Basic Concepts

Power delivery is a major challenge in present-day systems. This challenge is expected to increase in the next decade as systems become smaller and new materials are introduced into packages and boards. As devices scale and more transistors are integrated into a single integrated circuit, the power and current levels are expected to increase with a corresponding decrease in the voltage. With gigabit signals being propagated through the package and board, the ability to supply clean power to the transistor circuits becomes very critical. In addition, electromagnetic interference levels have to be kept low in the system to manage coupling and crosstalk.

In this chapter, the basics of power delivery are described. Along with a description of the components of a power delivery network (PDN), the analysis methodology of such networks is described with examples.

1.1 Introduction

1.1.1 Functioning of Transistors

Integrated circuits (ICs) such as microprocessors, field programmable gate arrays, memory devices, and other application-specific ICs contain transistors. Transistors are multiterminal switches that can be turned on or off on the basis of a control signal. The on or off position of the switch determines the current flowing through the device. In complementary metal oxide semiconductor (CMOS) field effect transistor (MOSFET) technology (which is the most popular technology used to
design microprocessors), two types of transistors are used, namely the NMOS (n-channel) transistor and the PMOS (p-channel) transistor. The detailed operation of these devices can be obtained from [1]. In this book, for simplicity, we will assume that both transistors are three-terminal devices that can be represented using switches, as shown in Figure 1-1. The three terminals are called the gate, source, and drain. By applying a voltage between the gate and source, the current through the transistor (from drain to source for NMOS and reverse for PMOS) can be turned on or off. The NMOS transistor is called as a normally open switch, since a gate voltage has to be applied to pass current through the transistor. Hence, if a binary 0 (logic level low) signal exists at the gate, the switch is OFF, and when a binary 1 (logic level high) signal is available at the gate, the switch turns ON. The reverse is true for the PMOS transistor, since a binary 1 level at the gate turns OFF the current, while a binary 0 level at the gate allows current to pass, that is, the switch is ON. Hence, the PMOS transistor is called as the normally closed switch. The drain and gate terminals of the NMOS and PMOS transistors can be connected together to form an inverter, which is one of the basic building blocks in any IC. We limit ourselves to the discussion of such an inverter in this section.

Figure 1-2 shows the inverter circuit. The gate connection is called as the input node, and the drain connection is called as the output node. The output node is connected to the input node of the following transistor circuit. Since the gate of

![Figure 1-1 NMOS and PMOS transistors represented as switches.](image_url)
the transistors acts as a capacitor (formed between the metal-oxide-semiconducting substrate), the inverter (also called the driver) is used to charge and discharge the input capacitance of the succeeding stage. The capacitor must be charged to reach the binary 1 voltage level. Similarly, discharging a capacitor to 0 voltage requires the removal of charge. The inverter circuit must be connected to a power supply (shown as Vdd and Gnd terminals), which provides the ability to charge and discharge a capacitor node within the IC. In Figure 1-2, a wire (interconnection) is used between the two inverters to act as a conduit for the charge, and Ron is the on resistance of the transistor. The speed at which the circuit operates determines how quickly charge can be either supplied or removed from the capacitor through the switches. A PDN in a system provides the interconnection framework to make this happen, supplying the transistors with sufficient voltage and current for them to switch states.
1.1.2 What Are the Problems with Power Delivery?

The power supply (which is the source of the voltage and current) is typically bulky and cannot be connected directly to the Vdd and Gnd terminals of the IC. Therefore, wires (interconnections), which have resistance and inductance in them, are used to establish this connection. The current flowing through these wires creates both a DC drop (not shown) and time-varying fluctuation of the voltage across the Vdd and Gnd terminals of the IC (shown in Figure 1-3), which is detrimental to the transistors in the IC. Hence, a suitable PDN must be created between the power supply and the IC, such that the voltage is well regulated for the required current to be supplied to the transistors over a required time period.

The voltage fluctuation across the Vdd and Gnd terminals of the transistors can cause the following problems with the transistors:

- Reduction in voltage across the power supply terminals of the IC that slows down the transistor or prevents the transistor from switching states.
- Increase in voltage across the power supply terminals of the IC, which creates reliability problems.
- Leakage of the voltage fluctuation into a quiet transistor, as shown in Figure 1-3, causing incorrect switching of quiet transistor circuits at the far end of a communication path along with crosstalk from neighboring signal lines.
- Timing margin errors caused by degraded waveforms at the output of the drivers.

![Figure 1-3 Voltage fluctuation](image-url)
The voltage fluctuation across the power supply of the IC is called power supply noise, delta I noise, or simultaneous switching noise (SSN), since it occurs only during the switching of the transistors.

1.1.3 Importance of Power Delivery in Microprocessors and ICs

Given the voltage fluctuations across the power supply of a transistor, it is helpful to understand how a microprocessor operates (as an example at a holistic level) and the impact of voltage fluctuations on microprocessor performance.

A microprocessor consists of millions of CMOS transistors interconnected through wires in a very complex fashion. The microprocessor speed can be limited by the gate (or transistor) delays, interconnect (or wire) delays, or both. The inverse of the gate delay (frequency) is proportional to the gate voltage. For a gate-dominated circuit, a 1% drop in the gate voltage results in nearly a 1% drop in frequency. The interconnect delay, however, is a very weak function of voltage. An important relationship exists between the operating voltage of the microprocessor and its speed (measured as frequency) around the nominal voltage of the microprocessor. This relationship is shown in Figure 1-4 for a 64-bit Scalable Processor Architecture (SPARC) microprocessor [3]. Around the nominal voltage of 1.6 V, the relationship between frequency and voltage is almost linear. As the graph shows, a reduction in voltage reduces the operating frequency of the microprocessor, while an increase in voltage increases its frequency. This important relationship is true in most microprocessors, and we use this example to explain the impact of power supply fluctuations on the operating frequency of the microprocessor. In reality, the relationship between processor performance and voltage is more complex and depends not only on the magnitude of power supply noise but also on the frequency of the noise.

Consider Figure 1-5, which assumes a linear relationship between the frequency of a microprocessor (along the y-axis) and voltage (along the x-axis), as in Figure 1-4. In Figure 1-5, \( F_{\text{MAX}} \) is the maximum operating frequency of the microprocessor. Any voltage above 1.65 V causes reliability problems and is shown as the reliability wall. Any voltage that falls within the reliability wall causes the dielectric breakdown of the gate oxide in the MOSFET due to excessive electric field. Hence, the power supply voltage cannot exceed 1.65 V for this example. Let’s assume initially that the operating voltage of the microprocessor is 1.55 V. According to the graph, the nominal \( F_{\text{MAX}} \) for the microprocessor is 720 MHz. However, voltage variations on the power supply cause the voltage to vary plus or minus 100 mV around the nominal voltage. On the high side, a voltage of
1.65 V (1.55 V + 100 mV) is below the maximum allowed voltage of 1.65 V, which ensures no reliability problems. On the low side, the power supply voltage reduces to 1.45 V (1.55 V – 100 mV). At 1.45 V, \( F_{\text{MAX}} \) now becomes 670 MHz. Hence, any drop or reduction in the power supply voltage causes the microprocessor to operate at a lower frequency. In other words, the PDN causing the variations on the power supply terminals of the IC results in the slowdown of the microprocessor. Similarly, a voltage rise across the power supply of the IC, if it exceeds the maximum voltage allowable, causes the IC to malfunction.

In the design of PDNs, the focus is always on minimizing the voltage droop on the power supply terminals of the transistor circuits within an IC and also on ensuring that the voltage maximum does not cause reliability problems [4].

### 1.1.4 Power Delivery Network

A PDN consists of a power supply, DC–DC converters (also called voltage regulator modules, or VRMs), lots of decoupling capacitors, and interconnections
that act as conduits for the supply and removal of charge to and from the switching circuits. In a typical computer system, the IC is packaged and placed on a motherboard (with or without a socket) with a power supply on the motherboard. The power supply provides high voltage and current to the motherboard. The voltage is reduced through a DC–DC converter and supplied to the IC through the interconnections in the motherboard and package. The decoupling capacitors are distributed on the motherboard, package, and IC; they act as reservoirs where charge can be stored. The charge is supplied as needed to the transistors from the decoupling capacitors. The proximity of the capacitors to the switching circuits determines the time required to supply the charge. The required time is controlled by the speed of light in the medium, which is the minimum time required to transfer the charge from the capacitor to the transistors. As an example, the minimum time required to supply charge from a capacitor placed on the motherboard 6 inches away from a transistor circuit is 1 ns, since the speed of light in typical printed circuit boards (PCBs) is 166 ps/inch.

A typical PDN for a semiconductor is shown in Figure 1-6 [5]. Since the inverse of time delay is frequency, the proximity of the capacitors to the transistors
determines if the capacitor supplies charge at high frequencies, middle frequencies, or low frequencies. The high-, mid-, and low-frequency capacitors are shown in the figure, where a capacitor farther away from the IC is always large and bulky, thereby operating at a lower frequency. The charge storage capacity of the large capacitors is of the order of thousands of microfarads, much higher than either the high- or mid-frequency capacitors, which are in the nanofarad range.

1.1.5 Transients on the Power Supply

Although the operating frequency of a microprocessor can be high (1 GHz and higher), power supply fluctuations can be caused over a range of frequencies, because a computer is a broadband system in which transistors switch at multiple frequencies. For example, a 1 GHz microprocessor in a system may be executing instruction at 1 GHz, causing voltage fluctuations at the 1 GHz frequency. At the same time, the microprocessor may be writing data to the cache on the PCB at 400 MHz and operating the Joint Test Access Group (JTAG) line for testing the hardware at 1 MHz. Such a switching activity can cause voltage fluctuations over a range of frequencies, which makes the design of the PDN very difficult. Voltage variation
on the power supply at multiple frequencies is shown in Figure 1-7 for a microprocessor [5]. For an IC, the transient current flowing through an inductor, $L$ in Figure 1-3, causes voltage drop, $V_L$, across it, given by

$$V_L = L \frac{dI}{dt}$$

where $dI/dt$ is the rate of change of current in the circuit. The inductor $L$ can be equal to $L_V$ or $L_G$, or a combination of the two depending on the current path. A positive $dI/dt$ through the inductor causes a voltage drop across it, resulting in a reduction in the supply voltage across the IC terminals and causing a performance problem due to a negative spike in the IC supply voltage. Similarly, a negative $dI/dt$ through the inductors increases the supply voltage across the IC terminals, resulting in a positive spike, which causes reliability problems. The power supply noise has four components (1) ultra-high-frequency noise in the 10 to 100 GHz range, (2) high-frequency noise in the 100 to 1000 MHz range, (3) mid-frequency noise in the 1 to 10 MHz range, and (4) low-frequency noise in the 1 to 100 KHz range. The inductance on-chip affects both the ultra-high- and high-frequency noise (>1 GHz), while the package has a large effect on the high-frequency and mid-frequency noise.
components (10 MHz – 1 GHz). The inductance of the motherboard and the voltage regulator module affect the mid-frequency and low-frequency noise components (<1 MHz), as shown in Figure 1-7.

The ultra-high-, high-, mid-, and low-frequency noise are also called, respectively, the first, second, third, and fourth droops or spikes on the power supply.

1.2 Simple Relationships for Power Delivery

In any IC, two kinds of circuits need to be powered: the core and I/O. The core consists of transistors that are contained within an IC and that communicate with each other. The I/O, on the other hand, has to communicate with other ICs through the package and motherboard. Because the wires connected to I/O circuits exit the IC, they are very noisy and often are isolated from the core circuits using a separate PDN, as shown in Figure 1-8, where both the core and I/O circuits during switching create voltage fluctuations across the power supply. In this section, simple relationships are derived for the voltage fluctuations on a power supply for both the core and I/O circuits.

1.2.1 Core Circuits

A very simple circuit is shown for the core circuits in Figure 1-9(a), where the driver and receiver circuits are shown as 2 and 1, respectively. The PDN contains some resistance and inductance due to the parasitics of the interconnections in the network. The resistance is assumed to be negligible here. A simple equivalent
circuit for Figure 1-9(a) is shown in Figure 1-9(b). In the simplified equivalent circuit, the switch represents the PMOS transistor that closes at time $t = 0$. The resistance $R$ is the on-resistance of the transistor, and $C$ is the input capacitance of receiver circuit 1 that needs to be charged. The total inductance of the voltage and ground paths is represented by a single inductance $L$.

