protects them. If extreme protection is required, the hybrid may be hermetically sealed in alloy cans filled with dry nitrogen (Figure 19-18).

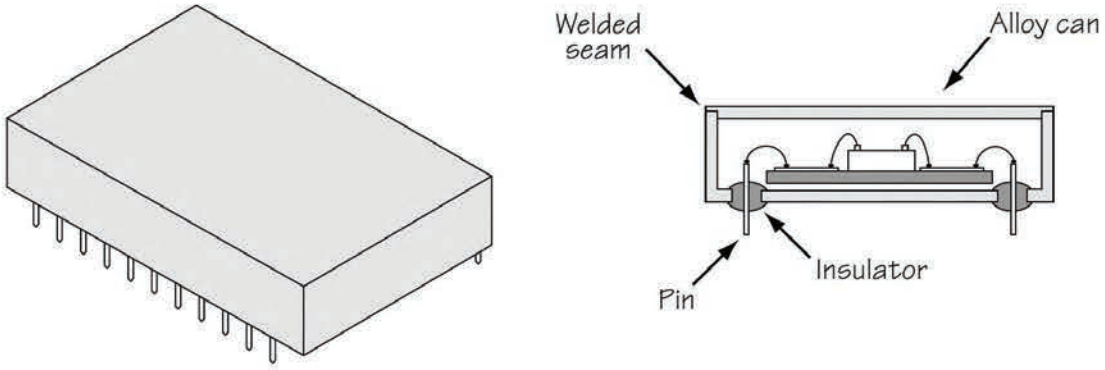


Figure 19-18. Hermetically sealed package

Finally, it should be noted that the majority of the assembly and packaging techniques introduced above are not restricted to hybrids, but may also be applied to individual integrated circuits (as discussed in Chapter 14) and multichip modules (as discussed in Chapter 20).

Chapter 20

Multichip Modules (MCMs)

Traditionally, integrated circuits, hybrids, and printed circuit boards were differentiated on the basis of their substrate materials: semiconductors (mainly silicon), ceramics, and organics, respectively. Unfortunately, even this simple categorization is less than perfect. Some hybrids use laminate substrates, which are, to all intents and purposes, small printed circuit boards. So, at what point does a printed circuit board become a hybrid with a laminate substrate?

The problem is exacerbated in the case of multichip modules—a generic name for a group of advanced interconnection and packaging technologies featuring unpackaged integrated circuits[1] mounted directly onto a common substrate (Figure 20-1).

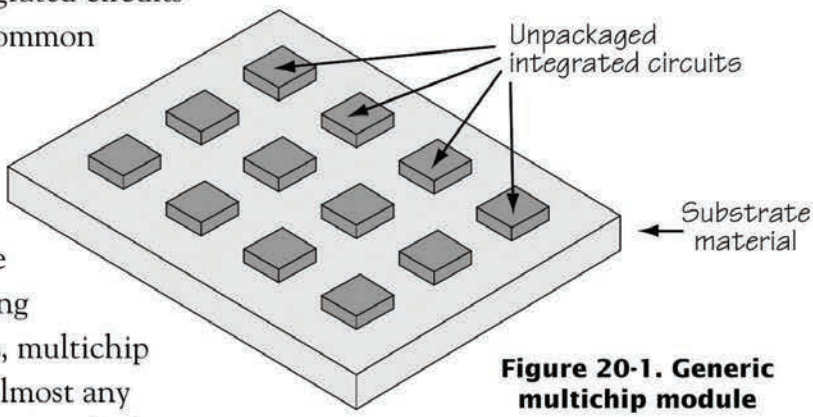


Figure 20-1. Generic multichip module

Multichip modules and hybrids can sometimes be differentiated by their choice of substrate materials and manufacturing processes. Having said this, multichip modules can be based on almost any substrate or even combinations of substrates, including laminates and ceramics, which are also used for hybrids. A slightly more reliable differentiation is that multichip modules usually have finer tracks and higher tracking and component densities[2] than hybrids. Even so, the classification of some devices is open for debate. In fact, many "hybrid designers"

---

[1] Including, but not limited to, analog, digital, application-specific, simple logic, memory devices, processor cores, and communications functions.

[2] In this context, the phrase "component density" does not necessarily refer to the number of integrated circuits mounted on the multichip module, but rather to the number of logic gates they represent. For example, a multichip module may consist of only four integrated circuits, but each of these could contain tens of millions of logic gates. See also the discussions on *Equivalent Integrated Circuits* later in this chapter.

in the late 1980s were actually designing multichip modules without knowing it. This is because the term *multichip module* and the first definitions of what constituted these devices were only introduced in 1990.

Categorization by Substrate

The original multichip module definitions were based on categorization by their substrate material:

MCM-L Laminates such as small, fine-line printed circuit boards with copper tracks and copper vias. These are usually made out of FR4 or polyimide and typically contain 5 to 25 tracking layers.

MCM-C Ceramic substrates, some of which are similar to those used for hybrids: that is, formed from a single, seamless piece of ceramic and carrying tracks that are created using thick-film or thin-film processes (or a mixture of both). However, a large proportion of ceramic multichip modules are of the cofired variety (more below), in which case they are formed from a material such as aluminum nitride or beryllium oxide and can contain as many as 260 layers (or more).

MCM-D Ceramic, glass, or metal substrates that are covered with a layer of dielectric material such as polyimide. The dielectric coat is used to modify the substrate's capacitive characteristics and tracks are created on the surface of the dielectric using thin-film processes. Modules of this type typically have around 5 tracking layers.

MCM-S Semiconductor substrates, predominantly silicon, with very fine tracks formed using opto-lithographic processes similar to those used for integrated circuits. Semiconductor substrates are also known as *active substrates*, because components such as transistors and logic gates can be fabricated directly onto their surface.[3] One additional benefit of using silicon as a substrate is that its coefficient of thermal expansion exactly matches that of any silicon chips that are attached to it.

[3] These logic gates are particularly useful for constructing special circuits to test the multichip module after it has been integrated into a larger system.

Why Use Multichip Modules?

One obvious question that may arise at this point is why should anyone wish to use multichip modules at all? If the only requirement is to increase tracking and component densities, why not simply build bigger and better integrated circuits? After all, integrated circuit fabrication processes offer the smallest components and the finest tracks known to man. The first answer to this question is based on yield. The processes used to manufacture semiconductors are as near perfect as modern technology allows, but they still leave a small number of molecular-level defects called *inclusions* in the crystalline structure. The presence of an inclusion will affect the manufacturing process and prevent an integrated circuit from functioning as planned.

For example, consider a semiconductor wafer that is to be used to build 10,000 integrated circuits, each 1 mm × 1 mm in size. Assuming that this wafer has 10 inclusions spread evenly across its surface, 10 of the chips would fail, 9,990 would work,[4] and the resulting yield would be 9,990/10,000 = 99.9%. Now consider the effect if the same wafer were used to build 100 integrated circuits, each 10 mm × 10 mm. Assuming that each inclusion affects a separate circuit, 10 of the chips would fail, 90 would work, and the yield would be 90.0%. Continuing this progression, if the wafer were used to build 25 circuits, each 20 mm × 20 mm, the yield would fall to only 60.0%.[5] Further increasing the size of the integrated circuits would rapidly reduce the yield to zero.

In addition to inclusions, problems can also arise during the doping and metalization stages—the larger the area, the more failure-prone is the process. Thus, the semiconductor game is a case of the smaller the chips, the greater the chance for successful yields. The solution is to construct a number of smaller integrated circuits and package them as a single multichip module.

Another consideration is that it may be required to implement a portion of the design using a high-speed technology such as *emitter-coupled logic (ECL)*, which uses a lot of power. If all of the logic for a multimillion-gate design were built in this technology, enough heat could be released to melt the device into a puddle. The solution is to put the speed-critical parts of the design on high-

[4] This is obviously an ideal case. In reality, there are a great number of other manufacturing considerations that will also reduce the yield.

