# 7

# On-chip Power Distribution Networks

The impedance characteristics of a power distribution system are analyzed in the previous chapter based on a one-dimensional circuit model. While useful for understanding the principles of the overall operation of a power distribution system, a one-dimensional model is not useful in describing the distribution of power and ground across a circuit die. The size of an integrated circuit is usually considerably greater than the wavelength of the signals in the power distribution network. Furthermore, the power consumption of on-chip circuitry (and, consequently, the current drawn from the power distribution network) varies across the die area. The voltage across the on-chip power and ground distribution networks is therefore non-uniform. It is therefore necessary to consider the two-dimensional structure of the on-chip power distribution network to ensure that target performance characteristics of a power distribution system are satisfied. The on-chip power distribution network should also be considered in the context of a die-package system as the properties of the die-package interface significantly affect the constraints imposed on the electrical characteristics of the on-chip power distribution network.

The objectives of this chapter is to describe the structure of an on-chip power distribution network as well as review related tradeoffs. Various structural styles of on-chip power distribution networks are described in Section 7.1. The influence of the electrical characteristics of the die-package interface on the on-chip power and ground distribution is analyzed in Section 7.2. The influence of the on-chip power distribution network on the integrity of the on-chip signals is discussed in Section 7.3. The chapter concludes with a summary.

## 7.1 Styles of on-chip power distribution networks

Several topological structures are typically used in the design of on-chip power distribution networks. The power network structures range from completely irregular, essentially *ad hoc*, structures, as in routed power distribution networks, to highly regular and uniform structures, as in gridded power networks and power planes. These topologies and other basic types of power distribution networks are described in Section 7.1.1. Design approaches to improve the impedance characteristics of on-chip power distribution networks are presented in Section 7.1.2. The evolution of on-chip power distribution networks in the family of Alpha microprocessors is presented in Section 7.1.3.

### 7.1.1 Basic structure of on-chip power distribution networks

Several structural types of on-chip power distribution networks are described in this section. Different parts of an on-chip network can be of different types, forming a hybrid network.

**Routed networks**

In routed power distribution networks, the local circuit blocks are connected with dedicated routed power trunks to the power I/O pads along the periphery of the die [202], as shown in Fig. 7.1. A power mesh is typically used to distribute the power and ground within a circuit block. The primary advantage of routed networks is the efficient use of interconnect resources, favoring this design approach in circuits with limited interconnect resources. The principal drawback of this topology is the relatively low redundancy of the power network. All of the current supplied to any circuit block is delivered through only a few power trunks. The failure of a single segment in a power distribution network jeopardizes the integrity of the power supply voltage levels in several circuit blocks and, consequently, the correct operation of the entire circuit. Routed power distribution networks are predominantly used in low power, low cost integrated circuits with limited interconnect resources.

**Mesh networks**

Improved robustness and reliability are offered by power and ground mesh networks. In mesh structured power distribution networks, parallel power and ground lines in the upper metal layers span an entire circuit or a specific circuit block. These lines are relatively thick and

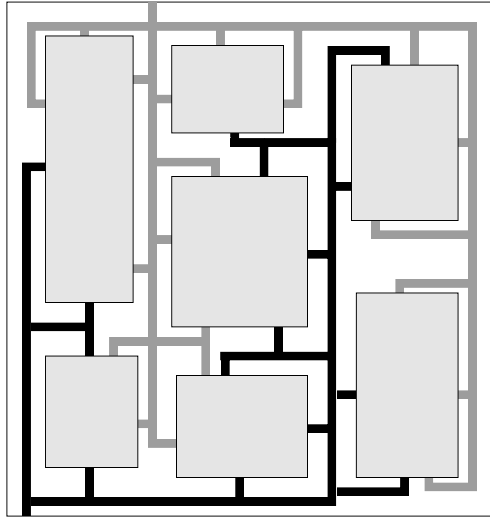

**Fig. 7.1.** Routed power and ground distribution networks. The on-chip circuit blocks are connected to the I/O power terminals with dedicated power (black) and ground (gray) trunks. The structure of the power distribution networks within the individual circuit blocks is not shown.

wide, globally distributing the power current. The lines are interconnected by relatively short orthogonal straps in the lower metal layer, forming an irregular mesh, as shown in Fig. 7.2. The power and ground lines in the lower metal layer distribute current in the direction orthogonal to the upper metal layer lines and facilitate the connection of on-chip circuits to the global power distribution network. Mesh networks are used to distribute power in relatively low power circuits with limited interconnect resources [203]. It is often the case in semiconductor processes with only three or four metal layers that a regular power distribution grid in the upper two metal layers cannot be utilized due to an insufficient amount of metal resources and an ensuing large number of routing conflicts. Mesh networks are also used to distribute power within individual circuit blocks.