The goal of the core PDN is to ensure that sufficient charge is supplied to the switching circuit so that the capacitance can be charged to the required voltage. To minimize delay, the charge has to be supplied within a short time. The circuit in Figure 1-9(b) has two time constants: $L/R$ and $RC$. The delay of the transistor circuit is defined by the $RC$ delay. Since the $L/R$ time constant should have minimum impact on the $RC$ delay of the transistor, it is desired that \[ L \ll R \] (1.2)

Under this assumption, the simplified equivalent circuit in Figure 1-9(c) can be used, where the voltage drop across the inductor can be obtained by solving equation (1.3),

![Figure 1-9](image-url)
where the current is obtained by solving the differential equation:

\[ L \frac{di(t)}{dt} + Ri(t) = v(t) \]  

(1.4)

In equation (1.4), \( v(t) \) is an equivalent source voltage with rise time \( t_r \) (that combines the switch and Vdd) given by

\[
v(t) = \begin{cases} 
  Vdd \times t/\tau & 0 \leq t \leq t_r \\
  Vdd & t \geq t_r
\end{cases}
\]  

(1.5)

The rise time is dictated by the speed of the switch. The maximum voltage across the inductor occurs at time \( t = t_r \) and is given by

\[
V_{L,\text{max}} = \Delta V = \frac{L \times Vdd}{Rt_r} \left(1 - e^{-t_r/(L/R)}\right)
\]  

(1.6)

Example

Assume \( t_r = 0.1 \) ns, \( L = 0.1 \) nH, \( R = 1 \) Ω, \( C = 1 \) nF, and Vdd = 1 V. The ratio \( L/R = 0.1 \) RC, and therefore the condition in equation (1.2) is met. The maximum voltage drop across the inductor can be obtained as 632 mV from equation (1.6).

By changing the inductance value to \( L = 0.01 \) nH, the maximum voltage drop across the inductor can be obtained as 100 mV.

The voltage waveform across the inductor is shown in Figure 1-10. As the inductance is changed from 0.1 nH to 0.01 nH, the voltage drop across the inductor looks more like a rectangular pulse. Therefore, when \( t_r \) is much greater than \( L/R \), equation (1.6) can be simplified to
\[ \Delta v \approx \frac{L \times V_{dd}}{R_{t_r}} \]  \hspace{1cm} (1.7)

with a pulse width of \( t_r \).

**Example**

Consider an IC in which the total capacitance to be charged is 10 nF and the on-resistance of all the transistors in parallel is 0.1 \( \Omega \). Let the required maximum voltage drop across the inductor be 10% of \( V_{dd} \) for a rise time \( t_r = 1 \) ns. The inductance required in the PDN to meet the voltage drop can be calculated from equation (1.6) by solving:

\[ \frac{\Delta v}{V_{dd}} = 0.1 = \frac{L}{(0.1*1)} \left(1 - e^{-1/(L/0.1)} \right) \]  \hspace{1cm} (1.8)

where \( L \) is measured in nanohenry. By iterating, the inductance can be calculated as \( L = 0.01 \) nH. The inductance satisfies the condition in equation (1.2). It can also be obtained from equation (1.7), since \( t_r \) is much greater than \( L/R \).

**Figure 1-10** Voltage drop across inductor.
1.2.2 I/O Circuits

I/O circuits, unlike core circuits, drive off-chip interconnections. With increase in frequency, the interconnections behave as transmission lines where the delay becomes important. The PDN used to drive an I/O circuit is shown in Figure 1-11(a): the transmission line has a characteristic impedance of \( Z_0 \) and delay \( T \). The far end of the transmission line is terminated with a resistor \( R = Z_0 \). The inductance \( L \) represents the PDN loop inductance from the power supply to the chip terminals. As before, the transistor is represented using a switch with an on-resistance \( R \), where \( R \) is much less than \( Z_0 \) to allow for the maximum voltage to be launched on the transmission line, as shown in Figure 1-11(b).

When the switch closes, the power supply inductance \( L \) acts as an open circuit and behaves as a short circuit at time \( t = \infty \). As in the previous section, the voltage source and the switch can be combined and represented as a pulse with rise time \( t_r \). Since the far end of the transmission line is terminated in the characteristic impedance of the transmission line, there are no reflections. The maximum voltage drop across the inductor occurs at time \( t = t_r \) and can be calculated as in the previous section by replacing \( R \) with \( Z_0 \):

\[
\Delta v_{L,max} = \frac{L \times V_{dd}}{Z_0 t_r} (1 - e^{-t_r/(L/Z_0)})
\]

(1.9)

Based on equation (1.9), a signal line with low \( Z_0 \) (highly capacitive) will always result in a larger voltage drop across the inductor, assuming the inductance is fixed, as described later in this chapter. When \( t_r \) is much greater than \( L/Z_0 \), the maximum voltage drop across the inductor simplifies to

\[
\Delta v \approx \frac{L \times V_{dd}}{Z_0 t_r}
\]

(1.10)

When \( N \) parallel transmission lines of characteristic impedance \( Z_0 \) are switched simultaneously, it is equivalent to switching a single transmission line of impedance \( Z_0/N \). Hence, the maximum voltage drop across the inductor can be obtained by replacing \( Z_0 \) by \( Z_0/N \) in equations (1.9) and (1.10).

Example

Assuming \( L = 1 \) nH, \( Z_0 = 50 \) \( \Omega \), \( V_{dd} = 1 \) V, and \( t_r = 0.1 \) ns, the maximum voltage drop across the inductor can be calculated as 200 mV from
equation (1.9). Assuming the inductance is reduced to \( L = 0.1 \) nH, the voltage drop becomes 20 mV. Figure 1-11(c) shows the voltage waveforms across the inductor: the shape of the voltage drop looks more like a rectangular pulse when \( t_r \) is much greater than \( L/Z_0 \).

**Example**

Consider a 32-bit bus with \( Z_0 = 50 \) Ω. The driver is switching with a \( t_r = 0.1 \) ns. Assuming \( V = 10\% \) of Vdd is desired as the voltage drop across the inductor, the maximum power supply inductance that must be supported can be obtained by solving the following equation iteratively:

\[
\Delta v \quad 0.1 = \frac{32 \times L}{50 \times 0.1} (1 - e^{-0.1/(L/(50/32))})
\]  

(1.11)

The result is an inductance of 16 pH. Since \( t_r \) is much greater than \( L/(Z_0/N) \), the same result can also be obtained from equation (1.10).

1.2.3 Delay Due to SSN

The presence of the inductor increases the delay of the I/O circuit. The voltage at the input end of the transmission line for a pulse with rise time \( t_r \) can be computed as

\[
v(t) = \frac{Z_0 \times Vdd}{L \times t_r} \left( \frac{L^2}{Z_0^2} [e^{(t/L)} - 1] + \frac{L}{Z_0} t \right) \quad t \leq t_r
\]  

(1.12)

and

\[
v(t) = A + B(1 - e^{-t/(L/Z_0)}) \quad t > t_r
\]  

(1.13)

where

\[
A = Vdd - [Vdd - v(t_r)]e^{-t_r/(L/Z_0)}
\]

\[
B = [Vdd - v(t_r)]e^{-t_r/(L/Z_0)}
\]  

(1.14)
and \(v(t_r) = v(t = t_r)\) from equation (1.12). A transistor circuit at the receiver requires a minimum voltage at its input to switch states. Let’s assume that the minimum voltage required for this to happen at the driver output \(V_{\text{chip}}\) (input end of the transmission line) is \(0.5 \times V_{\text{dd}}\). Equations (1.12) and (1.13) can be used to calculate the time required to reach \(0.5 \times V_{\text{dd}}\) and hence represent the delay incurred because of the power supply inductance. Equation (1.12) can be used when \(t_r\) is greater than \(L/Z_0\), and equation (1.13) can be used when \(t_r\) is less than \(L/Z_0\) to calculate a 50% delay. This delay does not include the transmission line delay and is valid for a matched load, as in Figure 1-11(b).

**Example**

Consider the previous example in which \(L = 0.1 \text{ nH}, Z_0 = 50 \text{ } \Omega, V_{\text{dd}} = 1 \text{ V},\) and \(t_r = 0.1 \text{ ns}\). Since \(t_r\) is greater than \(L/Z_0\), equation (1.12) can be used to calculate the 50% delay. Using an iterative process, the 50% delay at the input of the transmission line is 0.052 ns.

Let’s now assume that a 100-bit bus, each of impedance \(Z_0\), is switched simultaneously. This translates to an equivalent impedance \(Z_0 = 50/100 = 0.5 \text{ } \Omega\) (transmission lines in parallel). Since \(t_r\) is less than \(L/Z_0\), equation (1.13) can be used to calculate the 50% delay. Through iteration, this delay can be calculated as 0.191 ns. Hence, the delay increases because of an increased voltage drop across the inductor caused by an increase in current. The voltage at the input end of the transmission line is shown in Figure 1-11(d) along with the 50% delay.

### 1.2.4 Timing and Voltage Margin Due to SSN

Timing and voltage margins are affected by crosstalk, process variation, SSN, reflection, and other effects. In this section, we address only the effect of SSN. SSN can affect the voltage margin because power supply noise can corrupt the voltage levels of the signal waveform. In the previous section, a relationship was derived between the SSN and delay: as the SSN increased for a larger number of switching drivers, the 50% delay increased as well. This delay manifests itself as jitter that affects the signal integrity of the waveform and therefore increases the timing error; see Figure 1-11(e). As an example, consider an 8-bit-wide bus. If all the bits transition simultaneously from 0 to 1 (00000000 to 11111111 for the bus), the maximum transient current from the power supply is drawn, resulting in maximum noise and hence maximum delay. If only the alternate bits transition...
Design of PDNs

(00000000 to 10101010), fewer drivers switch and therefore the noise (and delay) is lower than in the previous case. For a pseudorandom bit stream (PRBS), the number of switching drivers changes at random, resulting in random SSN. Therefore, the 50% delay associated with the rising edge changes with the bit pattern, resulting in an uncertainty in the position of the rising edge. This effect is called jitter, shown in Figure 1-11(e). Jitter results in a timing uncertainty whereby a longer time interval may be required to latch the data for all the bit patterns if the jitter is large. Hence, the goal in I/O signaling is to ensure the smallest timing error by controlling jitter, which is possible by reducing SSN in addition to other parameters. This ensures a suitable timing margin. In Chapter 5, this effect is described in more detail through an example.

1.2.5 Relationship between Capacitor and Current

As mentioned earlier, decoupling capacitors serve as charge reservoirs and provide current to the switching circuits. Let’s assume that the power supply inductance is small such that equation (1.10) is valid. Consider a single 50-Ω driver, which requires a current of 0.1 A assuming Vdd = 5 V (ΔI = 5/50). Let’s assume that a 100-nF capacitor is available to provide charge to the switching circuits during a time interval of 10 ns (τr) that keeps the power supply fluctuations to within 10% of Vdd. The current that can be supplied by the capacitor that maintains Δv to be 10% of Vdd is given by [6]

\[ \Delta I = \frac{C \Delta v}{\tau_r} = \frac{100 \times 10^{-9} \times 0.5}{10 \times 10^{-9}} = 5 \text{ A} \]

(1.15)

Since a single driver requires 0.1 A to charge the interconnections, the 100-nF capacitor can provide the current to 50 I/O circuits over a period of 10 ns.

1.3 Design of PDNs

Since a computer system supports multiple frequencies, a PDN is best designed in the frequency domain. The response of the PDN to switching circuits can then be viewed in the time domain to evaluate the transient noise voltages generated on the power supply terminals of the IC or between any other nodes in the system. The response of the PDN in the frequency domain enables a designer to understand all the resonances and antiresonances in the system produced by the interaction of inductances and capacitances in the network. An antiresonance, when excited by a
Figure 1-11  (a) I/O circuit switching. (b) Simple equivalent circuit. (c) Voltage drop across inductor.
Uncertainty in delay due to SSN causing jitter

Varying voltage droop on power supply due to SSN

Figure 1-11  (d) Delay due to 1 driver and 100 drivers switching. (e) Jitter caused by simultaneous switching noise. (Courtesy of Sony.)
source, always generates the maximum noise voltage across the power supply terminals of the IC. Based on the frequency response of the PDN, the designer can evaluate the importance of the antiresonances in the system and decide if the source (switching circuits) will ever excite these antiresonances. Hence, the signature of the source along with the frequency response of the PDN decides the noise voltages on the power supply in the time domain. In this section, the concept of target impedance is introduced. The use of target impedance as a design parameter is discussed by evaluating a simple circuit in the frequency and time domain.

### 1.3.1 Target Impedance

The target impedance is based on Ohm’s law, which states that the ratio of voltage to current has to equal the impedance in the network. For a PDN, the voltage is the allowed ripple ($\Delta v$) on the power supply. The target impedance $Z_T$ (in ohms) of a PDN can then be calculated as

$$Z_T = \frac{V_{dd} \times \text{ripple}}{50\% \times I_{\text{max}}} (\Omega)$$

(1.16)

where the average current drawn by the switching circuits is assumed to be 50% of the maximum current and $V_{dd}$ is the power supply voltage. Assuming a voltage of 5 V with a ripple of 5% and a maximum current of 1 A, the target impedance can be calculated as

$$Z_T = \frac{5 \times 5\%}{50\% \times 1} = 0.5 \Omega$$

(1.17)

The maximum current drawn by an IC can always be calculated by using the relationship $P = V_{max} I_{max}$, since both power $P$ and voltage $V$ for an IC are known. The target impedance $Z_T$ establishes an upper limit for the maximum impedance for the PDN across the power supply terminals of the IC in the frequency domain. An impedance below $Z_T$ ensures that any current transients will always generate noise voltages of less than 5% of 5 V. Hence, $Z_T$ is a very useful parameter for designing PDNs in which the noise voltages have to be controlled within, say, 5% of the supply voltage.