[5] This value is also based on the assumption that each inclusion would affect a separate circuit.

speed, high-power chips; put the less critical functions on slower, lower-power chips; and package all of the chips as a single multichip module.

However, there still remains one unanswered question—why not package each chip individually and mount them on a standard printed circuit board? The answer here is one of scale. On-chip tracks are incredibly fine compared to bonding wires, package leads, and circuit board tracks. Similarly, the lengths of on-chip tracks are measured in millimeters, unlike their circuit board equivalents, which are measured in tens of centimeters. These differences in scale have a number of consequences:

a) The bonding wires combined with the package leads present very significant parasitic effects in terms of capacitance and inductance, which means that the logic gates driving off-chip have to handle relatively large currents. Thus, the logic gates driving off-chip must be made larger, which, in turn, means that they eat into the available silicon real estate, consume more power, and generate more heat.

b) The longer a track is, the more susceptible it is to picking up noise. The main technique to overcome noise is to reduce its significance by increasing the voltage differential between the logic 0 and logic 1 values. Because on-chip tracks are so short, they are relatively immune to external noise, and the voltages used to represent the logic values may differ by as little as one volt. However, circuit board tracks are much longer and are correspondingly more susceptible. Thus, the voltages used to represent the logic values at the board level typically differ by as much as three or five volts. The greater the difference between the voltages used to represent the logic values, the longer it takes for a signal to transition between values and the slower the speed of the system.

c) Last but not least, signals propagate through tracks with a finite speed—the longer a track, the longer it takes for a signal to travel down it.

Multichip modules address all of these problems. Unpackaged integrated circuits can be mounted much closer together than their packaged counterparts, resulting in shorter tracks and faster signals. Also, the fabrication processes used

for multichip modules result in much finer tracks than those found on printed circuit boards. The removal of the package leads and the shorter, finer, chip-to-chip tracks means that the logic gates driving off-chip can be smaller and less power-hungry, thereby saving silicon real estate and generating less heat. Additionally, the shorter chip-to-chip tracks are less susceptible to noise, allowing the voltage differential between logic values to be reduced, thereby minimizing the time it takes for a signal to transition between logic values.

Yet another consideration is that every solder joint and other form of connection in an assembly is a potential failure mechanism. Individually packaged integrated circuits already have a connection from each pad on the die to its corresponding package lead, and they require a second connection from this lead to the substrate. By comparison, in the case of a bare die, only a single connection is required between each pad on the die and its corresponding pad on the substrate. Thus, using bare die results in one less level of interconnect, which can be extremely significant when using die with hundreds or thousands of pads.

Cofired Ceramics

Ceramic substrates have numerous benefits (as were discussed in Chapter 19). However, once the ceramic has been fired, it is very difficult to machine and cannot be satisfactorily milled, drilled, or cut using currently available technology.[6] One solution to this problem is *cofired ceramics*, in which a number of slices of ceramic are preformed and then fired together.

The process begins with unfired, or "green," ceramic tape, which is rolled out into paper-thin layers. After each layer has been cut to the required size, microscopic holes are punched through it to form vias.[7] One technique used to punch these holes involves a steel die[8] studded with round or square pins (Figure 20-2).

---

[6] Modern laser technology offers some success in drilling and cutting ceramics, but not as much as one might hope.

[7] Cofired substrates typically use a tremendous number of blind and buried vias.

[8] In this context, the term "die" refers to a piece of metal with a design engraved or embossed on it for stamping onto another material, upon which the design appears in relief.

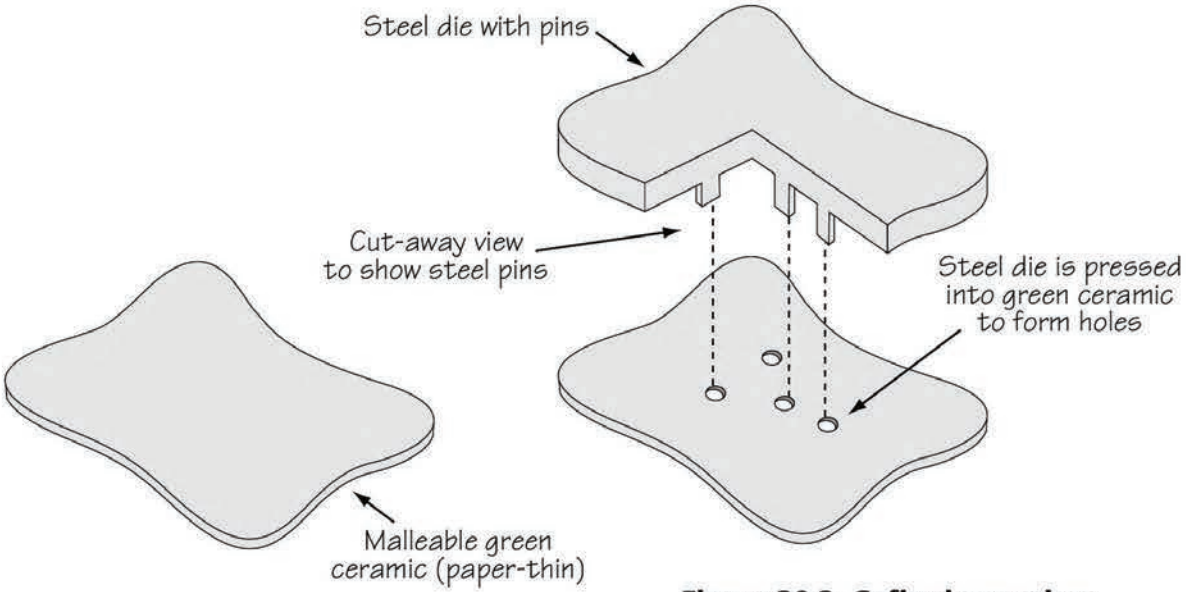


Figure 20-2. Cofired ceramics: punching holes to form vias

Unfortunately, although fast in operation, these dies take a long time to construct and the technique is incredibly expensive—each die may cost in the neighborhood of $10,000 and can typically be used for only a single layer of a single module. As an alternative, some manufacturers have experimented with programmable variants consisting of a matrix of pins, each of which may be retracted or extended under computer control. Programmable die are very expensive tools to build initially, but can be quickly reprogrammed for multiple layers and multiple modules. Their main drawback with this technique is that all of the pins are on a fixed grid. Another approach that is routinely used is that of a numerically (computer) controlled tool which punches the holes individually. This technique is relatively inexpensive, but requires much more time to complete the task.

After the holes have been punched, a screen-printing operation is used to backfill the holes with a conducting ink (a paste containing metallic powder). The mask used in this step is referred to as a *via-screen*. Next, tracks are screen-printed onto the surface of each layer using the same conducting ink. These processes are similar to the thick-film hybrid technique, but use much finer meshes for the screen print masks, thereby resulting in much finer lines and spaces (Figure 20-3).

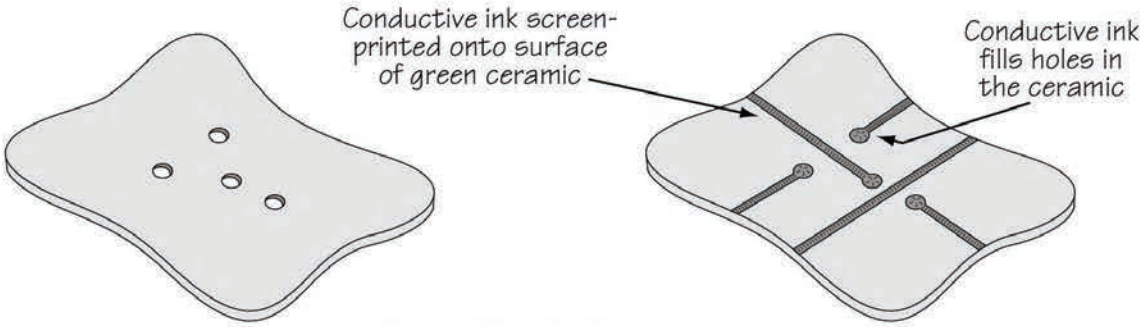


Figure 20-3. Cofired ceramics: screen-printing tracks

After all of the layers have been prepared in this way, they are stacked up, pressed together, and pre-fired at a relatively low temperature to burn off the binders used in the conductive inks and any moisture in the ceramic. The substrate is then fully fired at around 1,600°C, which melts the tracks to form good conductors and vias, and transforms the individual layers into one homogeneous piece of ceramic (Figure 20-4).