**Grid structured networks**

Grid structured power distribution networks, shown in Fig. 7.3, are commonly used in high complexity, high performance integrated circuits. Each layer of a power distribution grid consists of many equidistantly spaced lines of equal width. The direction of the power and ground lines within each layer is orthogonal to the direction of the

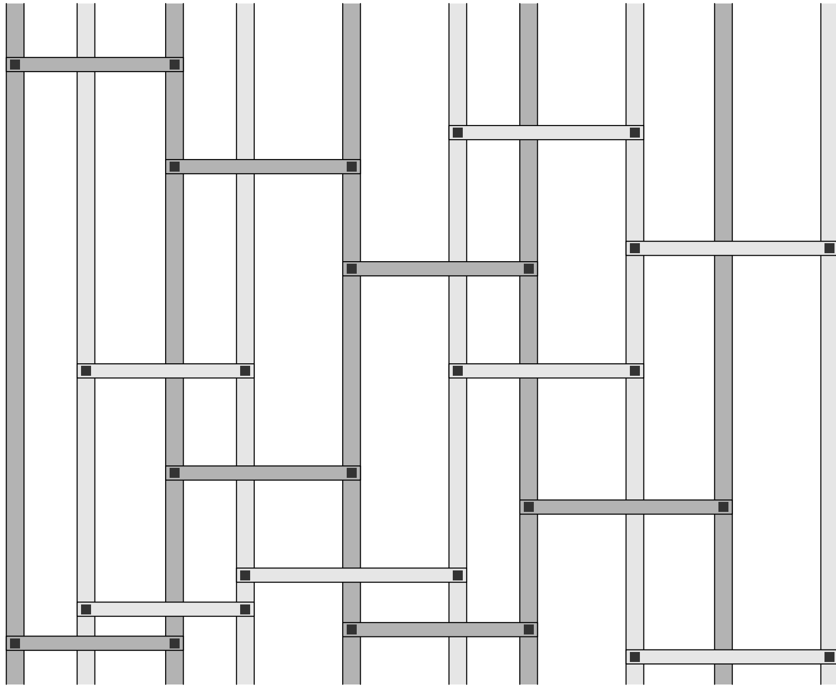

**Fig. 7.2.** A mesh structured power distribution network. Power (dark gray) and ground (light gray) lines in the vertically routed metal layer span an entire die or circuit block. These lines are connected by short straps of horizontally routed metal to form a mesh. The lines and straps are connected by vias (the dark squares).

lines in the adjacent layers. The power and ground lines are typically interdigitated within each layer. Each power and ground line is connected by vias to other power and ground lines, respectively, in the adjacent layers at the overlap sites. In a typical integrated circuit, the lower the metal layer, the smaller the width and pitch of the lines. The coarse pitch of the upper metal layer improves the utilization of the metal resources, conforming to the pitch of the I/O pads of the package, while the fine pitch of the lower grid layers brings the power and ground supplies in close proximity to each on-chip circuit, facilitating the connection of these circuits to power and ground.

Power distribution grids are significantly more robust than routed distribution networks. Multiple redundant current paths exist between the power terminals of each load circuit and the power supply pads. Due to this property, the power supply integrity is less sensitive to changes in the power current requirements of the individual circuit blocks. The