A plot of $Z_T$ versus frequency is shown in Figure 1-12. The frequency axis represents the frequency components associated with the source excitation. According to the figure, if the impedance exceeds the target impedance at any frequency where the current transients can excite the network, then the resulting
power supply noise will exceed 5% of 5 V = 250 mV. The figure assumes that the magnitude of the current transients is 50% of the maximum current.

The target impedance calculations for five microprocessors introduced between 1990 and 2002 are shown in Table 1-1. As can be seen, the target impedance has decreased 500-fold over a decade because of the lowering of the supply voltage and increase in power. Since the impedance of the PDN is also given by $Z = \sqrt{LC}$, where $L$ and $C$ are the inductances and capacitances in the network, a low target impedance always implies large capacitance and low inductance in the network. In Table 1-1, the frequency of the microprocessor has increased from 16 MHz to 1.2 GHz over a decade, which implies that the target impedance has to be maintained at least up to the fundamental frequency of the clock. However, this

**Figure 1-12**  $Z$ versus frequency.

Table 1-1 Target Impedance Trends

<table>
<thead>
<tr>
<th>Year</th>
<th>Voltage (Volts)</th>
<th>Power dissipated (Watts)</th>
<th>Current (Amps)</th>
<th>$Z_{\text{target}}$ (m$\Omega$)</th>
<th>Frequency (MHz)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1990</td>
<td>5.0</td>
<td>5</td>
<td>1</td>
<td>250</td>
<td>16</td>
</tr>
<tr>
<td>1993</td>
<td>3.3</td>
<td>10</td>
<td>3</td>
<td>54</td>
<td>66</td>
</tr>
<tr>
<td>1996</td>
<td>2.5</td>
<td>30</td>
<td>12</td>
<td>10</td>
<td>200</td>
</tr>
<tr>
<td>1999</td>
<td>1.8</td>
<td>90</td>
<td>50</td>
<td>1.8</td>
<td>600</td>
</tr>
<tr>
<td>2002</td>
<td>1.2</td>
<td>180</td>
<td>150</td>
<td>0.4</td>
<td>1200</td>
</tr>
</tbody>
</table>

Information from Smith [7].
overly restrictive condition may not be satisfied at all frequencies and can often increase the cost of the system. Hence, care should be taken to correlate the frequency response with the current transients in the system to better understand the frequencies at which the PDN will be excited. The target impedance should be maintained at these excitation frequencies.

### 1.3.2 Impedance and Noise Voltage

Consider the circuit shown in Figure 1-13(a). The circuit has a supply voltage of 2.0 V. The 3-mΩ resistance and 320-pH inductance are the spreading resistance and inductance from the power supply to the capacitor. Spreading resistance and inductance produce resistive and inductive drops when the current travels from the power supply to the capacitors (through the interconnects) for charging them. The capacitor parameters are equivalent series resistance ($ESR = 10$ mΩ), equivalent series inductance ($ESL = 1$ nH), and $C = 100$ μF, resulting in a resonant frequency of 0.5 MHz, which are explained in detail later. The on-chip capacitance is 800 nF in the circuit. The current source is 1 A between the voltage and ground terminals of the IC, and through an AC analysis, the voltage (or impedance in ohms) can be obtained as shown in Figure 1-13(b). In Figure 1-13(a), a 1-A current source is used to represent the current, and hence the voltage across it is the impedance in ohms ($Z = V/I$). In the frequency response, the resonant frequency of the decoupling capacitor can be seen, and the large impedance at approximately 13 MHz is caused by the antiresonance between the chip capacitance and ESL of the decoupling capacitor, which is explained later. The null in the impedance profile is called a resonance; the peak in the impedance profile is called the antiresonance. For a 2-V supply, 5% tolerance, and a 10-A average current, the target impedance is $10$ mΩ. Therefore, the maximum impedance allowed across the current source (which represents the switching circuit) is $10$ mΩ. Clearly, the target impedance is met up to a frequency of 5 MHz in Figure 1-13(b). In the frequency range from 5 MHz to 100 MHz, the target impedance has been exceeded.

Let’s now look at the response of this network to two current signatures. The circuit used to compute the time-domain response is shown in Figure 1-14(a). The switching circuit is represented using a time-dependent resistor, the resistance of which changes from 97 mΩ to 197 mΩ, which corresponds to a 10-A change in current in the circuit, assuming only 3 mΩ of resistive impedance (no inductance) is present in the PDN. The current changes from 20 A (2/100 mΩ) to 10 A (2/200 mΩ) in the circuit. The voltage across the time-dependent resistor is shown in Figure 1-14(b) for a current transient with rise time of 10 ns and period of 1 μs. As explained earlier, the transient voltage across the IC power supply contains
transients with both positive and negative peaks. The noise voltage settles to within 5% of 2 V after 50 ns following the switching activity. Hence, during a large part of the 1-μs period, the noise is below the 5% tolerance value. The 10-ns rise time has enough frequency components that exceed the target impedance initially, causing the first negative glitch to exceed the 100 mV tolerance value. If this negative glitch is a problem, then the impedance at frequencies corresponding to the rise time must be reduced.

Let’s now consider the noise voltage when the current transient has a rise time of 10 ns and period of 80 ns, corresponding to a frequency of about 13 MHz, which coincides with the antiresonant frequency. The noise voltage is shown in Figure 1-14(c), which is 200 mV for the entire period of 1 μs and hence exceeds the noise budget of 100 mV. This example shows the importance of managing the impedance of the PDN in the frequency domain to manage excessive noise caused by the current transients.

Figure 1-13  (a) Circuit of PDN. (b) Frequency response.
1.4 Components of a PDN

The elements of the PDN are shown in Figure 1-15 [8]; they include the chip-level power distribution with thin-oxide decoupling capacitors; the package-level power distribution with planes and mid-frequency decoupling capacitors; and the board-level power distribution with planes, low-frequency decoupling capacitors, and VRM. The frequency ranges covered by these elements are also shown in the figure: the power distribution operates at a higher frequency as the proximity to the active devices decreases because of the parasitic inductance and resistance of the interconnections between the active circuitry and the various elements of the PDN. These parasitic effects are explained in the next section.

1.4.1 Voltage Regulator

Computer systems require multiple DC voltages to operate. These voltages have to be well regulated and should be able to supply the required current over a range of frequencies. The trend of increasing power and lowering supply voltage requires designers to move AC–DC and DC–DC converters closer to the electronics.
A representative class of low-voltage high-current application is the core supply of central processing units (CPUs), digital signal processors (DSPs), and large switching chips. The voltage required may be in the 0.8 V to 2.5 V range, with the current in excess of 100 A for the largest devices. Since the core voltage is often unique and may be required only by the particular device, the DC–DC converters usually feed only one load and hence are also called point-of-load (POL) converters.

Besides the large current requirements, modern electronic circuits contain elements with several different supply voltages. Legacy 5-V and 3.3-V logic devices are still common, but newer devices often require 2.5 V, 1.8 V, 1.5 V, or even lower supply voltage. The pressing need for optimizing device speed while minimizing current consumption leaves little room to combine supply rails with similar but not exactly the same nominal voltage. The solution is therefore to place several DC–DC converters on the board to create the different supply voltages. The topology of these DC–DC converters is determined by two major system constraints: (1) most of the supply voltages are lower than the voltage of the primary source to the board (output of AC–DC converter or battery), so these converters usually have to step down the voltage; and (2) isolation is very seldom required in these converters. In AC-powered systems, the isolation can be easily provided in the AC–DC converters.

Because of these constraints, the single-phase, nonisolated buck converter is the most widely used DC–DC converter topology today, though for high-current...
applications, multiphase converters are also becoming popular. In a few applications, step-up boost converters and polarity-reversing buck-boost converters are also used.

1.4.1.1 Operating Principle

The VRM converts one DC voltage to another [9]. It has a reference voltage and a feedback loop. It senses the voltage near the load and adjusts the output current to regulate the voltage at the load. The bandwidth of the regulation loop is usually between one and several hundred kilohertz. At frequencies above the loop bandwidth, the VRM becomes high impedance, and therefore the voltage is no longer well regulated.

1.4.1.2 Four-Element Model

A four-element linear model for a VRM as described in [9] is explained in this section. Figure 1-16(a) is a simplified block diagram for a buck-switching regulator, commonly found in VRMs. At the left of the figure is an input voltage, assumed to be relatively constant. The function of inductor $L_1$ is to store up energy when switch $S_1$ is closed, and deliver current to the load. If $L_1$ has more current than the load is demanding, $S_1$ opens and $S_2$ closes. Current continues to flow to the load, but in an ever-diminishing amount until $S_2$ opens and $S_1$ closes again. There is an amplifier $A$ with frequency compensation that senses the load voltage with respect to a reference voltage. When the load voltage is too low, it causes the switches and inductor to ramp up the current. When the load voltage is too high, it causes the switches and inductor to ramp down the current. The inductor current is integrated in $C_1$, which smooths the voltage. $C_1$ has an ESR. The buck regulator is nonlinear because switches open and close as a function of time.

Figure 1-16(b) shows the linearized model of the VRM consisting of an ideal voltage source and four passive elements. In the linear model, $R_0$ is the value of the resistor between the VRM sense point and the actual load and is usually only a few milliohms. $L_{out}$ represents the output inductance of the VRM. It may be the inductance of cables that connect the VRM to a system board or it may be the inductance of pins that connect a VRM to a module (about 200 and 4 nH, respectively). The maximum effective frequency for the VRM is determined by $L_{out}$. $R_{flat}$ represents the ESR (explained in the next section) of the capacitor associated with the VRM. Generally, the capacitor determines the output impedance of the VRM at frequencies beyond the response time of the loop. The ideal voltage source has the value of the power supply voltage. The value of $L_{slew}$ is chosen so that current will be ramped up in the linear model in about the same
time that it is ramped up in a real VRM. It is calculated from the equation \( V = \frac{L_{\text{t}}dI}{dt} \). In the equation, \( V \) is the amount of voltage droop or spike that can be accepted (say 5% of 1.8 V). The maximum transient current is used for \( dI \). The total amount of time for the VRM to ramp this transient current either up or down is used for \( dt \). As described in [7], typical model values for a VRM are \( R_0 = 1 \text{ m}\Omega \), \( L_{\text{out}} = 4 \text{ nH} \), \( R_{\text{flat}} = 30 \text{ m}\Omega \), and \( L_{\text{slew}} = 67.5 \text{ nH} \).

### 1.4.1.3 Design Challenges

The challenges for the DC–DC converters are multifold [8]. As a first challenge, the converters have to feed the low-voltage load with reasonable efficiency over a widely varying load-current range, which often requires synchronous rectification to keep losses low. Since the POL converters have to be placed close to the load, which will eventually dissipate the full output power, increasing the efficiency of the POL converter barely reduces the total power dissipation. However, higher converter efficiency can result in a smaller converter volume, which is usually the
driving factor. Depending on the size and cost of the converter, their efficiencies are in the 85% to 95% range.

A second challenge is to optimize the converter’s control loop to provide sufficiently low output transient ripple against the varying load current. Especially in the case of cascaded DC–DC converters, where the converter’s input may have little transient filtering, the upstream converter’s output has to deal with large current fluctuations. For example, a POL converter with 1.0-V output voltage and 30-A maximum current rating with a maximum of 60 mVpp load transient noise (excluding switching ripple) requires an output impedance below 2 m\(\Omega\) (including the capacitors) \((Z = 60 \text{ mV}/30 \text{ A} = 2 \text{ m}\Omega)\). At DC, providing low output resistance is relatively easy. With increasing frequency, however, the dropping loop gain creates an increasing output impedance of the converter. For guaranteeing unconditional stability against the unknown load impedance, some converters have very low bandwidth. If the converter’s output impedance, for instance, exceeds the required 2 m\(\Omega\) at 1 kHz, the on-board capacitors have to provide the impedance. At 1 kHz, 80,000 \(\mu\)F capacitance is required for providing a 2-m\(\Omega\) capacitive reactance \((Z = 1/(2\pi \times 1000 \times 80,000 \times 10^{-6}))\). Figure 1-17 shows the small-signal output impedance of a POL converter at 1.5-V 20-A load, with a 680-\(\mu\)F external capacitor.

A third challenge is to keep the conducted and radiated emissions of the converters under control. The converters are often placed very close to high-speed, low-swing digital interconnections and sensitive analog circuits. Since the peak AC current ripple is always higher in the converters than their DC output current, care must be taken to minimize the switching noise the converters introduce to nearby circuits. To reduce this interference, spread-spectrum converters have been introduced [9].