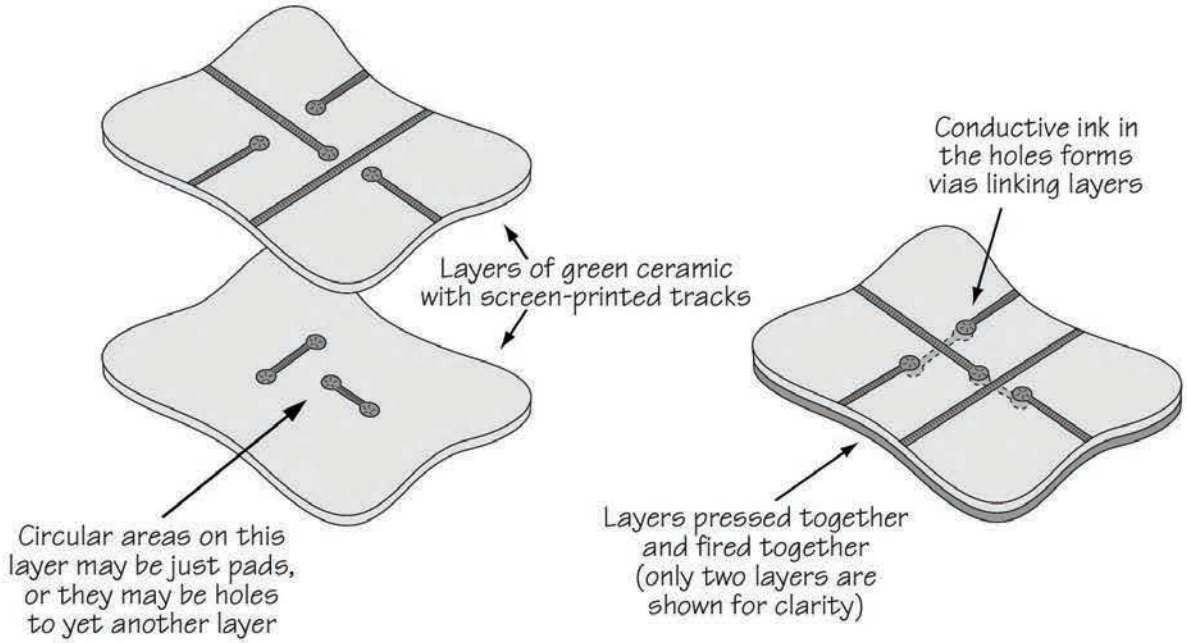


Figure 20-4. Cofired ceramics: assembling and firing the layers

Cofired ceramic substrates constructed in this way can contain an amazing number of layers. Substrates with 100 layers are not uncommon, while leading-edge processes may support 260 layers or more. To provide a sense of scale, an 85-layer substrate may be approximately 4.5 mm thick, while a 120-layer substrate will be around 6.25 mm thick. Finally, after the substrate has been fired, any tracks and pads on the surface layers are plated with a *noble metal*[9] such as gold.

One not-so-obvious consideration with cofired ceramics is that the act of rolling out the ceramic tape lines up the particles of ceramic in the direction of the roll. This causes the layers to shrink asymmetrically when the ceramic is fired. To counter this effect, the layers are alternated such that adjacent layers are rolled in different directions. The net result is symmetrical shrinkage with respect to the major axes.

Low-fired Cofired Ceramics

One problem with traditional cofired processes is that the tracks must be created using *refractory metals*.[10] Although these are reasonable conductors, they are not ideal for high-reliability and high-performance applications. For high reliability, one would ideally prefer a noble metal like gold, while for high performance at a reasonable price one would ideally prefer copper. (Surprisingly, copper is one of the best conductors available for high-frequency applications, especially those operating in the microwave arena.) Unfortunately, metals like gold and copper simply vaporize at refractory temperatures.

Low-fired cofired processes use modern ceramic materials whose compositions are very different from those of traditional ceramics. To differentiate the two, the term *HTCC (high-temperature cofired ceramics)* is used to refer to high-temperature cofired ceramics requiring tracks to be formed using refractory metals, while the term *LTCC (low-temperature cofired ceramics)* refers to low-fired cofired ceramics. These new materials can be fired at temperatures as low as 850ºC. This allows tracks to be screen-printed using copper-based or noble metal-based conductive inks (the firing must take place in a nitrogen-muffled furnace to prevent the metals from oxidizing).

---

9 Noble metals are those such as gold and platinum, which are extremely inactive and are unaffected by air, heat, moisture, and most solvents.

10 Refractory metals are those such as tungsten, titanium, and molybdenum, which are capable of withstanding the extremely high, or refractory, temperatures necessary to fire the ceramic.

Assembly and Packaging

To a large extent, multichip module packages are created on a design-specific basis. Due to the high lead counts involved, the package tends to be an integral part of the module itself. This means that packaging cannot be the afterthought it is in the case of some hybrids. For our purposes in the discussion below, we will assume a module with a cofired ceramic substrate carrying four bare die.

Because cofired substrates are constructed from multiple layers, designers have the ability to create quite sophisticated structures. For example, holes called *cavities*, or *wells*, can be cut into a number of the substrate's upper layers before they are combined together and fired (Figure 20-5).

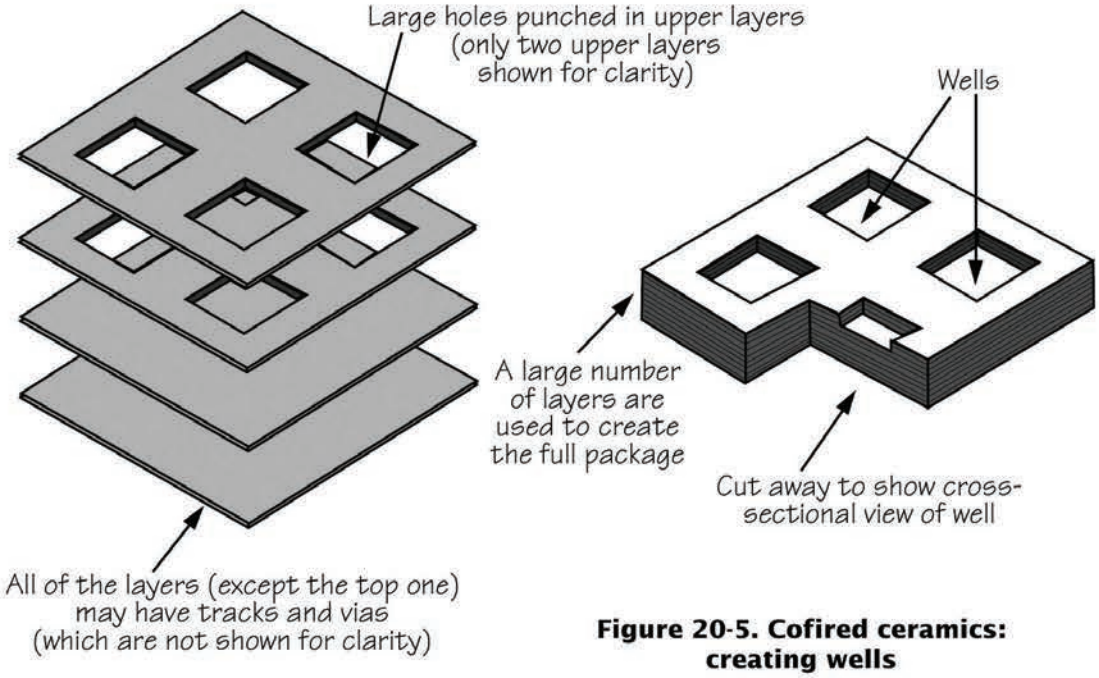


Figure 20-5. Cofired ceramics: creating wells

Remembering that the layers of ceramic are paper thin, the wells are actually formed from whatever number of layers is necessary to bring them to the required thickness. These wells will eventually accommodate the die, while the remaining ceramic will create a bed for the package's lid.[11] Before the die are mounted on the substrate, however, the external leads must be attached.