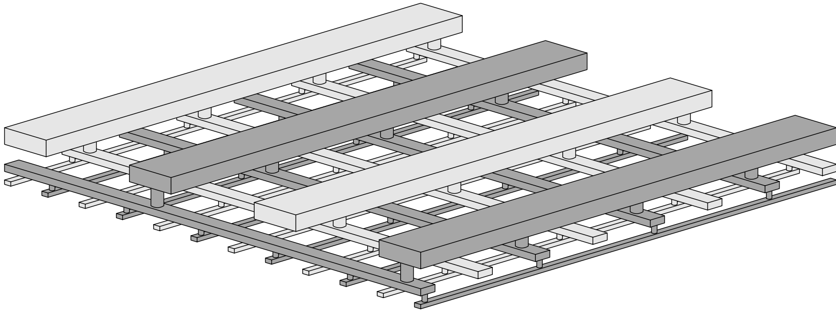

**Fig. 7.3.** A multi-layer power distribution grid. The ground lines are light gray, the power lines are dark gray. The pitch, width, and thickness of the lines are smaller in the lower grid layers than in the upper layers.

failure of any single segment of the grid is not critical to delivering power to any circuit block. An additional advantage of power distribution grids is the enhanced integrity of the on-chip data signals due to the capacitive and inductive shielding properties of the power and ground lines. These advantages of power distribution grids, however, are achieved at the cost of a significant share of on-chip interconnect resources. It is not uncommon to use from 20% to 40% of the metal resources to build a high density power distribution grid in modern high performance microprocessors [24], [25], [202].

### Power and ground planes

Dedicated power and ground planes, shown in Fig. 7.4, have also been used in the design of power distribution networks [204]. In this scheme, an entire metal layer is used to distribute power current across a die, as shown in Fig. 7.4. The signal lines above and below the power/-ground plane are connected with vias through the holes in the plane. Power planes also provide a close current return path for the surrounding signal lines, reducing the inductance of the signal lines and therefore the signal-to-signal coupling. This advantage, however, diminishes as the interconnect aspect ratios are gradually increased with technology scaling. While power planes provide a low impedance path for the power current and are highly robust, the interconnect overhead is typically prohibitively large, as entire metal layers are unavailable for signal routing.

### Cascaded power/ground rings

A novel topology for on-chip power distribution networks, called a

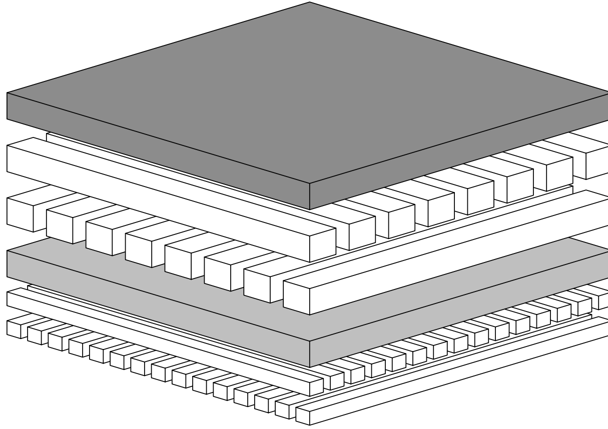

**Fig. 7.4.** On-chip power distribution scheme using power and ground planes. Two entire metal layers are dedicated to the distribution of power (dark gray layer) and ground (light gray layer).

"cascaded power/ground ring," has been proposed by Lao and Krusius [205] for integrated circuits with peripheral I/O. This approach is schematically illustrated in Fig. 7.5. The power and ground lines are routed from the power supply pads at the periphery of the die toward the die center. The power and ground lines at the periphery of the die, where the power current density is the greatest, are placed on the thick topmost metal layers with the highest current capacity. As the power current decreases toward the die center, the power and ground interconnect is gradually transfered to the thinner lower metal layers.

**Hybrid-structured networks**

Note that the boundaries between these network topologies are not well defined. A routed network with a large number of links looks quite similar to a meshed network, which, in turn, resembles a grid structure if the number of "strapping" links is large. The terms "mesh" and "grid" are often used interchangeably in the literature.