### 1.4.2 Bypass or Decoupling Capacitors

Switching transistor circuits requires current to charge the load. This current must be supplied by the PDN. When the VRM is unable to respond because of high output impedance, the current should be supplied by an alternative source for maintaining the voltage. In other words, when the output impedance of the VRM exceeds the desired impedance, then an alternative method is necessary to pull down the impedance. Bypass capacitors perform this function. Since capacitors store charge, they bypass the VRM and supply current to the switching circuits when there is a demand for it. These capacitors are also called decoupling capacitors because they decouple the VRM from the switching circuit. Bypass capacitors are
classifying components of a PDN into low-frequency, mid-frequency, and high-frequency capacitors depends on their operating range and proximity to the transistor circuits.

1.4.2.1 Factors Affecting the Performance of Bypass Capacitors

The bypass capacitors are surface mount devices (SMDs) attached to pads on the PCB or package. The SMD capacitors have two terminals, one attached to the voltage plane and the other to the ground plane, as shown in Figure 1-18. When SMD capacitors supply charge (or current), the current leaves the voltage plane, travels through the voltage via, flows through the capacitor, and returns through the ground via and then to the ground plane, as shown in Figure 1-18. Given the current path, the factors affecting the capacitor performance are as follows:

- Since the capacitor electrodes are made with conductors that have finite conductivity, they have resistance associated with them called as the equivalent series resistance (ESR) of the capacitor. Time-varying current

![Figure 1-17](image-url)
flowing through the capacitor produces a magnetic field, resulting in inductance, called the equivalent series inductance (ESL) of the capacitor. The ESL interacts with the capacitance of the capacitor, causing it to resonate. The capacitors are capacitive below the resonant frequency and become inductive above the resonant frequency. The impedance of the capacitor can be written as

\[ Z = R + j\omega L + \frac{1}{j\omega C} \]  

From equation (1-18), the capacitor resonates at a frequency

\[ f = \frac{1}{2\pi\sqrt{LC}} \]  

and at the resonance frequency, the impedance of the capacitor is \( Z = R \). Hence, the minimum impedance achievable with the capacitor is \( R \). At low

frequencies, since the resistance and inductance contribution is low, the magnitude of the impedance (in dB) is given by

\[
20 \log(|Z|) = -20 \log(2\pi fC)
\]  

(1.20)

Hence, below the resonant frequency of the capacitor, the impedance has a negative slope of –20dB/decade. At higher frequencies, beyond the resonant frequency, the inductance contribution begins to exceed the resistance and capacitance contribution and the impedance therefore becomes

\[
20 \log|Z| = 20 \log(2\pi fL)
\]  

(1.21)

which has a positive slope of 20dB/decade (see Figure 1-19). To obtain low impedance for a given capacitance over a broad frequency range, both the ESR and ESL should be reduced so that the minimum impedance is small and the impedance beyond the resonant frequency is decreased as well.

- The currents flowing through the vias in Figure 1-18 produce a time-varying magnetic field. For both the voltage and ground vias, the magnetic field surrounds the via in such a way that the flux passes through the loop formed by the voltage and ground currents. Such a magnetic flux results in inductance that can be reduced by reducing the loop area by placing the voltage and

![Figure 1-19](#)  

**Figure 1-19** Frequency response of a capacitor.
ground vias next to each other. Hence, the proximity of the voltage and ground pads that connect to the capacitor and the position of the vias with respect to the pad position become very critical; see Figure 1-20, which shows that the inductance can be halved by moving the pads and vias close to each other.

- Using multiple via connections per pad can decrease the overall loop inductance [10]. With state-of-the-art low-inductance capacitor constructions, the inductance limitation becomes the external connections formed by pads, escape traces, and vias. This realization has given rise to capacitor case styles with multiple terminals. Today, the lowest inductance can be achieved with the various C4 or BGA capacitor packages [11]. When connected to the power and ground planes, the vertical via connections remain as the ultimate limiting factor for lowering inductance, which is discussed in Chapter 5.

- The current from the capacitors must travel on the voltage plane to reach the transistor circuit and return through the ground plane back to the capacitor, thus forming a current loop. The voltage and ground planes therefore add additional inductance given by

\[
L = \frac{\mu kl}{w}
\]  

(1.22)

where $L$ is the per-unit-length (pul) inductance, $d$ is the separation between the planes, and $w$ is the plane width. The ability of the capacitor to supply current at higher frequencies can be enhanced by reducing the spacing between the voltage and the ground planes, thereby reducing the loop inductance.

- Computer systems often have multiple capacitors of various types in parallel, which produces an antiresonance when the first capacitor becomes inductive while the second capacitor is in its capacitive region. The parallel circuit formed by the inductance and capacitance in parallel produces an impedance peak, which can be controlled by reducing the inductance (ESL) and resistance (ESR) of the capacitor. A similar effect can be seen when a decoupling capacitor is attached to a power and ground plane in its capacitive region, as explained by the following example.

**Example**

Consider a plane pair (power plane over ground plane) of lateral dimensions 250 mm by 250 mm. The space between the planes is filled with a dielectric material of relative permittivity $\varepsilon_r = 4$ and thickness 200 $\mu$m. A decoupling capacitor with $L = 2.53$ nH (ESL) and $C = 100$ nF is connected between the power and ground planes. The resonant frequency of the capacitor is 10 MHz from equation (1.19). Around this frequency, the plane pair behaves as a lumped capacitor with capacitance $C_p = 11.07$ nF. After the decoupling capacitor is mounted on the plane pair, the impedance of the plane pair becomes

$$Z = \frac{1 - \omega^2 LC}{j\omega \left( C + C_p \left( 1 - \omega^2 LC \right) \right)}$$

(1.23)

At antiresonance, the impedance becomes infinity (high impedance), and therefore, by setting the denominator to zero, the antiresonance frequency is

$$f = \frac{1}{2\pi \sqrt{\frac{L}{C + C_p} \frac{C_p C}{C}}} = 31.7 \text{ MHz}$$

(1.24)
The characteristics of various capacitors used in a system are described in the next sections.

1.4.2.2 Bulk Decoupling Capacitors

Bulk decoupling capacitors maintain the PDN impedance at the required value beyond the VRM frequency and until the frequency at which mid-frequency capacitors become useful (typically from a few kilohertz to a few megahertz). Since bulk capacitors are connected to the VRM on one side, their values must be estimated in conjunction with the VRM output impedance.

Example

The approximate value of the bulk capacitance can be estimated using a simple example described by Smith and colleagues [7]. Assuming there is a 20-A current transient, the VRM responds in 15 μs, and the PDN must remain within 5% of a 1.8 V power supply. The amount of bulk capacitance required can be estimated as

\[
C = I \frac{dt}{dv} = 20 \times \frac{15 \times 10^{-6}}{1.8 \times 0.05} = 3333 \, \mu F
\]

Equation (1.25) assumes that capacitance is available at all frequencies, which is not true, since the parasitics of the capacitor come into play as frequency increases.

Bulk capacitors typically have an ESR value in the range of 2 to 100 mΩ. Capacitors with a large ESR produce a flat impedance over a large frequency range, while a reduced ESR results in a sharper resonance, as the next example demonstrates.

Example

Consider two bulk capacitors, one with \( C = 10,000 \, \mu F, ESR = 50 \, m\Omega \), and \( ESL = 10 \, nH \), and another with \( C = 22 \, \mu F, ESR = 2 \, m\Omega \), and \( ESL = 1 \, nH \). The frequency response of both the capacitors is shown in Figure 1-21 on a log-log scale. The capacitor with lower ESR produces a sharper resonance.
As expected, at lower frequencies, the impedance of the 10,000-μF capacitor is lower because it has larger capacitance.

An important parameter used to assess the usefulness of a capacitor is its quality factor $Q$. The quality factor of a capacitor is the ratio of its reactance to resistance and, for a series $RLC$ circuit representation as in equation (1.18), can be written as

$$Q = \frac{\omega L - \frac{1}{\omega C}}{R}$$  \hspace{1cm} (1.26)

In the inductive part of the frequency response, $Q$ can be approximated as

$$Q = \frac{\omega L}{R}$$  \hspace{1cm} (1.27)
The $Q$ factor should be minimized for a bypass capacitor either by reducing the ESL for a low ESR capacitor or by increasing the ESR for a high ESL capacitor (which is typically not preferred because it increases the impedance at resonance).

For wide frequency portions, the target impedance requirement may be flat, which corresponds to resistive impedance. Most available decoupling capacitors, however, have moderate or high $Q$ factor, making it a challenge to create the flat impedance profile required. It has been shown that bypass capacitors with $Q$ much less than 1 help to create flat impedance profiles with a minimum number of components [12], [13]. For fixed ESR and ESL, the $Q$ of the capacitor varies inversely with capacitance. Therefore, creating smooth impedance transitions with large bulk capacitors, even with a low ESR, is an easier task. Providing a smooth impedance profile with multiple lower-valued, low-ESR ceramic capacitors is difficult and challenging, especially when the frequency dependency of capacitance, resistance, and inductance are taken into account. Reference [12] introduced the bypass quality factor (BQF) as a measure of effectiveness of the capacitor to cover a wide frequency range, where $\text{BQF} = \frac{C}{L}$ ($C$: capacitance; $L$: inductance), indicating that a capacitor is more effective if the $C/L$ ratio is higher.

For several hundred microfarad and higher capacitance values, tantalum, niobium, and various electrolytic capacitors have been used. The large capacitance dictates relatively large capacitor bodies, which in turn represent large inductance. Electrolytic capacitors in standard radial packages require a bottom seal in the can, creating few nanohenries of inductance. Tantalum and niobium capacitors are usually offered in brick case styles. The typical construction has a clip connection for the anode, introducing more than one nanohenry of inductance in spite of the smaller case style. Recently, low-inductance face-down constructions have been introduced with significantly lower inductance [14].

1.4.2.3 Mid-Frequency Decoupling Capacitors
The mid-frequency SMD capacitors are useful in the 10 to 100 MHz range and higher. These capacitors are primarily ceramic capacitors that come in several dielectric types (NPO, X7R, X5R, and Y5V) and several sizes (1206, 0805, 0603). NPO capacitors have the lowest ESR and best temperature and voltage properties but are only available up to a few nanofarads. X7R capacitors have reasonable voltage and temperature coefficients and are available from several nanofarads to several farads. X5R capacitors are similar to X7R but have reduced reliability and are being extended to 100 $\mu$F. Y5V dielectric capacitors are used to achieve high capacitance values but have very poor voltage and temperature characteristics [7].
The mid-frequency capacitors are much smaller than the bulk capacitors and therefore can be placed closer to the transistor circuit. Since the ceramic capacitors are smaller, they have lower ESR and ESL and lower capacitance than bulk capacitors, leading to a higher resonance frequency with a smaller impedance at resonance. Therefore, ceramic capacitors can be used at higher frequencies. Typical mid-frequency capacitors have capacitance in the range of 1 to 100 nF, ESR in the range of 10 to 100 mΩ, and ESL in the range of 0.5 to 1 nH.

Figure 1-22 shows the impedance versus frequency for several X7R and NPO ceramic decoupling capacitors. Notice that as the capacitance value decreases and the resonant frequency goes up, the impedance minimum does not dip down as low, since ESR is lower for high-valued capacitors. Capacitor effectiveness can be optimized by using low-inductance pads to increase the resonant frequency of a given capacitor, as shown in Figure 1-20 [7].

More recently, embedded decoupling capacitors are being investigated for decoupling at higher frequencies. These capacitors are integrated within a package as an extra capacitor layer or between the voltage and ground planes. These capacitors are discussed in more detail in Chapter 5.

1.4.3 Package and Board Planes

Planes play a very important role at high frequencies by acting as high-frequency capacitors, serving as conduit for the transportation of current, and supporting the return currents of the signal lines referenced to it. Planes are large metal structures separated by a thin dielectric and are invariably used in all high-frequency packages and boards for power delivery and shielding. A plane pair is shown in Figure 1-23, which depicts the following:

- Voltage and ground planes transport the current from the SMD capacitors to the switching circuits. The planes therefore control the inductance and delay from the capacitor to the switching circuit. At high frequencies much beyond the resonant frequency of the capacitor, the plane inductance dominates the impedance of the power distribution network.

- The capacitance formed between the voltage and ground planes can be used to decouple the power supply at high frequencies and hence becomes a useful contributor.

- Planes carry the return current of the signal lines, and hence the voltage fluctuations between the voltage and ground planes across the package and board are dictated by the plane behavior.
Package planes are effective for power distribution in the mid- to high-frequency range [8]. However, a major problem with power and ground planes is their behavior as electromagnetic resonant cavities, where dielectric constant of the insulator and the dimensions of the cavity determine the resonance frequency. When excited at the resonance frequency, the planes can become a significant source of noise in the package and board and also can act as a source.