[11] Note that, with the exception of the upper layer, each of the layers forming the wells may also be used to carry tracks.

Pin Grid Arrays

In the case of a packaging style known as *pin grid arrays (PGAs)*, the module's external connections are arranged as an array of conducting leads, or pins, in the base. In many PGAs, the leads are flat brazed directly onto the bottom layer of the device. Alternatively, in a similar manner to the way in which the wells were created in the upper layers, smaller wells may be cut into a number of the substrate's lower layers before they are combined together and fired. Each of these wells is connected to a track constructed on one of the inner layers. After firing, the wells are lined with a thin layer of metal and the leads are inserted (Figure 20-6).

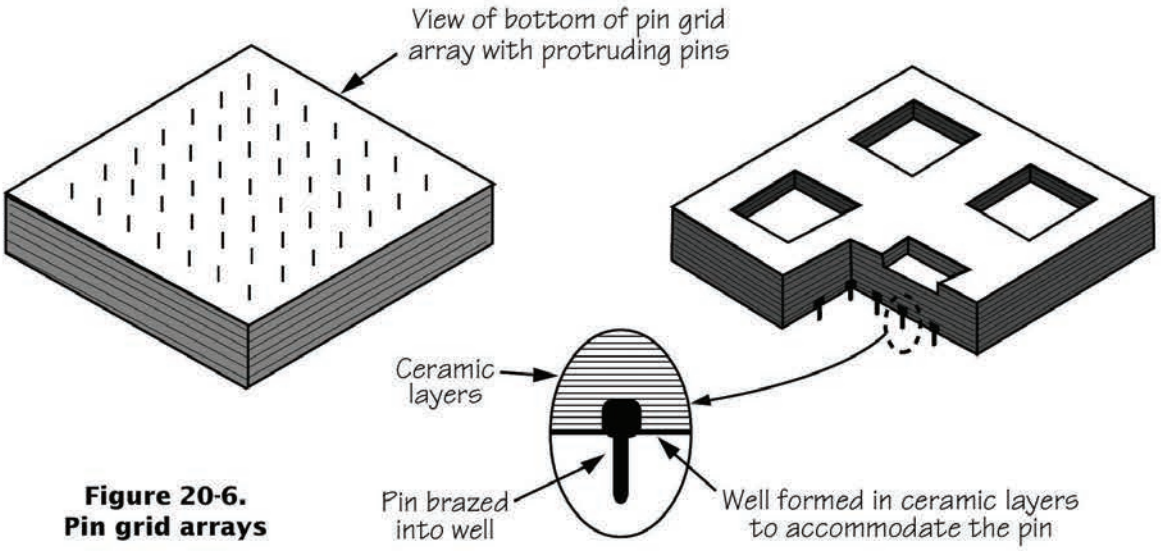


Figure 20-6. Pin grid arrays

The lead pitch (the spacing between adjacent leads) is typically in the order of 1.25 mm, and a module can contain more than a thousand such leads (although the average device contains rather less than this). After the leads have been inserted, the entire assembly is heated to approximately 1,000°C, which is sufficient to braze the leads to the layer of metal lining the wells. Finally, the leads are gold-plated to ensure that they will make good electrical connections when the module is eventually mounted on a circuit board.

Pad, Ball, and Column Grid Arrays

As an alternative to pin grid arrays, some modules are presented as *pad grid arrays (PGAs)*, in which the external connections are arranged as an array of

conducting pads on the base. In yet another alternative, which has become one of the premier packaging technologies currently in use, devices may be presented as *ball grid arrays (BGAs)*[12] or *column grid arrays (CGAs)*. These forms of package are similar to a pad grid array, but with the addition of small balls or columns of solder attached to the conducting pads (Figure 20-7).

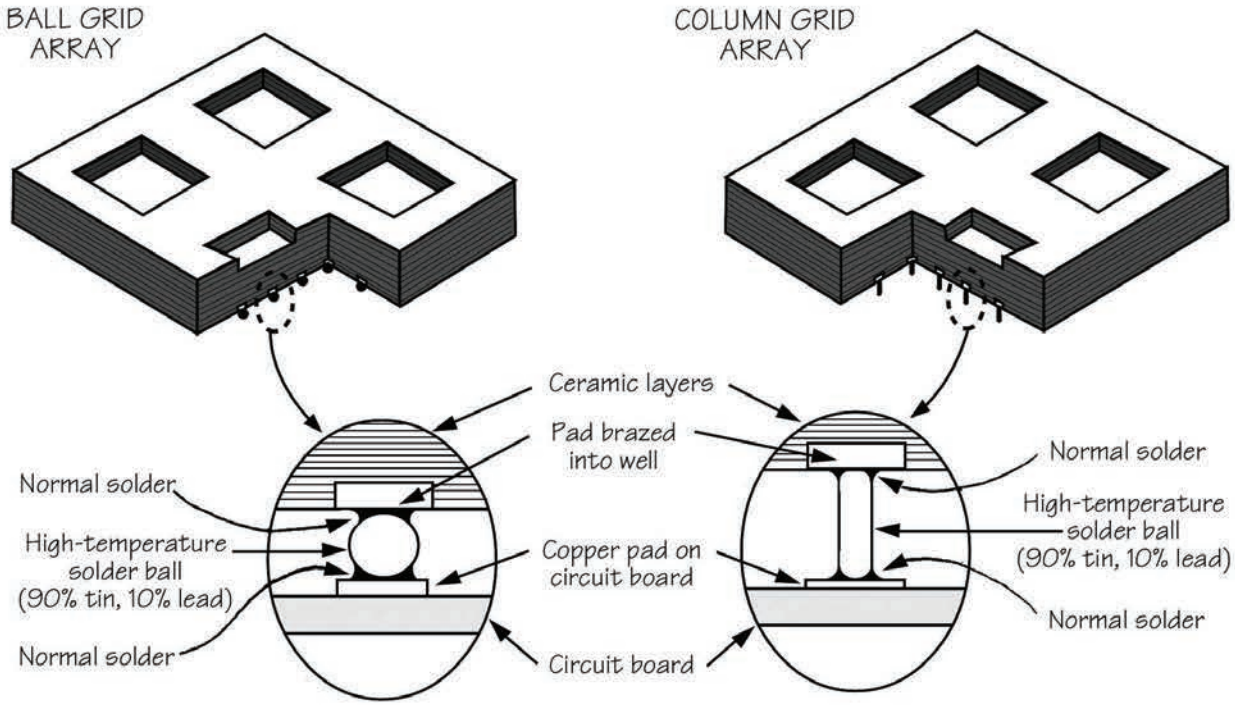


Figure 20-7. Ball grid arrays and column grid arrays

The balls (or columns) are made out of a high-temperature solder (90% tin and 10% lead), which does not melt during the soldering process. Instead, a standard low-temperature solder, referred to as *eutectic solder*, is used to connect the balls (or columns) to the pads. These packages are said to be self aligning because, when they are mounted on a circuit board and heated in a reflow oven, the surface tension of the eutectic solder causes the device to automatically align itself.