Furthermore, the structure of a power distribution network in a complex circuit often comprises a variety of styles. For example, the global power distribution can be performed through a routed network, while a meshed network is used for the local power distribution. Or the global power distribution network is structured as a regular grid, while the local power network is structured as a mesh network within one circuit block and a routed network within another circuit block. The common

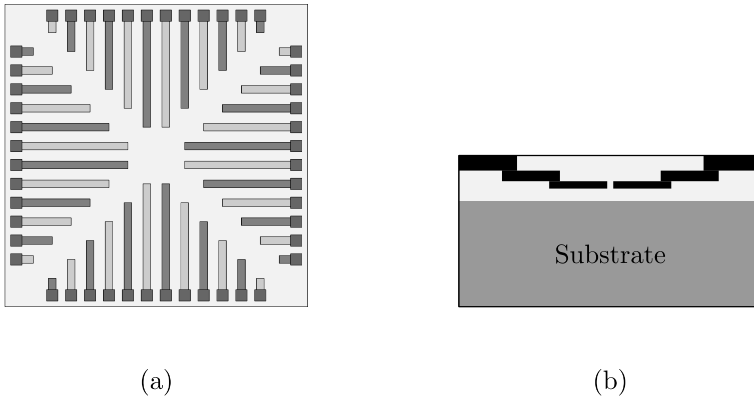


**Fig. 7.5.** Power distribution network structured as a cascaded power/ground ring; (a) cross-sectional view and (b) top view.

style of a *global* power distribution network has, however, evolved from a routed network to a global power grid, as the power requirements of integrated circuits have gradually increased with technology scaling.

### 7.1.2 Improving the impedance characteristics of on-chip power distribution networks

The on-chip power and ground interconnect carries current from the I/O pads to the on-chip capacitors and from the on-chip capacitors to the switching circuits, acting as a load to the power distribution network. The flow of the power current through the on-chip power distribution network produces a power supply noise proportional to the network impedance, $Z = R + j\omega L$. The primary design objective is to ensure that the resistance and inductance of a power distribution system is sufficiently small so as to satisfy a target noise margin.

Several techniques have been employed to reduce the parasitic impedance of on-chip power distribution networks. The larger width and smaller pitch of the power and ground lines increase the metal area of the power distribution network, decreasing the network resistance. The resistance is effectively lowered by increasing the area of the power lines in the upper metal layers since these layers have a low sheet resistance. It is not uncommon to allocate over half of the topmost metal layer for global power and ground distribution [24], [165].

The inductance of on-chip power and ground lines has traditionally been neglected because the overall inductance of a power distribution

network has been dominated by the parasitic inductance of the package pins, planes, vias, and bond wires. This situation is changing due to the increasing switching speed of integrated circuits [206], [207], the lower inductance of advanced flip chip packaging, and the higher on-chip decoupling capacitance which terminates the high frequency current paths. The requirement of achieving a low inductive impedance is in conflict with the requirement of a low resistance, as the use of wide lines to lower the resistance of a global power distribution network increases the network inductance. Replacing a few wide power and ground lines with multiple narrow interdigitated power and ground lines, as shown in Fig. 7.6, reduces the self inductance of the supply network [208], [209] but increases the resistance. The tradeoffs among the area, resistance, and inductance of on-chip power distribution grids are explored in greater detail in Chapter 11.

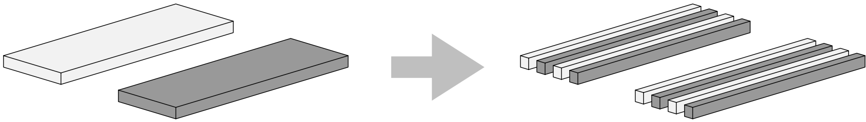

**Fig. 7.6.** Replacing wide power and ground lines (left) with multiple narrow interdigitated lines (right) reduces the inductance and characteristic impedance of the power distribution network.

### 7.1.3 Evolution of power distribution networks in Alpha microprocessors

The evolution of on-chip power distribution networks in high speed, high complexity integrated circuits is well illustrated by several generations of Digital Equipment Corporation Alpha microprocessors, as described by Gronowski, Bowhill, and Preston [165]. Supporting the reliable and efficient distribution of rising power currents have required adapting the structure of these on-chip power distribution networks, utilizing a greater amount of on-chip metal resources.