Components of a PDN

1.4.3.1 Power Plane Resonance

Figure 1-24 depicts the distribution of voltage fluctuations on the power and ground planes for an open-ended board of size $a \times b$. As can be seen from the figure, the voltage distribution on the plane depends on the resonance mode (discussed in the next section), while the resonance frequency is determined by the mode number, dielectric constant of the insulator, and physical size of the planes. The resonance frequency of the planes is given by

$$f_{mn} = \frac{1}{2\pi \sqrt{\mu \epsilon}} \sqrt{\left(\frac{m\pi}{a}\right)^2 + \left(\frac{n\pi}{b}\right)^2}$$

(1.28)
where the first mode in Figure 1-23 (assuming $b > a$) corresponds to a resonant frequency of

$$f_{01} = \frac{1}{2\pi\sqrt{\mu\varepsilon}} \frac{\pi}{b} = \frac{1}{2\sqrt{\mu\varepsilon}b}$$

(1.29)

In equations (1.28) and (1.29), $\mu$ and $\varepsilon$ correspond to the permeability (typically that of air) and permittivity of the material between the planes. The distribution of the voltage across the plane depends on the source location.

As can be seen in Figure 1-24, at the resonant frequencies, the voltage distribution is maximum and minimum at certain points on the planes. This variation in the voltage fluctuation across the plane is called plane bounce. Large voltage fluctuations on planes can cause significant coupling to signal lines referenced to it, which can propagate to quiet receiver circuits and can also cause electromagnetic emission from the edges that are unterminated.

1.4.3.2 Plane Impedance

Since planes are passive structures, they can be represented using transfer functions. The parameter that is often used to understand plane characteristics is the impedance. The impedance of any two-port circuit can be defined as

$$V_1 = Z_{11}I_1 + Z_{12}I_2$$

$$V_2 = Z_{21}I_1 + Z_{22}I_2$$

(1.30)
where $V$ and $I$ are the voltage and current at the two ports, and $Z$ is the impedance. Equation (1.30) can be extended to an arbitrary number of ports.

Consider Figure 1-23. It consists of two planes, with the top plane assigned the DC voltage $V$ and the bottom plane assigned to the DC voltage zero or ground. At the far end (right corner), the bottom plane is connected to an ideal ground, meaning that this is the position where the voltage is kept constant at 0 V by connecting to ideal ground. A port consists of two points (or nodes). For each port, the first node is on the top plane and the second node is on the bottom plane located under the first node. The bottom nodes represent the reference points for each port so that a voltage can be measured between the voltage and ground nodes. Assuming $I_1 = 1$-A current source is connected at port 1 between the voltage and ground nodes with zero current at port 2, then the measured voltage $V_1$ at port 1 is the impedance $Z_{11}$, while the measured voltage $V_2$ at port 2 is the impedance $Z_{21}$. The same procedure can be repeated by applying a 1-A current source at port 2 and leaving port 1 as an open circuit to compute $Z_{12}$ and $Z_{22}$. Impedances $Z_{11}$ and $Z_{22}$ are called as the self-impedances, while $Z_{21}$ and $Z_{12}$ are called as the transfer impedances. These impedances can be modeled and measured, and can be used to represent the behavior of the planes. Since impedances are circuit quantities, they can be used in some circuit simulators to capture the behavior of planes.

1.4.3.3 Practical Considerations

Because the size of the package is smaller than the board (dimension $a$ or $b$), the plane resonance frequencies of the board appear at a lower frequency than the package. In reality, package and board planes are connected to each other using solder balls and vias, and they can contain slits and cut-outs, resulting in a complicated behavior of the PDN. In addition, when numerous decoupling capacitors are connected to a power or ground plane cavity through the power or ground vias, the resonance frequency and the associated field distribution change because of the change in the effective capacitance and inductance of the plane cavity. The degree of the field distribution change and the resonance frequency shift depends on the capacitance and ESL of the decoupling capacitors and vias. Furthermore, the field distribution and the resonance frequencies can be slightly modified by the die attachment onto the package substrate.

At the plane antiresonance frequency, the power distribution impedance reaches its highest value, with the maximum value dictated by the losses in the structure. The loss includes radiation loss, conductor loss, dielectric loss, and losses associated with any components mounted on the package or board. The loss lowers the quality factor at resonance and hence reduces the noise [15]. In general,
radiation and dielectric loss do not provide enough damping to completely eliminate resonances. Conductor loss can damp the resonance between power and ground planes when thin dielectrics are used [15].

At the resonance frequencies of the power and ground planes, the self-impedance and transfer impedance magnitudes may be large enough to create signal-integrity and electromagnetic interference (EMI) problems. The resonance can be suppressed in several ways. It has been shown that dielectric thickness below 10 μm between the power and ground planes forces a large part of the electromagnetic field to travel in the conductor rather than in the dielectric, thus effectively suppressing the plane resonance through conductive loss [16]. Lossy dielectric layers have also been proposed [17], though their potential impact on signals has not been published yet. For power-ground laminates with thickness greater than 50 μm, the plane resonance must be suppressed by other means. The smoothest impedance profile can be achieved with the lowest number of parts if the cumulative ESR of bypass capacitors equals the characteristic impedance of the planes, which requires either ceramic bypass capacitors with controlled ESR or low-inductance external resistors in series with low-ESR bypass capacitors [18].

### 1.4.4 On-Chip Power Distribution

High-performance on-chip power distribution networks [8] are constructed as multilayer grids, as shown in Figure 1-15. Since the on-chip power distribution is in close proximity to the switching circuits, it operates at frequencies above 1 GHz. Although on-chip power distribution is not the focus of this book, it is important to understand a few characteristics of this network, since it is one component in the system-level PDN.

Designing on-chip power distribution networks in high-performance microprocessors has become very challenging because of the continual scaling of CMOS process technology [19]. Each new technology generation results in a rapid increase in circuit densities and interconnect resistance, faster device switching speeds, and lower operating voltages. These trends lead to microprocessor designs with increased current densities and transition rates and reduced noise margins. The large currents and interconnect resistance cause large, resistive IR voltage drops, while the fast transition rates cause large inductive \( \frac{dI}{dt} \) voltage drops in on-chip power distribution networks. Along with large voltage drops due to large \( \frac{dI}{dt} \), electromigration (EM) is one of the critical interconnect failure mechanisms in ICs [20]. Electromigration, which is the flow of metal atoms under the influence of high current densities, causes increased resistance
and opens in on-chip interconnects, causing further IR drops and potential reliability problems.

1.4.4.1 On-Chip Capacitors

On-chip power distribution systems for high-performance CMOS microprocessors must provide a low impedance path over a wide frequency range beyond 1 GHz. The impedance of the power distribution inductance increases with frequency according to \( Z = j\omega L \), where \( \omega = 2\pi f \) is the frequency, and \( L \) is the inductance. On-chip decoupling capacitance is used as a local charge supply, which effectively lowers the power distribution impedance at high frequencies. Hence, high-frequency switching currents are “decoupled” from the inductance in the power distribution system, and switching noise is therefore reduced. The on-chip decoupling capacitance includes both the intrinsic decoupling capacitance (n-well and quiet circuit) and the add-on capacitance [21]. Intrinsic decoupling capacitance alone is not sufficient for acceptable noise suppression in high-performance microprocessor designs. Additional capacitance, often in the form of thin-oxide capacitors, which uses a thin-oxide layer between the n-well and polysilicon gate, is required.

1.4.4.2 Chip-Package Antiresonance

A major problem in combining the chip and package power distribution is chip-package antiresonance, which is explained in detail in a later section. The package inductance and chip decoupling capacitance form a parallel \( RLC \) circuit, which resonates at the frequency\( f' = 1/2\pi\sqrt{LC} \), where \( L \) is the equivalent inductance of the package and \( C \) is the total nonswitching capacitance on-chip between voltage and ground. At this frequency, the power distribution seen by the circuits on the chip has a high impedance. If the chip operating frequency is near or at the chip-package resonant frequency, the noise voltage will be high. A large voltage fluctuation can build up over many cycles at this frequency when excited. In future generations of CMOS microprocessors, large amounts of on-chip decoupling capacitance, which shifts the chip-package resonant frequency well below the operating frequency, must be used to aggressively control switching noise [22].

A typical signature of AC differential noise at the center of a microprocessor operating at 3 GHz that consumes 150 W of power with a 1-V supply is shown in Figure 1-25 [8]. For this illustration, 210 nF of on-chip decoupling capacitance was used along with a low-inductance flip-chip package. The mid-frequency step response occurs when chip power changes abruptly from zero to the maximum power. The magnitude is decreased with on-chip decoupling capacitance. The
oscillation frequency is the chip-package \( LC \) resonance. The mid-frequency noise is eventually damped, resulting in a residual high-frequency AC noise in the steady state. This steady-state response is due to the periodic switching of the microprocessor. The high-frequency steady-state noise rides on a DC offset, which is the \( IR \) drop caused by the chip power distribution resistance.

### 1.4.4.3 Design Challenges

In the past, CMOS active power, which is the power used to do useful work such as switching circuits, has been the main focus of power delivery and management. However, as CMOS scales to 90 nm and below, process-related device leakage current represents a significant passive power component. This passive power includes many sources of device leakage current, such as junction leakage, gate-induced drain leakage, subthreshold channel currents, gate-insulator tunnel currents, and leakages due to defects [23], [24], [25], [26]. Two of these leakage currents, the gate-insulator tunnel current and the subthreshold channel current, are major problems with the scaling of technology. Beyond the 90 nm technology node, more than 50% of the total power can be attributed to leakage, which represents wasted power since no useful work is done.

Gate leakage current can be reduced by using high-K dielectric materials instead of silicon dioxide as the gate dielectric. The subthreshold component of power remains one of the most fundamental challenges as it equals the active component near the 65-nm technology node. This passive power component places a further strain on the on-chip power distribution system because it erodes the DC IR drop noise budget and compounds the electromigration problem.

On-chip voltage islands (logic and memory regions on the chip supplied through separate, dedicated power feeds) is becoming a design approach for managing the active and passive power problems for high-performance designs [26]. In such designs, the voltage level of an island is independent of other islands and is supplied from an off-chip source or on-chip embedded regulators. The design goal is to define regions of circuits within the chip that can be powered by a lower supply voltage while maintaining performance objectives and providing a reduction in active and passive power. Performance-limited critical paths are powered by the maximum voltage that the technology is optimized for, while paths with sufficient timing slack are powered with a lower supply. Hence, transistor libraries with multiple threshold voltages are used. Voltage islands in on-chip power distribution present challenges because isolation of decoupling capacitance reduces its effectiveness for nearby islands. Additional transients due to the activation and deactivation of islands must be managed. The distribution of multiple power supplies complicates the on-chip power grid design and introduces a potential wiring density loss.

1.4.5 PDN with Components

Figure 1-26 shows the complete PDN with the VRM, bypass capacitors, planes, and IC. The proximity of the various components to the IC can be inferred from Figure 1-26(a), where the VRM and bulk capacitors are farthest from the IC and the package planes and package capacitors are much closer. The frequency response of the individual components, their distance from the IC, and the parasitics between the component and the IC dictate the ability of any component to respond to the current demand from the IC. The circuit representation of the core PDN (no signal lines) is shown in Figure 1-26(b). The current flows from the capacitors or VRM through the planes in the package or PCB. The current always flows as a loop, returning to the source through the ground connection.

1.5 Analysis of PDNs

The impedance of a PDN is a parameter that can be calculated in the frequency domain through an AC analysis using a circuit simulator in which a 1-A current
A source is used as the excitation and the voltage at the various nodes of the system are calculated. The computed voltage represents either the self-impedance (voltage measured at the same node as the current excitation) or transfer impedance (voltage measured at a different node). Design of PDNs is always done in the frequency domain, and its analysis in the frequency domain therefore becomes very critical, as explained in the earlier discussion of target impedance. However, power supply noise is a time-domain event, and therefore, the computation of the noise in the time domain on the PDN is also important. Since the frequency response and time response are related to each other, the signature and amplitude of the time-domain response can be controlled by managing the impedance of the PDN in the frequency domain. This approach is very critical for the design of PDNs, since power supply noise in the time domain is a function of the slew rate \(\frac{dI}{dt}\) or, more generally, signature of the current excitation. In other words, different current signatures will always result in varying levels of power supply noise. Because it is difficult to excite all frequencies in the time domain (infinite current signatures would be necessary), a frequency domain analysis is preferable. The frequency domain analysis is explained in this section.

Consider the circuit representation of the PDN described earlier and shown in Figure 1-26. The impedances of such a network can be computed at any node or between any nodes. Two examples are shown in Figure 1-27. In Figure 1-27(a), the IC circuitry is removed, a 1-A current source is connected between the voltage

![Figure 1-26](a) Power delivery with components. (b) Circuit representation. (Courtesy of Professor Joungho Kim, KAIST, South Korea.)
Analysis of PDNs

and ground terminals of the IC, and the voltage at the same node is calculated as a function of frequency. Since a 1-A current source is used, the calculated voltage is the self-impedance (in ohms) seen by the IC looking down into the package toward the VRM. Similarly, in Figure 1-27(b), the board is analyzed separately where a current source of 1 A is connected across a board capacitor (e.g., by removing one of the capacitors in a measurement) and the voltage is measured across the same capacitor and also at the input end of the VRM across a bulk capacitor. The calculated voltage $V_1$ is the self-impedance (in ohms), and voltage $V_2$ is the transfer impedance (also in ohms), and both can be calculated as a function of frequency. Hence, the frequency domain analysis can be used to characterize any part of a PDN as long as the nodes in the network can be probed to connect the current source and measure the voltage. In practical measurements, this analysis is done using a vector network analyzer (VNA), where the $S$-parameters are measured as a function of frequency and can be transformed into impedances (explained in a later section).