[12] Ball grid arrays may also be referred to as solder grid arrays, bump grid arrays, or land grid arrays.

Ball grid arrays are physically more robust than column grid arrays. However, column grid arrays are better able to accommodate any thermal strains caused by temperature mismatches between the multichip module and the circuit board.

Fuzz-Buttons

In yet another variant, which may be of use for specialist applications, small balls of fibrous gold known as *fuzz-buttons* are inserted between the pads on the base of the package and their corresponding pads on the board. When the package is forced against the board, the fuzz-buttons compress to form good electrical connections. One of the main advantages of the fuzz-button approach is that it allows broken devices to be quickly removed and replaced.

Populating the Die

However, we digress. Once the external leads have been connected, the module is ready to be populated with its die (Figure 20-8). In many ways the assembly processes used for multichip modules are similar to those used for hybrids. Bare die can be physically mounted on the substrate using adhesive, eutectic, or flipped-chip techniques, and electrically connected using wire bonding, tape automated bonding, or solder bumping methodologies. After the die have been mounted, the wells are back-filled with nitrogen, and protective lids are firmly attached.

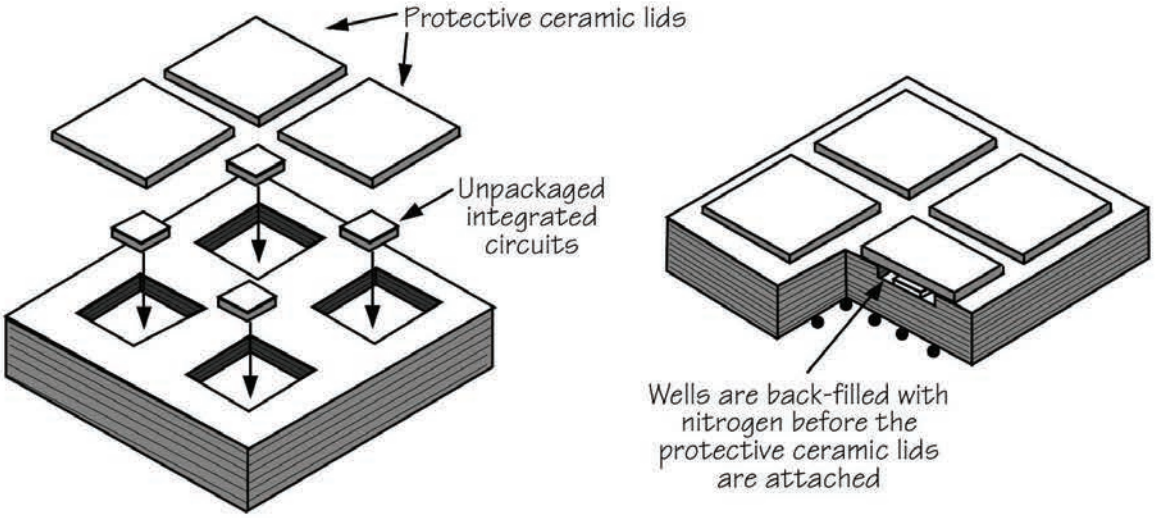


Figure 20-8. Populating the die

Equivalent Integrated Circuits

In previous discussions, it was noted that printed circuit boards may contain thousands of components, and hybrids may contain hundreds of components, but the multichip module described above contained only four integrated circuits. However, simply counting the components does not provide an appropriate basis for comparison. Although today's multichip modules typically contain relatively few devices, the actual logic represented by those devices can be huge. Thus, as opposed to *component density*, a more appropriate comparison is that of *functional density*.

There are a number of methods that may be used to compare these different interconnection strategies, including the average number of vias per square centimeter and the average number of tracks per layer per square centimeter. However, one of the most popular techniques is based on the concept of an *equivalent integrated circuit*.[13] The definition of an equivalent integrated circuit is one that contains a certain number of pins or leads;[14] a commonly used value is 20.

First, all of the leads associated with the real components attached to the substrate in question are added together and divided by the number of leads on the hypothetical equivalent integrated circuit. Thus, the result returned by this operation is in units of equivalent integrated circuits. For example, 300 leads divided by 20 equals 15 equivalent integrated circuits. This value is then divided by the area of the substrate to determine the number of equivalent integrated circuits per square centimeter.

To illustrate what this means in real terms, the author performed these calculations on a real printed circuit board, a real hybrid, and a real multichip module that he happened to find lying around his office. The printed circuit board was 27.5 cm × 20.0 cm and contained 627 components with a total of 1,736 leads; the hybrid was 2.5 cm × 5.0 cm and contained 36 components with a total of 352 leads; while the multichip module was 5.0 cm × 5.0 cm and contained four die each with 300 leads giving a total of 1,200 leads. By some

---

[13] This is similar in principle to the concept of an *equivalent gate*, which was introduced in Chapter 14.

[14] As usual, the actual definition of what constitutes an equivalent integrated circuit depends on whom one is talking to.

strange quirk of fate, this multichip module was the same as the one illustrated in the figures above. Performing the equivalent integrated circuit calculations returned the results shown in Table 20-1.

Table 20-1. Comparison based on equivalent integrated circuits

	Area of substrate (cm^2)	*# of real component leads*	*# of equivalent integrated circuits*	*Equivalent integrated circuits per cm^2*
Printed circuit board	550	1,736	86.8	0.16
Hybrid	12.5	352	17.6	1.41
Multichip module	25	1,200	60	2.40

Although this sample (one each of a circuit board, hybrid, and multichip module) is statistically meaningless, it does serve to illustrate the trend, which is that multichip modules offer the most sophisticated interconnection technology currently available.

The Mind Boggles

In addition to the substrates introduced above, some multichip modules use more esoteric materials such as quartz, garnet, and sapphire.[15] These modules are targeted towards specialized applications. For example, the ultra-high frequency capabilities of sapphire are of particular interest in the microwave arena. Additionally, some multichip modules use a mixture of substrates: for example, ceramic substrates with areas of embedded or deposited silicon or diamond film. Other devices may use a mixture of tracking techniques: for example, thin-film processes for surface signals combined with thick-film processes for inner layers, or vice versa.

In addition to complex materials selection and process problems, designers of multichip modules are faced with a number of other challenges. Parasitic and thermal management are key issues due to the high interconnection density and high circuit speeds. Thus, it is simply not possible to split a multichip module design into an interconnection problem, a thermal management problem, and a

[15] See also the discussions on diamond substrates in Chapter 21.

packaging problem. Each multichip module presents a single mammoth problem in concurrent design and typically requires a unique solution. The result is that working with multichip modules can be an expensive hobby in terms of tooling requirements and long design times.

The multichip module designer's task has migrated from one having primarily electronic concerns to one incorporating components of materials science, chemistry, and packaging. In addition to basic considerations—such as selecting the packaging materials and determining the optimal connection strategy (pin grid array versus pad grid array versus ball grid array)—the designer has to consider a variety of other packaging problems, including lead junction temperatures and thermal conduction and convection characteristics. The magnitude of the task is daunting, but if it was easy then everybody would be doing it, and you can slap me on the head with a kipper[16] if it doesn't make life exciting!

[16] A kipper is a smoked and salted fish (especially tasty when brushed with butter, toasted, and served as a breakfast dish with strong, hot English tea).

Chapter 21

Alternative and Future Technologies

Electronics is one of the most exciting and innovative disciplines around, with evolutionary and revolutionary ideas appearing on almost a daily basis. Some of these ideas skulk around at the edges of the party, but never really look you in the eye or take the trouble to formally introduce themselves; some surface for a short time and then disappear forever into the twilight zone from whence they came; some tenaciously manifest themselves in mutated forms on a seasonal basis; and some leap out as if from nowhere with a fanfare of trumpets and join the mainstream so quickly that before you know it they seem like old friends.