**Alpha 21064**

The Alpha 21064, the first microprocessor in the Alpha family, is implemented in a 0.7 $\mu$m CMOS process in 1992. The Alpha 21064 consumes 30 watts of power at 3.3 volts, resulting in nine amperes of average power current. Distributing this current across a 16.8 mm $\times$ 13.9 mm

die would not have been possible in an existing two metal layer 0.75 $\mu$m CMOS process. An additional thick metal layer was added to the process, which is used primarily for the power and clock distribution networks. The power and ground lines in the third metal layer are alternated. All of the power lines are interconnected at one edge of the die with a perpendicular line in metal level three, and all of the ground lines are interconnected at the opposite edge of the die, as shown in Fig. 7.7. The resulting comb-like power and ground global distribution networks are interdigitated. The parallel lines of the global power and ground distribution networks are strapped with the power and ground lines of the second metal layer, forming a mesh-structured power distribution network.

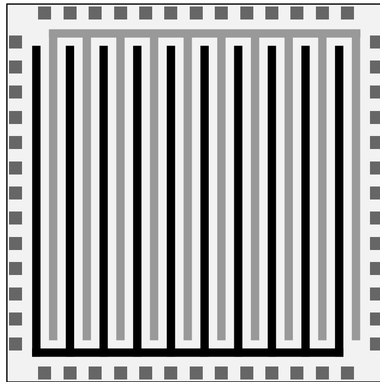

**Fig. 7.7.** Global power distribution network in Alpha 21064 microprocessor.

**Alpha 21164**

The second generation microprocessor in the series, the Alpha 21164, appeared in 1995 and was manufactured in a 0.5 $\mu$m CMOS process. The Alpha 21164 nearly doubled the power current requirements to 15 amperes, dissipating 50 watts from a 3.3 volt power supply. The type of power distribution network used in the previous generation could not support these requirements. An additional forth metal layer was therefore added to a new 0.5 $\mu$m CMOS process. The power and ground lines in the forth metal layer are routed orthogonally to the lines in the third layer, forming a two layer global power distribution grid.

**Alpha 21264**

The Alpha 21264, the third generation of Alpha microprocessors, was introduced in 1998 in a 0.35 $\mu$m CMOS process. The Alpha 21264 consumes 72 watts from a 2.2 volt power supply, requiring 33 amperes of average power current distributed with reduced power noise margins. Utilization of conditional clocking techniques to reduce the power dissipation of the circuit increased the cycle-to-cycle variation in the power current to 25 amperes, exacerbating the overall power distribution problem [165]. The two layer global power distribution grid used in the 21164 could not provide the necessary power integrity characteristics in an integrated circuit with peripheral I/O. Therefore, two thick metal layers were added to the four layer process to allow the exclusive use of two metal layers as power and ground planes. More recent implementations of the Alpha 21264 microprocessor in newer process technologies utilize flip-chip packaging with a high density area array of I/O contacts. This approach obviates the use of on-chip metal planes for distributing power [210].

Power distribution grids are the design style of choice in most modern high performance integrated circuits [165], [204], [211], [212]. The focus of the material presented herein is therefore on on-chip power distribution grids.

## 7.2 Die-package interface

At high frequencies, the impedance of a power distribution system is determined by the impedance characteristics of the on-chip and package power distribution networks, as discussed in Chapter 5. On-chip decoupling is essential to maintain a low impedance power distribution network at the highest signal frequencies of interest, as discussed in Section 5.5. The required on-chip decoupling capacitance is determined by the frequency where the package decoupling capacitors become inefficient. This frequency, in turn, is determined by the inductive impedance of the current path between the package capacitance and the integrated circuit. The required minimum on-chip decoupling capacitance is proportional to this inductance, as expressed by (5.20) (or by the analogous constraints presented in Section 5.6). Minimizing this inductance achieves the target impedance characteristics of a power distribution network with the smallest on-chip decoupling capacitance.

Achieving a low impedance connection between the package capacitors and an integrated circuit is, however, difficult. Delivering power is only one of the multiple functions of an IC package, which also include connecting the I/O signals to the outside world, maintaining an acceptable thermal environment, providing mechanical support, and protecting the circuit from the environment. These package functions all compete for physical resources within a small volume in the immediate vicinity of the die. Complex tradeoffs among these package design goals are made in practice, often preventing the realization of a resonance-free die-to-package interface.