The role of the various components in a PDN is shown in Figure 1-28. The y-axis is the self-impedance at a node on the IC, and the x-axis is the frequency. At low kilohertz frequencies, the VRM and bulk capacitor impedance is low and begins to increase with a positive slope on the basis of the ESL of the bulk capacitors.
In the megahertz frequencies, the capacitance of decoupling capacitors on the board reduces the impedance, thus producing a negative slope, which then becomes positive after the resonant frequency of the capacitors. The positive inductive slope caused by the ESL of the capacitors is then compensated by the plane capacitance, which pulls down the impedance through the negative slope shown in the figure up to several hundred megahertz. After the planes resonate, the frequency response of the planes becomes inductive and continues to have a positive slope into the gigahertz range, which is then compensated by the on-chip capacitance. As can be seen in the figure, the impedance fluctuates with multiple resonances (minimum impedance) and antiresonances (maximum impedance). Using the target impedance, the design goal is therefore to ensure that the impedance in Figure 1-28 does not exceed the target impedance at the desired frequencies, which can be for several hundreds of megahertz or several gigahertz.

1.5.1 Single-Node Analysis

A single-node analysis is a simple analysis that can be done using a circuit simulator [7]. The assumption is that any variations in the voltage occur simultaneously across all of the PDN components, and therefore the separation between components is not critical. During the design of a PDN, the single-node analysis is the first step toward meeting the target impedance goals. By placing the VRM in
parallel with all the decoupling capacitors, the number and value of the capacitors can be estimated to meet the target impedance goal.

---

**Example**

To illustrate single-node analysis, consider a plane pair (voltage plane above ground plane) of dimension 250 mm by 250 mm with a dielectric thickness of 0.2 mm. The dielectric constant is FR-4 with a relative permittivity of 4.0. In Figure 1-29, two ports are placed at positions P1 (1.25 mm, 126.25 mm) and P2 (248.75 mm, 126.25 mm). P1 is the input port (current source), and P2 is the output port (voltage measurement). The plane pair is meshed with a unit cell size of 2.5 mm (100 × 100 mesh) and solved using the multilayered finite difference method (M-FDM) described in detail in Chapter 2. The normalized differential voltage between the voltage and ground plane is shown in Figure 1-29(a) as a voltage distribution map at 30 MHz. As can be seen, except at the source, all unit cells are at the same voltage, implying the voltage distribution is constant along the entire surface of the plane. Hence, the plane pair behaves as a lumped capacitor at 30 MHz, and the nodes of any decoupling capacitors connected to these planes will all be at the same potential, irrespective of the capacitor position. The capacitance of the plane pair is 11.067 nF, and the magnitude of its impedance at 30 MHz is $0.479 \Omega (1/[2\pi \times 30 \times 10^6 \times 11.067 \times 10^{-9}])$.

Assuming a VRM of inductance 100 nH is connected to the plane at port P1, the impedance at port P2 is as shown in Figure 1-29(b). The antiresonance is caused by the parallel resonance between the inductance and the plane capacitance at approximately 4.78 MHz.

As explained earlier, when a capacitor is attached to a bare plane, it causes both a resonance and an antiresonance. For a capacitor with ESL = 2.53 nH and $C = 100$ nF, the resonance occurs at 10 MHz, while the antiresonance occurs at 31.7 MHz, according to equation (1.24). Both the resonance and antiresonance are shown in Figure 1-29(c). As can be seen from the figure, though the antiresonance is close to 31.7 MHz, there is a shift in resonance to a slightly lower frequency of approximately 30 MHz. This shift can be attributed to the plane inductance. Although the plane inductance is small compared to the ESL of the capacitor, it does have a small effect in shifting both the resonance and antiresonance frequencies. As the separation between the voltage and ground planes increases, the
Figure 1-29  Single node analysis.  (a) Voltage distribution on a plane pair at 30 MHz.  (b) VRM + plane.
corresponding increase in the plane inductance causes a larger shift in the resonance and antiresonance frequencies. This shift is shown in Figure 1-29(d), where the antiresonance frequency for a capacitor with ESL = 2.53 nH and C = 100 nF mounted on a plane with dielectric separation from 0.05 mm to 0.4 mm is shown using equation (1.24) and the result of analysis using M-FDM. The error is 16% for a 0.4-mm separation and decreases to an
error of 3% for a 0.05-mm separation. In Figure 1-29(d), the impedance is measured at port P2.

With the plane inductance included, the resonance and antiresonance frequencies are modified as follows:

\[
f_{\text{resonance}} = \frac{1}{2\pi \sqrt{(L + L_p)C}}
\]

\[
f_{\text{antiresonance}} = \frac{1}{2\pi \sqrt{(L + L_p) \frac{C_p C}{C + C_p}}}
\]

(1.31)

where \( L \) is the ESL of the capacitor, \( C \) is its capacitance, \( L_p \) is the plane inductance, and \( C_p \) is the plane capacitance.

1.5.1.1 Calculation of Plane Inductance at Low Frequencies

At low frequencies in the 1 to 50 MHz range, the plane can be represented using a lumped T-model, as shown in Figure 1-30(a) between ports P1 and P2. With a capacitor placed at port P1, the impedance seen at port P2 can be calculated. The T-model can be derived using a plane solver such as M-FDM by computing the two-port impedance between ports P1 and P2 at low frequencies. As an example, in Figure 1-29(a), the two-port Z-parameters at 10 MHz are \( Z_{11} = -j1.4122 \), \( Z_{12} = -j1.44089 \) and \( Z_{22} = -j1.4122 \). The inductance and capacitance of the T-model can then be derived from

\[
L_1 = \frac{(Z_{11} - Z_{12})}{2\pi f}
\]

\[
C = \frac{1}{2\pi f Z_{12}}
\]

\[
L_2 = \frac{(Z_{22} - Z_{12})}{2\pi f}
\]

(1.32)

leading to an inductance of 0.456 nH and capacitance of 11.04 nF. The variation of the inductance with plane separation for the structure in Figure 1-29(a) is shown in Figure 1-30(b), where the inductance increases almost linearly as the plane separation increases, as expected from equation (1.22).
1.5.1.2 Meeting Target Impedance at Frequencies below Cavity Resonance

Cavity resonance represents the frequencies at which the planes resonate, as described by equation (1.28). At frequencies below cavity resonance, the placement of capacitors is not critical and hence they can be uniformly distributed across the plane to meet the target impedance.

As an example, let’s assume that the target impedance required is 20 mΩ from DC to 30 MHz, which is not satisfied in Figure 1-29(b) because the impedance exceeds 20 mΩ for frequencies larger than 30 KHz. The goal is therefore to choose capacitors that resonate at frequencies at which the impedance must be reduced to meet the target impedance. As the capacitors are connected to the planes, antiresonances are created at frequencies given by equation (1.31), which we’ll call LC antiresonance. The capacitors therefore have to be chosen such that they minimize the impedance at the antiresonance frequencies. The capacitors...
available to reach the 20-mΩ target impedance are shown in Table 1-2 along with their ESR, ESL, and resonant frequency values.

### Table 1-2 Decoupling Capacitors

<table>
<thead>
<tr>
<th>Capacitance (F)</th>
<th>ESL (nH)</th>
<th>ESR (mΩ)</th>
<th># of Capacitors</th>
<th>Resonant Frequency (Hz)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.00 × 10⁻²</td>
<td>1</td>
<td>23</td>
<td>2</td>
<td>5.03 × 10⁴</td>
</tr>
<tr>
<td>5.00 × 10⁻⁵</td>
<td>1</td>
<td>5</td>
<td>1</td>
<td>7.11 × 10⁵</td>
</tr>
<tr>
<td>2.20 × 10⁻⁵</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1.07 × 10⁶</td>
</tr>
<tr>
<td>1.00 × 10⁻⁵</td>
<td>1</td>
<td>3.1</td>
<td>1</td>
<td>1.59 × 10⁶</td>
</tr>
<tr>
<td>4.70 × 10⁻⁶</td>
<td>1</td>
<td>4.7</td>
<td>1</td>
<td>2.32 × 10⁶</td>
</tr>
<tr>
<td>2.20 × 10⁻⁶</td>
<td>1</td>
<td>2.2</td>
<td>1</td>
<td>3.39 × 10⁶</td>
</tr>
<tr>
<td>1.00 × 10⁻⁶</td>
<td>1</td>
<td>15.4</td>
<td>1</td>
<td>5.03 × 10⁶</td>
</tr>
<tr>
<td>4.70 × 10⁻⁷</td>
<td>1</td>
<td>23.3</td>
<td>2</td>
<td>7.34 × 10⁶</td>
</tr>
<tr>
<td>2.20 × 10⁻⁷</td>
<td>1</td>
<td>35.3</td>
<td>2</td>
<td>1.07 × 10⁷</td>
</tr>
<tr>
<td>1.00 × 10⁻⁷</td>
<td>1</td>
<td>67.1</td>
<td>4</td>
<td>1.59 × 10⁷</td>
</tr>
<tr>
<td>4.70 × 10⁻⁸</td>
<td>1</td>
<td>99.7</td>
<td>5</td>
<td>2.32 × 10⁷</td>
</tr>
<tr>
<td>2.20 × 10⁻⁸</td>
<td>1</td>
<td>148.5</td>
<td>8</td>
<td>3.39 × 10⁷</td>
</tr>
</tbody>
</table>

At the resonant frequency, the capacitor has the minimum impedance. Hence, the number of capacitors of each type required to meet the target impedance at the resonant frequency can be calculated using

$$
\text{# Capacitors} = \frac{ESR}{Z_T} = \frac{ESR}{20\text{mΩ}}
$$

Using equation (1.33), the number of each capacitor required is shown in Table 1-2. These capacitors can be placed and connected to the plane at any arbitrary position, since location does not matter. By connecting the capacitors in parallel in a circuit simulator, as shown in Figure 1-31(a), the impedance response similar to Figure 1-31(b) can be obtained. In Figure 1-31(a), multiple capacitors of the same type in parallel have been replaced with a single capacitor containing scaled values for $ESR$, $ESL$, and $C$. For example, two capacitors in parallel with $C = 10$ mF, $ESL = 1$ nH, and $ESR = 23$ mΩ have been replaced with a single capacitor with $C' = 20$ mF, $ESL' = 0.5$ nH, and $ESR' = 11.5$ mΩ without changing
Analysis of PDNs

its resonant frequency. The impedance in Figure 1-31(b) was obtained using the M-FDM method in which the 20 mΩ target impedance is not exceeded up to 95 MHz. The inductive slope in Figure 1-31(b) is caused by the parallel combination of the ESL of all the capacitors. With 29 capacitors in parallel, each with an inductance of 1 nH, an equivalent inductance of 34.48 pH results, which corresponds to the impedance at ~100 MHz.

A practical example is shown in Figure 1-32 [7]: the VRM plus 144 capacitors of various types in parallel result in a frequency response that meets the target impedance up to 50 MHz. Beyond this frequency, the impedance becomes inductive (positive slope), as shown in the figure.

1.5.2 Distributed Analysis

As the frequency extends beyond 50 MHz, the voltage distribution on the plane changes and standing waves are generated. From equation (1.28), a 250 mm by 250 mm plane resonates at 300 MHz (1,0 mode), 423 MHz (1,1 mode), 600 MHz (2,0
mode), 670 MHz (2,1 mode), 847 MHz (2,2 mode), 900 MHz (3,0 mode), and so on. A mode represents a voltage distribution. The voltage distribution for some of these resonant modes is shown in Figure 1-33; both the maximum voltage and minimum voltage are shown as a function of position on the planes. As can be seen, the plane is no longer an equipotential surface. At 300 MHz, a half wave stands along the length; at 600 MHz, a half wave stands along the length and width; at 670 MHz, a full wave stands along the width while a half wave stands along the length; and at 847 MHz, a full wave stands along the length and width. A full wave represents a standing wave with a 360-degree phase reversal, while a half wave represents a 180-degree phase reversal. For example, at 300 MHz, the maxima (plotted as magnitude) occurs at the near and far edge with a minima at the center, representing a 180-degree phase shift (phase not shown) and representing half of a standing wave. At frequencies at which the planes resonate, the location of the decoupling capacitors becomes important because they need to be located at a voltage maximum (or at the source) to reduce the bounce on the planes.