In this, the penultimate chapter, we introduce a smorgasbord of technologies, many of which have only recently become commercially available or are on the cutting-edge of research and development. Even the most outrageous topics presented below have undergone experimental verification, but nature is a harsh mistress and natural selection will take its toll on all but the fittest. Although some of the following may seem to be a little esoteric at first, it is important to remember that a good engineer can easily believe three impossible things before breakfast. Also remember that the nay sayers proclaimed that it was impossible for bumble bees to fly (although they obviously could), that man would never reach the moon, and that I would never finish this book![1]

Reconfigurable Hardware and Interconnect

The term *hardware* is generally understood to refer to any of the physical portions constituting an electronic system, including nuts and bolds, connectors, components, circuit boards, power supplies, cabinets, and monitors (see sidebar). However, this discussion concentrates on the subset of hardware consisting of circuit boards, electronic components, and interconnect.

---

[1] Ha! I pluck my chest hairs threateningly in their general direction.

The term *software* refers to programs, or sequences of instructions, that are executed by hardware. Additionally, *firmware* refers to programs that are hard-coded into non-volatile memory devices, while *vaporware* refers to either hardware or software that exists only in the minds of the people who are trying to sell them to you. Last but not least, *wetware* may refer either to human brains or programs that are very new and are not yet as robust as one might hope.

Unfortunately, the phrase *reconfigurable hardware* is akin to the phrase "stretch-resistant socks"—they both mean different things to different people. To the young and innocent, "stretch-resistant" would tend to imply a pair of socks that will not stretch. But, as those of us who are older, wiser, and a little sadder know, "stretch-resistant" actually refers to socks that will indeed stretch—they just do their best to resist it for a while! Similarly, the term *reconfigurable* is subject to myriad diverse interpretations depending on the observer's point of view.

As a starting point, *reconfigurable* hardware refers to an electronic product whose function can be customized to a specific system or application. There are obvious benefits to making one product (that can be customized) many times, as opposed to making many application-specific products once. The problem is, the perception of what is implied by "reconfigurable" is a moving target which evolves over time as new techniques and technologies become available. Throughout most of the 1980s, the most sophisticated level of customization was displayed by products based on *programmable logic devices (PLDs)* such as PROMs, PLAs, and PALs,[2] or variants of read-only memory such as PROMs,[3] EPROMs, and E^{2}PROMs.[4] Products using these devices were usually targeted at a particular application and then focused towards a specific implementation. By comparison, the advent of *field-programmable gate arrays (FPGAs)* in the mid-1980s and early 1990s opened the door to products that could be almost totally customized for diverse applications.

As you have hopefully surmised by now, *Bebop to the Boolean Boogie* does not simply describe "where we are," but also attempts to show "how we got here." However, the pace of technological development since the birth of the

---

[2] Programmable logic devices were introduced in Chapter 16.

[3] PROMs may be regarded as being programmable logic devices or as memory devices depending on the applications for which they are being used.

[4] Memory devices were introduced in Chapters 15 and 16.

integrated circuit poses particular problems in the case of reconfigurable logic, because there is no consistent terminology that has survived the ravages of time. To permit the presentation of a consistent view, the solution offered by Bebop is to adopt the following terminology:[5]

Configurable Hardware:	A product whose function may be customized once or a very few times.
Reconfigurable Hardware:	A product whose function may be customized many times.
Remotely Reconfigurable Hardware:	A product whose function may be customized remotely, by telephone or radio, while remaining resident in the system.
Dynamically Reconfigurable Hardware:	A product whose function may be customized "on-the-fly" while remaining resident in the system.
Virtual Hardware:	An extension of dynamically reconfigurable hardware.

And so, armed with our trusty definitions, it only remains for us to gird up our loins and delve into the ramifications of hardware whose very reason for being can be transmogrified by an idle whim.

Configurable Hardware

A classic example of a product whose function may be customized only once is provided by a car radio, of which there may be several versions: a cheap, no-frills little number for the cost-conscious buyer, a mid-range model for the young man about town, and an ultra-flash, no-holds-barred version for the rising business executive (the sort who delights in thrusting his gold Rolex under your nose). But, little do they know that it is not unusual for all of these variations to be constructed on identical circuit boards, which can be configured by adding or removing certain components and modifying certain

[5] The terminology presented here differs to some extent from current industry standards, which only tend to reflect the technology of the day (or flavor of the month).

switches or jumpers. In fact, on some occasions, the only major difference between the different models is the quality of their cases and the number of buttons and dials that they support. From the viewpoint of the manufacturer, the circuit boards used in these radios are *configurable*; but from the perspective of the user, their function is cast in stone (Figure 21-1).

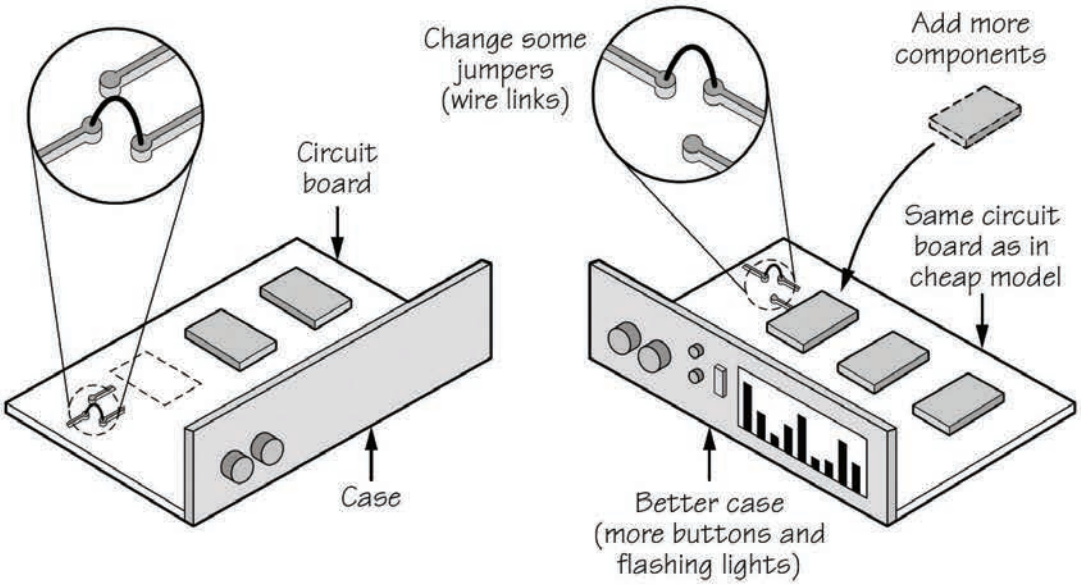


Figure 21-1.
Configurable hardware: car radios

A similar example that may be a little closer to home revolves around digital wristwatches. One down-side of the electronics era is the inevitability of being cornered by someone who insists on regaling you with the 1,001 details you didn't want to know about his new digital watch that can simultaneously display the current time in Tokyo, Paris, London, New York, and Moscow, play sixteen immediately annoying tunes, and has a calculator rivaling the control panel of the space shuttle thrown in for good measure. By some strange quirk of fate, this almost invariably occurs on those days when the timepiece you are sporting arrived in the form of a free gift at the bottom of a box of cornflakes. But there is no need to lower your head in shame, because it is not beyond the bounds of possibility that both of these instruments contain identical integrated

circuits! In the case of the simpler model, a hard-wired voltage level applied to one of the device's pins instructs it to pretend to be "cheap and cheerful." Once again, the major difference between the two models is the quality of their cases and . . . the price tag.