**Wire-bond packaging**

Maintaining a low impedance die-package interface is particularly challenging in wire-bonded integrated circuits. The self inductance of a wire bond connection typically ranges from 4 nH to 6 nH. The number of bond wires is limited by the perimeter of the die and the pitch of the wire bond connections. The total number of power and ground connections typically does not exceed several hundred. It is therefore difficult to decrease the inductance of the package capacitor connection to the die significantly below one to two nanohenrys. In wire-bond packages, the inductance of the current loop terminated by the package capacitors is not significantly smaller than the inductance of the current loop terminated by the board decoupling capacitors. Under these conditions, the package capacitors do not significantly improve the impedance characteristics of the power distribution system and therefore no appreciable gain in circuit speed is achieved [105], [213]. Decoupling a high inductive impedance at gigahertz frequencies typically requires an impractical amount of on-chip decoupling capacitance (and therefore die area), limiting the operational frequency of a wire bonded circuit.

Providing a low impedance connection between the off-chip decoupling capacitors and a wire-bonded integrated circuit requires special components and packaging solutions. For example, a so-called “closely attached capacitor” has been demonstrated to be effective for this purpose in wire-bonded circuits [214]. A thin flat capacitor is placed on the active side of an integrated circuit. The dimensions of the capacitor are slightly smaller than the die dimensions. The bonding pads of an integrated circuit and the edge of the capacitor are in close proximity, permitting a connection with a short bond wire, as illustrated in Fig. 7.8. The die-to-capacitor wires are several times shorter than the wire connecting the die to the package. The impedance between the

circuit and the off-chip decoupling capacitance is therefore significantly decreased.

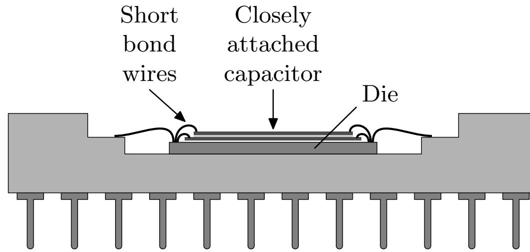


**Fig. 7.8.** Closely attached capacitor. Stacking a thin flat capacitor on top of a circuit die allows connecting the capacitor and the circuit with bond wires of much shorter length as compared to bond wires connecting the circuit to the package.

This wiring technique reduces the power switching noise in CMOS circuits by at least two to three times, as demonstrated by Hashemi *et al.* [214]. For example, attaching a 12 nF capacitor to a 32-bit microprocessor decreases the internal logic power supply noise from 900 mV to 150 mV, a sixfold improvement (the microprocessor is packaged in a 169-pin pin grid array package and operates at 20 MHz with a 5 volt power supply). Similarly, in a 3.3 volt operated bus interface circuit, a 12 nF closely attached capacitor lowers the internal power noise from 215 mV to 100 mV. A closely attached capacitor is used in the Alpha 21264 microprocessor, as the on-chip decoupling capacitance is insufficient to maintain adequate power integrity [165], [215].

**Flip-chip packaging**

The electrical characteristics of the die-package interface are significantly improved in flip-chip packages. Flip-chip bonding refers to attaching a die to a package with an array of solder balls (or bumps) typically 50 $\mu$m to 150 $\mu$m in diameter. In cost sensitive circuits, the ball connections, similar to wire-bond connections, can be restricted to the periphery of the die in order to reduce the interconnect complexity of the package. In high complexity, high speed integrated circuits, however, an area array flip-chip technology is typically used where solder ball connections are distributed across (almost) the entire area of the die. The inductance of a solder ball connection, typically from 0.1 nH to 0.5 nH, is much smaller than the 4 nH to 10 nH typical for a bond wire [103], [132], [216]. Area array flip-chip bonding also provides

a larger number of die to package connections as compared to wire bonding. Modern high performance microprocessors have thousands of flip-chip contacts dedicated to the power distribution network [25], [217], [218], [219]. A larger number of lower inductance power and ground connectors significantly decreases the overall inductance of the die to package connection.

Also important in integrated circuits with peripheral I/O, the power current is distributed on-chip across a significant distance: from the die edge to the die center, as shown in Fig. 7.9. In integrated circuits with a flip-chip area array of I/O bumps, the power current is distributed on-chip over a distance comparable to the size of the power pad pitch, as shown in Fig. 7.10. This distance is significantly smaller than half the die size. Flip-chip packaging with high density I/O, therefore, significantly reduces the effective resistance and inductance of the on-chip power distribution network, mitigating the resistive [220], [221] and inductive voltage drops.