Figure 1-32  Single-node analysis to meet target impedance—a practical example. By permission from L.D. Smith, et al. [7], © 1999 IEEE.
Example

Consider Figure 1-33(a): the voltage peaks have to be reduced at 300 MHz. To monitor the voltage maximum, port 3 is placed at (7.5 mm, 7.5 mm). With a current source excitation of 1 A at port 1, the magnitude of the voltage measured at port 3 is plotted in Figure 1-34 as a function of frequency. An impedance peak can be seen at 300 MHz with an impedance magnitude of ~3 Ω. To reduce this peak, a capacitor with a resonant frequency of 300 MHz must be placed at (7.5 mm, 7.5 mm). The voltage distribution resulting from a capacitor with $C = 1 \text{nF}$, $ESR = 10 \text{mΩ}$, and $ESL = 0.28 \text{nH}$ placed at (7.5 mm, 7.5 mm) is shown in Figure 1-35(a). Clearly, the voltage distribution on the plane changes with voltage minima occurring at the position where the capacitor is placed. The impedance at port...
3 with a current source excitation of 1 A at port 1 is plotted in Figure 1-35(b). The impedance peak of ~3 Ω has been reduced significantly. It is important to note that the placement of the capacitor at the center (125 mm, 125 mm) will not reduce the voltage variation on the plane at 300 MHz because, as
per Figure 1-33(a), this represents a voltage minima point. Hence, the placement of the capacitors on the plane becomes very important, requiring a distributed analysis.

### 1.5.2.1 Meeting Target Impedance at Cavity Resonances

Meeting target impedance at cavity resonances is more tricky than at frequencies below cavity resonances because capacitors that resonate at higher frequencies are required. The principle of choosing the right capacitor and the number of capacitors required is similar to that of single-node analysis in which the capacitors are chosen such that they resonate at the cavity resonance frequency and the number of capacitors are chosen on the basis of equation (1.33). The primary difference is the placement of the capacitor.

In any computer system, the goal is always to minimize plane bounce, since any voltage variation on the planes can couple to signal lines referenced to it. To minimize voltage variations through the placement of capacitors, consider as an example the 250 mm by 250 mm plane described earlier. As before, the position of ports 1 and 2 remains the same. In addition, nine ports are uniformly added across the surface of the planes to monitor the voltage fluctuation. The source is at port 1, and the goal is to minimize the voltage fluctuations at the remaining ports. As discussed earlier, the first cavity resonance occurs at 300 MHz. Below this frequency, capacitors can be uniformly distributed using single-node analysis such that the target impedance is met. The transfer impedance between the source (port 1) and the remaining ports is shown in Figure 1-36(a), where the impedance of 20 mΩ has been met at most frequencies between DC and 300 MHz using single-node analysis and by uniformly distributing capacitors on the plane. Also shown is the first cavity resonance, which has shifted from the original frequency of 300 MHz to 340 MHz because of the loading of the planes by the ESL of the capacitors.

At the cavity resonance of 340 MHz, the transfer impedance magnitude is not the same at the 11 ports. Hence, to reduce the impedance to meet target impedance, capacitors that resonate at 340 MHz have to be chosen with the number decided by equation (1.33). These capacitors have to be placed at the position next to where the transfer impedance needs to be reduced. The transfer impedances obtained by placing capacitors with $ESR = 10 \text{ mΩ}$, $ESL = 1 \text{ nH}$, and $C = 0.219 \text{ nF}$ next to the appropriate ports is shown in Figure 1-36(b), where the impedance at 340 MHz has been reduced. However, new antiresonances have been created near
280 MHz because of parallel resonance. The parallel resonance now must be reduced by using capacitors that resonate at 280 MHz.

Hence, the methodology for the placement of capacitors should include an initial single-node analysis followed by distributed analysis whereby the impedances are iteratively minimized across the plane surface to minimize plane bounce. This methodology can be extended to packages or boards containing multiple plane layers.

Figure 1-36 (a) Transfer impedance between ports. (b) Transfer impedance between ports with first cavity resonance suppressed.
1.6 Chip-Package Antiresonance: An Example

The interaction between the chip and package can cause a large antiresonance in the impedance profile of a chip, as described briefly earlier. Large voltage fluctuations can occur if the operating frequency of the chip coincides with the chip-package antiresonance frequency. This effect is best illustrated through an example. Consider a multilayered board of dimension 30 cm by 25 cm, as shown in Figure 1-37(a). The board contains a VRM at one corner with a number of decoupling capacitors. The package assembled on the board is a laminate package with cross section, as shown in Figure 1-37(b), and is 40 mm by 40 mm in size. The package contains an 800-μm core with four sequential build-up layers on either side of the core. The package is assembled on the board through solder balls. The chip of size 15 mm by 15 mm is mounted on the package through C4 solder bumps.

Consider the core PDN where the VRM supplies current to the core transistors of the IC. The core PDN can be represented using an equivalent circuit, as shown in Figure 1-38, where the 1-V VRM is assumed to be represented as an ideal power supply. The two-port impedance between the VRM and bottom of the package is a 2 × 2 matrix with elements $Z_{11}$, $Z_{12}$, and $Z_{22}$, which represent the self-impedance at the VRM, transfer impedance, and self-impedance at the bottom of the package (assumed as a single node at the center of the package). This 2 × 2 matrix can be represented as a T-network with elements $Z_{11} - Z_{12}$, $Z_{12}$, $Z_{22}$.

![Figure 1-37](image-url) (a) Board showing package location. (b) Chip on package.
Z₂₂ – Z₁₂, which are all functions of frequency, as shown in Figure 1-38. This circuit representation is one of many possible ways to represent a two-port network. The impedance Z_{board}(f) at the bottom of the package can be calculated from Figure 1-38 as

\[ Z_{board}(f) = \frac{Z_{11}(f)Z_{22}(f) - Z_{12}(f)^2}{Z_{11}(f)} \]  

(1.34)

The impedance at the bottom of the package looking toward the VRM is plotted in Figure 1-39. The frequency response assumes a certain number of decoupling capacitors mounted on the board at various locations. As expected, the impedance has resonances and antiresonances due to the resonant frequency of the capacitors and becomes inductive beyond 100 MHz. Let’s assume for simplicity that the package is represented through an equivalent inductance whose impedance is represented as Z_{package} in Figure 1-38. The inductance of such a package can be calculated approximately by separating the package into layers consisting of solder bumps, sequential build-up layers, core, and solder balls, then modeling the inductance of the bumps, vias, and balls. This is a good approximation, since the current flows vertically through such multilayered ball grid array packages. A tool called FastHenry [27], available in the public domain, has been used here to generate Table 1-3. In the table, each power/ground bump/via/ball pair has been modeled as a cylinder to extract the self- and mutual partial inductance. The loop inductance for a pair is then calculated using \( L_{loop} = L_{self} + L_{self} - 2M \) (since current flows in opposite directions for a power/ground loop). The equivalent inductance for a number of these pairs in parallel is then computed by dividing the loop inductance by the total number of pairs present in the package.
As can be seen from Table 1-3, the largest contributors to package inductance are the core and the solder balls. Adding the equivalent loop inductances (in series), the package inductance from the solder ball to the C4 bump can be estimated as 3.37 pH. Since Table 1-3 is an approximate method for calculating the equivalent loop inductance, let’s assume that the package inductance is ~4 pH. The impedance of a 4-pH inductor is plotted in Figure 1-39, which is a straight line with positive slope on a log-log graph. The total impedance from the top of the package looking into the VRM can now be calculated from Figure 1-38 as

\[ Z_{pkg+board}(f) = Z_{board}(f) + Z_{package}(f) \]  

(1.35)

The plot of equation (1.35) is shown in Figure 1-39. At low frequencies (<10 MHz), the impedance is dominated by the board, while above 50 MHz, the inductance of the package dominates and the inductance is primarily due to the package inductance. The chip can be represented as a capacitor. Let’s assume that the IC capacitance is 500 nF. The impedance of the chip capacitor is plotted in Figure 1-39. The impedance seen from the IC looking toward the VRM can be calculated from Figure 1-38 as
since the impedances are in parallel.

This calculation is plotted in Figure 1-39. As can be seen from the figure, an antiresonance occurs at around 100 MHz with a corresponding large impedance. This large impedance is caused by chip-package antiresonance. From Figure 1-39, the chip-package antiresonance occurs at the frequency where $Z_{chip}$ and $Z_{package}$ intersect. Hence, the chip-package antiresonance is caused by the parallel resonance of the chip capacitance and package inductance. The board inductance has little to do with causing the antiresonance.

The importance of the chip, package, and board for managing the PDN impedance is shown qualitatively in Figure 1-40. At frequencies from DC to 10 MHz, the board design becomes very critical for managing the target impedance with no contribution coming from either the IC or package. Between 10 and 100 MHz, the interaction between the package and board becomes important. It can cause a shift in the resonance and increase the overall inductance. From 100 MHz to 1 GHz, the interaction between the chip and package can cause an impedance peak due to

\[
Z_{total}(f) = Z_{pkg+board}(f) \parallel Z_{chip}(f)
\]  

\[(1.36)\]
the antiresonance between the package inductance and IC capacitance. This is the frequency range at which chip–package co-design becomes critical. Beyond 1 GHz, the PDN impedance is dictated solely by the IC capacitance with neither the package nor the board having any influence.

1.7 High-Frequency Measurements

High-frequency measurements represent a very important part of evaluating the effectiveness of decoupling capacitors, planes, and PDNs with surface-mount components. The magnitude and slope of change in resistance and inductance above the self-resonant frequency (SRF) depend on the relative dimensions of the capacitor body, pads, vias, and closest planes. Since $ESR(f)$ and $ESL(f)$ depend on the geometry, it is important to measure the parts in a fixture containing pads, vias, and planes with geometries similar to actual usage to ensure that the measured complex impedance reflects both the device under test (DUT) and, if necessary, the fixture. The low impedance values associated with today’s bypass capacitors can be conveniently measured using vector network analyzers (VNA) in two-port connections [28].

The VNA consists of a tunable sinusoidal source and tracking receiver. The receiver can be connected to measure the incident wave ($a_i$) or reflected wave ($b_i$) to and from the DUT, or the transmitted wave ($b_2$) through the DUT. These waves are measured by measuring voltages. The VNA measures and displays the various combinations of incident and reflected waves, commonly known as $S$-parameters [29]. In a VNA measurement, the equivalent circuit can be represented as shown.
in Figure 1-41. In the figure, $Z_0$ is the connecting impedance of the instruments, usually 50 Ω. A state-of-the-art VNA may have around one-tenth of a decibel repeatability and a fraction of a decibel absolute error in $S_{11}$ measurements, which allows the accurate measurements of impedances not much smaller or greater than 50 Ω. The impedance of a one-port DUT can be calculated from its voltage reflection coefficient ($S_{11}$) by using this formula:

$$Z_{DUT} = \left(\frac{1 + S_{11}}{1 - S_{11}}\right)Z_0$$

(1.37)

The impedances of PDNs are typically low, much smaller than 50 Ω. With very low impedances, the nominator in equation (1.37) approaches zero, and therefore the accuracy gradually breaks down. As shown in [30], the two-port impedance measurement setup can extend the capabilities of one-port vector network analyzer measurements by more than an order of magnitude, enabling the measurement of PDN impedances below a milliohm. Hence, a two-port measurement is always preferable while characterizing PDNs.