A final example of configurable products that is too good not to share revolves around a number of well-known computer manufacturers who used to offer a choice between the fast, deluxe, and expensive version of a machine, or the slower, somewhat cheaper model. However, unbeknownst to the innocent purchaser, the only difference between the two systems was a simple switch on the main circuit board. Depending on the position of this switch, the system's clock either ran at full speed or was "slugged" to half its normal operating frequency. When the owners of one of the slower systems decided that they just had to upgrade, large amounts of money would change hands and, after appropriate sacrifices had been offered to the Gods, the computer engineer would (eventually) arrive. After making a ceremony of preparing to exchange the circuit boards, the engineer would suddenly leap to his feet, point excitedly out of the window, and cry "Good grief! What's that?" Then, while everybody's backs were turned, he would flick the switch and have the board halfway back into the system before anybody knew what was happening![6]

Reconfigurable Hardware

Hardware that is simply configurable is obviously limited, because everything that the product can do has to be designed into its base configuration, which has to encompass all possible variants. One technique for producing a product whose function may be extended beyond its original design objectives is to base that product on devices that can be reprogrammed. For example, a PROM could be employed in the role of a hardware truth table (Figure 21-2).

Similarly, a PLA could be used to implement a state machine that performs a certain sequence of operations, or, based on the inputs presented to it, a PAL device could be used to implement Boolean equations and to generate appropriate outputs. In all of these cases, the functions of the truth table, state machine, or Boolean equations could be modified by simply exchanging the programmable device with an upgraded version.

---

[6] And there are those who would say that electronics is dull and boring!

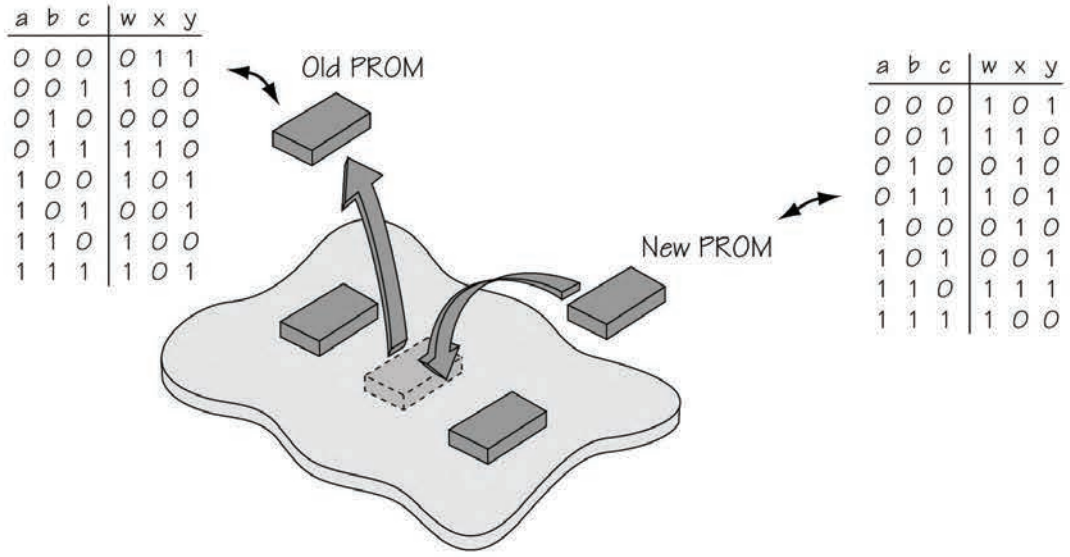


Figure 21-2. Reconfigurable hardware: PROM as a hardware truth table

Another technique would be to use non-volatile memory devices to store firmware programs for use by a microprocessor or a microcontroller. An example could be a set of instructions used by a microprocessor to play a tune such as the National Anthem on a musical door chime. Different versions of the PROM could be used to allow the product to be marketed in different countries (hopefully with the correct anthem). Additionally, in the case of those countries that count revolution as a national sport, the product could be easily reconfigured to reflect the "*tune of the day*."

The non-volatile memory examples introduced above could use PROM, EPROM, E²PROM, or FLASH devices; similarly, the programmable logic device examples could employ PLD, EPLD, E²PLD, or FLASH-PLD components. In all of these cases, from a board-level perspective, the board itself would be classified as *reconfigurable*. However, from a device-level viewpoint, PROMs (which can only be programmed a single time) would fall into the category of *configurable*, while their more sophisticated cousins—EPROM, E²PROM, and FLASH—would be categorized as *reconfigurable* (similarly in the case of PLDs versus EPLDs, E²PLDs, and FLASH-PLDs). Additionally, E²-based and FLASH-based components may be referred to as *in-system programmable (ISP)*, because they can be reprogrammed while remaining resident on the circuit board.

Dynamically Reconfigurable Hardware

The advent of SRAM-based FPGAs presented a new capability to the electronics fraternity: *dynamically reconfigurable hardware*, which refers to designs that can be reconfigured "on-the-fly."

FPGAs contain a large number of diverse logic gates and registers, which can be connected together in widely different ways to achieve a desired function. SRAM-based variants allow new configuration data to be downloaded into the device by the main system in a fraction of a second. In the case of these devices, a few of the external pins are dedicated to the task of loading the data, including an *enable* control and *clock* and *data* inputs. When the *enable* control is placed in its active state, edges on the *clock* are used to load the device's SRAM with a stream of *0*s and 1s, which is presented to the serial *data* input. Although all of the logic gates and SRAM cells are created on the surface of a single piece of silicon substrate, it may be useful to visualize the device as comprising two distinct strata: the logic gates and the programmable SRAM "switches" (Figure 21-3).

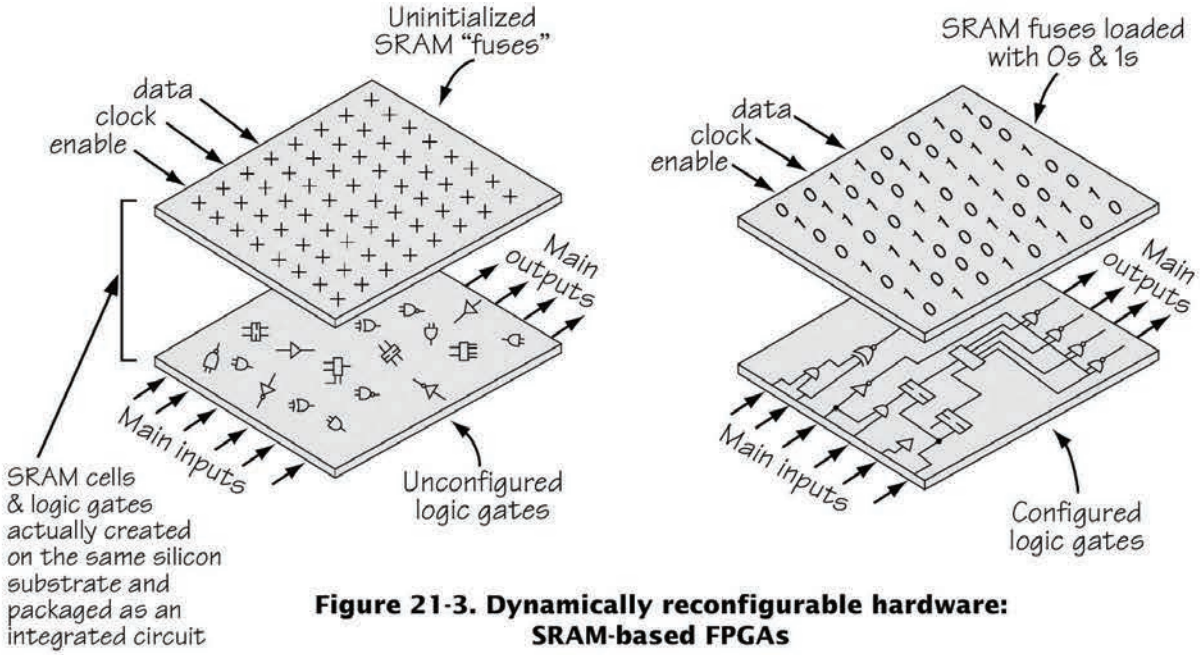


Figure 21-3. Dynamically reconfigurable hardware: SRAM-based FPGAs

The versatility of these devices opens the floodgates to a wealth of possibilities. For example, the creation of circuit boards whose interfaces can be

configured to meet a variety of communications protocols, or devices acting as *digital signal processors (DSPs)*, whose core algorithms can be modified to process data in a variety of ways. As alternative protocols become available or improved algorithms are invented, the patterns that are used to configure the FPGAs can be modified to take full advantage of these new developments.