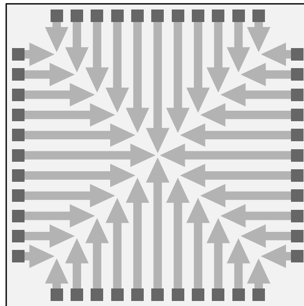

**Fig. 7.9.** Flow of power current in an integrated circuit with peripheral I/O. The power current is distributed on-chip across a significant distance: from the edge of the die to the die center.

Flip-chip packaging with high density I/O therefore decreases the area requirements of the on-chip power distribution network, improving the overall performance of a circuit [210], [222], [223], [224]. The dependence of the on-chip power voltage drop on the flip-chip I/O pad density and related power interconnect requirements are discussed in greater detail in Chapter 12.

Another advantage of area array flip-chip packaging is the possibility of placing the package decoupling capacitors in close physical proximity to the die, significantly enhancing the efficacy of the capacitors. A bank

of low parasitic inductance capacitors can be placed on the underside of the package immediately below the die, as shown in Fig. 7.11. The separation between the capacitors and the on-chip circuitry is reduced to 1 mm to 2 mm, minimizing the area of the current loop and associated inductance. The parasitic inductance of a package decoupling capacitor with the package vias and solder bumps connecting the capacitor to the die can be reduced well below 1 nH [213], enhancing the capacitor efficiency at high frequencies. Placing a package decoupling capacitor immediately below the circuit region with the greatest power current requirements further improves the efficiency of the decoupling capacitors of the package [225].

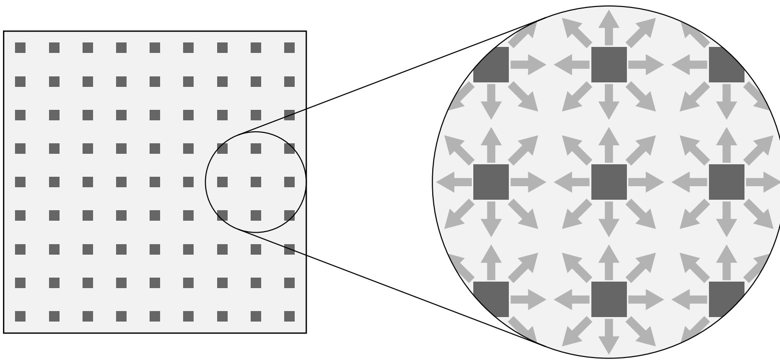

**Fig. 7.10.** Flow of power current in an integrated circuit with flip-chip I/O. The power current is distributed on-chip over a distance comparable to the pitch of the I/O pads.

Overall, flip-chip packaging significantly decreases the impedance between the integrated circuit and the package decoupling capacitors, relaxing the constraints on the resistance of the on-chip power distribution network and on-chip decoupling capacitors. Flip-chip packaging therefore can significantly improve the power supply integrity while reducing the die area.

**Future packaging solutions**

The increasing levels of current consumed by CMOS integrated circuits as well as the high switching speeds require power distribution systems with a lower impedance over a wider frequency range. An increasingly lower inductance between the integrated circuit and the package decoupling capacitance is essential to maintain a lower

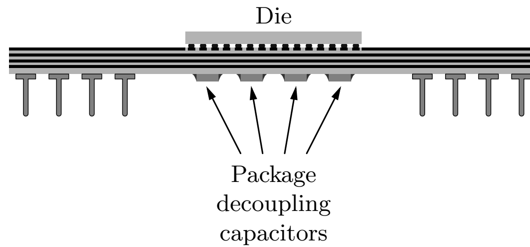


**Fig. 7.11.** Flip-chip pin grid array package. The package decoupling capacitors are mounted on the bottom side of the package immediately below the die mounted on the top side. In this configuration, the package capacitors are in physical proximity to the die, minimizing the impedance between the capacitors and the circuit.

impedance at higher frequencies. Providing a low impedance die-to-package connection remains a challenging task [226]. Future generations of packaging solutions, such as chip-scale and bumpless build-up layer (BBUL) packaging, are addressing this problem with higher density die-to-package contacts, a smaller separation between the power and ground planes, and a lower package height [114], [227], [228].