However, the true power of these devices, which are referred to as *in-circuit reconfigurable (ICR)*, resides in their ability to be reconfigured on-the-fly. For example, when a system is first turned on, it might configure all of the FPGAs to perform diagnostic functions, both on themselves and on the circuit board. Once the diagnostic checks have been completed, the system can dynamically reconfigure the FPGAs to fulfill the main function of the design.

Another example is illustrated by the Tomahawk cruise missile, which uses one navigational technique to control itself while flying over water and another while soaring over land. When the Tomahawk crosses the boundary from water to land or vice versa, it causes its on-board FPGAs to be dynamically reconfigured, changing from water-navigation mode to land-navigation mode in a fraction of a second.[7]

Dynamically Reconfigurable Interconnect

Wonderful as all of the above is, even these techniques only scratch the surface of the possibilities offered by today's emerging technologies. Designers would ideally like to create board-level products that can be reconfigured to perform radically improved, or completely different functions from the ones that they were originally designed for. The solution is to be able to dynamically configure the board-level connections between devices.

A breed of devices offer just this capability: *field-programmable interconnect devices (FPIDs)*, which may also be known as *field-programmable interconnect chips (FPICs)*.[8] These devices, which are used to connect logic devices together, can be dynamically reconfigured in the same way as standard SRAM-based FPGAs. Due to the fact that each FPID may have around 1,000 pins, only a few such devices are typically required on a circuit board (Figure 21-4).

---

[7] Of course, some of us might take the view that it is inherently unwise to have an armed cruise missile rocketing around the sky in a mindless state while it reprograms its own brain!

[8] FPIC is a trademark of Aptix Corporation.

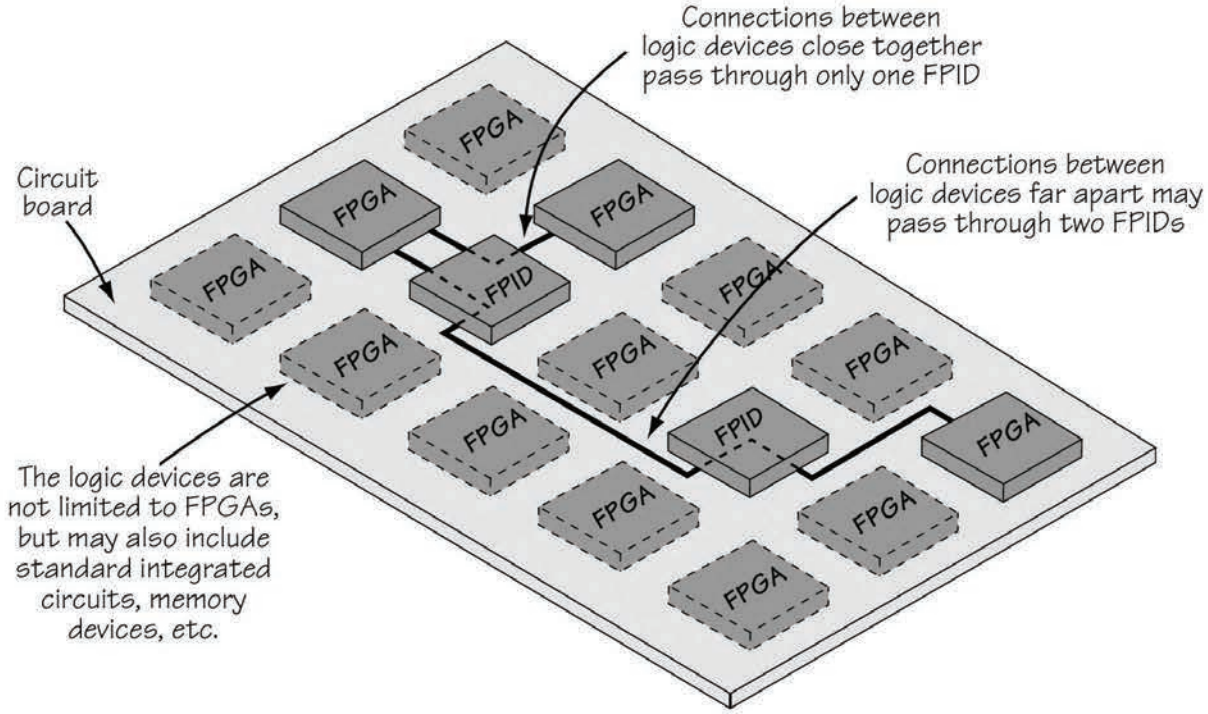


Figure 21-4. Dynamically reconfigurable interconnect: SRAM-based FPIDs

In fact, the concepts discussed here are not limited to board-level implementations. Any of the technologies discussed thus far may also be implemented in hybrids and multichip modules. Additionally, ultra-large-scale ASIC and system-on-chip (SoC) devices are becoming available; these little rapscallions can combine microprocessor cores, blocks of memory, and communication functions with embedded FPGA-style and FPID-style functions.

Virtual Hardware

The main limitation with the majority of SRAM-based FPGAs is that it is necessary to load the entire device. Apart from anything else, it is usually necessary to halt the operation of the entire circuit board while these devices are being reconfigured. Additionally, the contents of any registers in the FPGAs are irretrievably lost during the process.

To address these issues, a new generation of FPGAs was introduced around the beginning of 1994. In addition to supporting the dynamic reconfiguration of selected portions of the internal logic, these devices also feature:

a) No disruption to the device's inputs and outputs.

b) No disruption to the system-level clocking.

c) The continued operation of any portions of the device that are not undergoing reconfiguration.

d) No disruption to the contents of internal registers during reconfiguration, even in the area being reconfigured.

The latter point is of particular interest, because it allows one instantiation of a function to hand over data to the next function. For example, a group of registers may initially be configured to act as a binary counter. Then, at some time determined by the main system, the same registers may be reconfigured to operate as a linear feedback shift register (LFSR),[9] whose seed value is determined by the final contents of the counter before it is reconfigured.

Although these devices are evolutionary in terms of technology, they are revolutionary in terms of the potential they offer. To reflect their new capabilities, appellations such as *virtual hardware*, *adaptive hardware*, and *Cache Logic*[10] have emerged.

The phrase *virtual* hardware is derived from its software equivalent, *virtual memory*, and both are used to imply something that is not really there. In the case of virtual memory, the computer's operating system pretends that it has access to more memory than is actually available. For example, a program running on the computer may require 500 megabytes to store its data, but the computer may have only 128 megabytes of memory available. To get around this problem, whenever the program attempts to access a memory location that does not physically exist, the operating system performs a sleight-of-hand and exchanges some of the contents in the memory with data on the hard disk. Although this practice, known as *swapping*, tends to slow things down, it does allow the program to perform its task without having to wait while someone runs down to the store to buy some more memory chips.

Similarly, the phrase *Cache Logic* is derived from its similarity to the concept of *Cache Memory*, in which high-speed, expensive SRAM is used to store active

---

[9] Linear feedback shift registers (LFSRs) are introduced in detail in Appendix F.

[10] Cache Logic is a trademark of Atmel Corporation, San Jose, CA, USA.