The electrical characteristics of a package have become one of the primary factors that limit the performance of an integrated circuit [229], [230]. The package design is now crucial in satisfying both the speed and overall cost targets of high performance integrated circuits. Achieving these goals will require explicit co-design of the on-chip global interconnect and the package interconnect networks [112], [229].

## 7.3 Other considerations

Dependence of the power supply integrity on the impedance characteristics of the on-chip power distribution network has been discussed previously. The on-chip power distribution network also significantly affects the integrity of the on-chip data and clock signals. The integrity of the on-chip signals depends upon the structure of the power distribution network through two primary mechanisms: inductive interaction among the power and signal interconnect and substrate coupling. These two phenomena are briefly discussed in this section.

**Dependence of on-chip signal integrity on the structure of the power distribution network**

The structure of the power distribution grid is one of the primary

factors determining the integrity of on-chip signals. The power and ground lines shield adjacent signal lines from capacitive crosstalk. Sensitive signals are typically routed adjacent to the power and ground lines. Co-designing the power and signal interconnect has become an important consideration in the design of high speed integrated circuits [231], [232], [233].

Power and ground networks provide a low impedance path for the signal return currents. The structure of the on-chip power distribution network is therefore a primary factor that determines the inductive properties of on-chip signal lines, such as the self and mutual inductance. Modeling the inductive properties of the power distribution grid is necessary for accurately analyzing high frequency phenomena, such as return current distribution, signal overshoot, and signal delay variations, as demonstrated by on-chip interconnect structures using full-wave partial element equivalent circuit (PEEC) models [234], [235]. These conclusions are also supported by the analysis of commercial microprocessors. Inadequate design of the local power distribution network can lead to significant inductive coupling of the signals, resulting in circuit failure [212].

**Interaction between the substrate and the power distribution network**

In high complexity digital integrated circuits, the ground distribution network is typically connected to the substrate to provide an appropriate body bias for the NMOS transistors. The substrate provides additional current paths in parallel to the ground distribution network, affecting the current distribution in the network, as illustrated in Fig. 7.12. This effect is significant in most digital CMOS processes which utilize a low resistivity substrate to prevent device latch-up [236]. A methodology for analyzing power distribution networks together with the silicon substrate requires a complete model, which includes both the power distribution system and the substrate, as described by Panda, Sundareswaran, and Blaauw [159]. The substrate significantly reduces the voltage drop in the ground distribution network (assuming an N-well process) by serving as an additional parallel path for the ground current to flow, as demonstrated by an analysis of three Motorola processor circuits [159]. The placement of substrate contacts also affects the on-chip power supply noise and the substrate noise [159], [237].

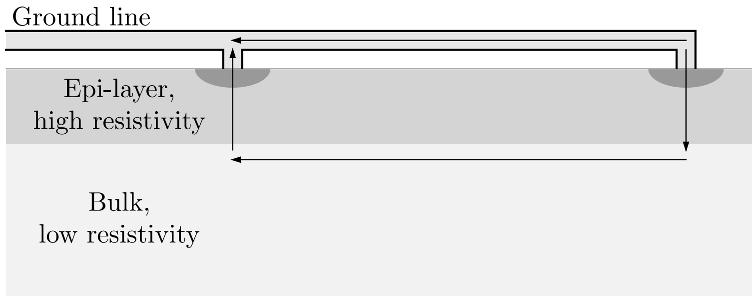


**Fig. 7.12.** Interaction of the substrate and power distribution network. The low resistivity bulk substrate provides additional current paths between the points where the ground network is connected to the substrate. These current paths are connected in parallel to the ground distribution network.

## 7.4 Summary

The structure of the on-chip power distribution network and related design considerations are described in this chapter. The primary conclusions are summarized as follows.

- On-chip power distribution grids are the preferred design style in high speed, high complexity digital integrated circuits
- Constraints placed on the impedance characteristics of the on-chip power distribution networks are greatly affected by the electrical properties of the package
- The high frequency impedance characteristics of a power distribution system are significantly enhanced in packages with an area array of low inductance I/O contacts