1.7.1 Measurement of Impedance

The equivalent circuit of a PDN impedance measurement is shown in Figure 1-42. In the figure, $L_{p1}$ and $L_{p2}$ represent the inductive discontinuities of the cable–DUT interface, including the loop inductance due to the current loop from the signal tip to the ground tip of a probe [31]. The PDN is represented as a two-port network with impedances $Z_{11}$, $Z_{12}$, $Z_{21}$, and $Z_{22}$ associated with it, as defined earlier. The
reference planes are set as shown in Figure 1-42, assuming that both the connectors (or probes) are calibrated up to the DUT interface. The DUT includes the entire PDN; hence the goal is to extract the PDN impedances between the ports of interest. Assuming a two-port measurement, the measured $S$-parameters can be related to the $Z$-parameters of the PDN through the transformation

$$
\begin{bmatrix}
S_{11} & S_{12} \\
S_{21} & S_{22}
\end{bmatrix} = \begin{bmatrix}
Z_{11} + j\omega L_{p1} + Z_0 & Z_{12} \\
Z_{21} & Z_{22} + j\omega L_{p2} + Z_0
\end{bmatrix}^{-1} \begin{bmatrix}
Z_{11} + j\omega L_{p1} - Z_0 & Z_{12} \\
Z_{21} & Z_{22} + j\omega L_{p2} - Z_0
\end{bmatrix}
$$

(1.38)

where $Z_0 = 50$ Ω. Setting $Z_1 = 50 + j\omega L_{p1}$ and $Z_2 = 50 + j\omega L_{p2}$, the above equation can be rewritten in the form

$$
\begin{bmatrix}
S_{11} & S_{12} \\
S_{21} & S_{22}
\end{bmatrix} = \frac{1}{(Z_{11} + Z_1)(Z_{22} + Z_2) - Z_{12}Z_{21}} \begin{bmatrix}
Z_{22} + Z_2 & -Z_{12} \\
-Z_{21} & Z_{11} + Z_1
\end{bmatrix} \begin{bmatrix}
Z_{11} + j\omega L_{p1} - Z_0 & Z_{12} \\
Z_{21} & Z_{22} + j\omega L_{p2} - Z_0
\end{bmatrix}
$$

(1.39)

Since $S_{12}$ measurement is preferable for low impedance measurements, from equation (1.39), the measured $S_{12}$ parameter can be related to the impedance as

$$
S_{12} = \frac{2Z_{12}Z_0}{(Z_{11} + Z_1)(Z_{22} + Z_2) - Z_{12}Z_{21}}
$$

(1.40)
1.7.2 Measurement of Self-Impedance

Assuming the self-impedance of a PDN is to be measured, the two connectors (or probes) can be placed next to each other such that $Z_{11} = Z_{12} = Z_{21} = Z_{22}$. Under such a measurement, the equivalent circuit simplifies to the circuit shown in Figure 1-43. Then equation (1.40) simplifies to

$$S_{12} = \frac{2Z_{11}Z_0}{Z_{11}(Z_1 + Z_2) + Z_1Z_2}$$

(1.41)

from which the self-impedance can be extracted from the measured $S$-parameter as

$$Z_{11} = S_{12} \frac{Z_1Z_2}{2Z_0} \frac{1}{1 - S_{12}} \frac{1}{2Z_0} \left( \frac{Z_1 + Z_2}{2Z_0} \right)$$

(1.42)

Assuming the inductive discontinuities are negligible, resulting in a nearly 0 dB through measurement, the self-impedance simplifies to

$$Z_{11} = S_{12} \times \frac{50 \times 50}{2 \times 50} \times \frac{1}{1 - S_{12}} = 25 \times \frac{S_{12}}{1 - S_{12}}$$

(1.43)

Although equation (1.42) is an exact expression, inductances $L_{p1}$ and $L_{p2}$ result in very small perturbation to the magnitude of $Z_{11}$ in equation (1.43), which is illustrated in the following example. Equation (1.43) is especially useful for PDNs that have a very small impedance since $S_{11}$ measurements are not very reliable.

![Figure 1-43](image-url)
Example

Assuming \( L_{p1} = L_{p2} = 0.5 \, \text{nH} \), and the measured \( S_{12} = 0.1 \angle 60^\circ \) (–20 dB) at a frequency of 1 GHz, using equation (1.42) to calculate magnitude of \( Z_{11} \) introduces a 0.4% difference in the measurements as compared to equation (1.43).

When small impedance values are measured with a 50-Ω system, \( S_{12} \) is much less than 1 and therefore a good approximation for the self-impedance measurement is

\[
Z_{11} \approx 25 \times S_{12}
\] (1.44)

The insertion loss \( S_{12} \) is typically measured in decibels and hence must be converted to the appropriate scale. Using equation (1.42), (1.43), or (1.44), accurate measurements are possible for small values of impedances typical of PDNs provided the measurement setup is well calibrated up to the probe tips. A simple calibration method is the short-open-load-through (SOLT) calibration.

Example

Consider a capacitor embedded in the package, discussed in detail in Chapter 5. The capacitor consists of two square electrodes with size in the range 2 mm to 10 mm separated by a dielectric layer. For such a structure, a two-port measurement can be conducted to extract the impedance of the structure, as shown in Figure 1.44(a). Using equation (1.43), by separating the real and imaginary parts, the impedance of the capacitor can be extracted as

\[
\begin{align*}
\text{Re}(Z_{11}) &= 25 \times \frac{\text{Re}(S_{21}) \times [1 - \text{Re}(S_{21})] - [\text{Im}(S_{21})]^2}{[1 - \text{Re}(S_{21})]^2 + [\text{Im}(S_{21})]^2} \\
\text{Im}(Z_{11}) &= 25 \times \frac{\text{Im}(S_{21})}{[1 - \text{Re}(S_{21})]^2 + [\text{Im}(S_{21})]^2}
\end{align*}
\] (1.45)

The measured impedance magnitude for four different capacitor sizes is shown in Figure 1-44(b). The low-frequency capacitance, ESR, and ESL can be seen in the measurements.
1.7.3 Measurement of Transfer Impedance

The transfer impedance $Z_{12}$ between two ports placed far apart can be extracted from equation (1.40) as

$$
Z_{12} = S_{12} \frac{Z_1 Z_2}{2Z_0} \frac{(1 + \frac{Z_{11}}{Z_1})(1 + \frac{Z_{22}}{Z_2})}{1 + \frac{S_{12} Z_{21}}{2Z_0}}
$$

(1.46)

where $Z_1$ and $Z_2$ have the same definitions as in the self-impedance measurements, while $Z_{11}$, $Z_{22}$, and $Z_{12}$ are the two-port impedance parameters of the structure. Assuming $L_{p1} = L_{p2} = 0$, equation (1.46) can be simplified to

$$
Z_{12} = S_{12} \times \frac{50 \times 50}{2 \times 50} \frac{(1 + \frac{Z_{11}}{50})(1 + \frac{Z_{22}}{50})}{1 + \frac{Z_{21} S_{12}}{2 \times 50}}
$$

(1.47)

For low impedance measurements $S_{12} Z_{21}/100 \ll 1$, $Z_{11}/50 \ll 1$, and $Z_{22}/50 \ll 1$. Therefore, the transfer impedance measurement can be simplified similar to equation (1.44). Several of the transfer impedance measurements in the book make use of the simplified equation in (1.44).

1.7.4 Measurement of Impedance by Completely Eliminating Probe Inductance

For PDNs with a moderate impedance and measurements involving a high probe inductance, a different methodology can be used. Let’s once again assume that the $2 \times 2$ impedance matrix, including the probe parasitics, can be approximated as

$$
Z = \begin{bmatrix}
Z_{11} + j\omega L_{p1} & Z_{12} \\
Z_{21} & Z_{22} + j\omega L_{p2}
\end{bmatrix}
$$

(1.48)

This equation shows that the transfer impedance is not affected by the probe inductance. Hence, whatever the value of $L_{p1}$ and $L_{p2}$ is, $Z_{12}$ can be extracted very accurately, assuming the VNA is calibrated up to the probe tips. This property is...
Signal Lines Referenced to Planes

used in Chapter 5, where the material properties are extracted from $Z_{12}$. The $2 \times 2$ impedance matrix in equation (1.48) can be obtained from measured $S$-parameters.

For accurate measurement of an input impedance, say $Z_{11}$, the probe parasitic inductances must be known. However, there is a simple way of avoiding this problem as well. A separate measurement is taken, where the two probes are placed very close to this input port. Since the two ports are very close to each other, the input and transfer impedances of such a network can be assumed to be equal ($Z_{11} = Z_{12} = Z_{21} = Z_{22}$), as mentioned earlier. Hence, once again, the transfer impedance $Z_{12}$ can be measured, which is not affected by the probe inductance. Since the input and transfer impedances of this DUT are equal, the input impedance $Z_{11}$ can be found to be equal to the measured transfer impedance $Z_{12}$. This procedure provides accurate measurement results for PDNs in the presence of high probe inductance.

1.8 Signal Lines Referenced to Planes

In high-frequency packages and PCBs, the signal lines are often routed over or between planes. The planes provide shielding for the signal lines by minimizing coupling between the signal lines routed on the same layer or on different layers. In addition, the planes perform a very important function by carrying return current of the signal lines. Any discontinuities in the return current of the signal lines can cause SSN. Hence, to model SSN arising from the switching of I/O drivers, a good knowledge of the return path discontinuities is necessary. Placement of the capacitors at the return path discontinuities reduces SSN. In this section, the
transmission line equations are derived along with a discussion, using two examples, on the importance of return currents.

1.8.1 Signal Lines as Transmission Lines

Signals cannot propagate faster than the speed of light. The velocity of $3 \times 10^8$ m/s in air is roughly equivalent to a time delay of 1 ns per foot of travel. In FR-4 (a common dielectric material used to fabricate PCB), a signal takes 1.66 ns to travel a distance of 10 inches. This delay can be an entire clock period behind and hence can slow down the system. Analyzing such signal lines using Kirchhoff’s voltage and current laws neglect this time delay and can provide erroneous results. Early ICs involved transistor–transistor logic (TTL) devices with internal delays of 15 ns or more, limiting the clock rates to lower levels and making the transit time of the signals negligible in comparison. Modern devices have improved to the point that signal delays are the limiting factor in digital circuit design. Transmission line theory in which signal lines are analyzed as distributed circuits explicitly includes this time delay and therefore is applicable to these situations.

To better explain transmission line theory, consider a coaxial cable, as shown in 1-45(a), consisting of two conductors separated by a dielectric material. When current flows on the coaxial cable, there is a physical movement of charge carriers (electrons) down one conductor and back on the other. If there is a current flow into the inner conductor, then the return current flows in the opposite direction on the outer conductor. Because of this movement, the charge has some momentum, and the current therefore wants to keep moving once it has started. This effect is equivalent to some series inductance in the cable.

There is also some series resistance, since the metal is not a “perfect” conductor of electricity, and some of the electrical energy is therefore converted to heat as the current flows. At the same time, equal and opposite charge is stored instantaneously on the two conductors, giving rise to some shunt capacitance. If the material separating the conductors is not a perfect insulator, there will also be some leakage current from one conductor to another, which is the shunt conductance [32].

Although coaxial cables are excellent transmission lines, since the return current is always in close proximity to the forward current, it is difficult to use them in ICs, packages, and boards because of their three-dimensional structure. Hence, planar structures such as microstrip and striplines are used where the reference conductor is in the form of planes. The planes carry the return current, so it is important to locate the reference plane in close proximity to the signal conductor. An example of a microstrip line is shown in Figure 1-45(b).
Since the cross section of a transmission line is uniform and its length is much larger than its cross section, it can be described using the pul resistance \( (R) \), inductance \( (L) \), conductance \( (G) \), and capacitance \( (C) \) parameters. A small section of a transmission line can be represented using the equivalent circuit shown in Figure 1-46, where the pul parameters \( R, L, G, \) and \( C \) are each multiplied by the length \( \Delta z \). The pul inductance \( L \) is the loop inductance between the signal line and the reference plane, where the forward current on the signal line returns in the opposite direction through the reference plane. The pul resistance \( R \) in Figure 1-46 is the sum total of the signal and reference plane resistance. The pul capacitance \( C \) and conductance \( G \) are measured between the singal and reference conductors.

Applying Kirchoff’s voltage and current laws to the circuit in Figure 1-46, the following two equations can be obtained:

\[
\begin{align*}
v(z + \Delta z, t) + L \Delta z \frac{\partial}{\partial t} i(z, t) + R \Delta z l(z, t) &= v(z, t) \\
i(z + \Delta z, t) - i(z, t) &= -G \Delta z v(z + \Delta z, t) - C \Delta z \frac{\partial v(z + \Delta z, t)}{\partial t}
\end{align*}
\]

(1.49)

![Figure 1-46](image-url) Transmission-line equivalent circuit.
Under the limit \( \Delta z \to 0 \), the transmission line equations can be written in the form

\[
\begin{align*}
\frac{\partial v}{\partial z} &= -Ri - L \frac{\partial i}{\partial t} \\
\frac{\partial i}{\partial z} &= -Gv - C \frac{\partial v}{\partial t}
\end{align*}
\]

(1.50)

Equation (1.50) can be extended to multiple coupled lines and form the multiconductor transmission line equations.

### 1.8.2 Relationship between Transmission-Line Parameters and SSN

Two important parameters that describe a transmission line are its characteristic impedance \( (Z_0) \) and delay \( (T) \). For a lossless line \( (R = 0 \text{ and } G = 0) \), these parameters are given by

\[
Z_0 = \sqrt{\frac{L}{C}} \quad (\Omega)
\]

\[
T = \sqrt{LC}l \quad (s)
\]

(1.51)

where \( l \) is the length of the transmission line. The SSN has an inverse relationship with \( Z_0 \), as described by equation (1.10). A transmission line with high characteristic impedance will always require less current to charge the line and hence a smaller \( dI/dt \), which translates into lower SSN. A transmission line with low \( Z_0 \) is capacitive and hence requires a larger current, resulting in larger SSN. Figure 1-47(a) shows the variation of the SSN as a function of \( \tau_t/(L/Z_0) \) based on equation (1.9). From the figure, for \( \tau_t = 0.1 \text{ ns} \) and \( L = 1 \text{ nH} \), a 50-\( \Omega \) line will result in an SSN voltage of 0.2 V as compared to 0.1 V for a 100-\( \Omega \) line.

Noise on the power supply always results in an extra delay for the signal to rise to the required voltage in addition to the delay of the transmission line, as described by equations (1.12) and (1.13). Therefore, timing errors can occur when the SSN is large, which in turn is related to the \( Z_0 \) of the transmission line. Figure 1-47(b) shows the variation of the 50% delay as a function of \( \tau_t/(L/Z_0) \) for a bus, based on equation (1.12) when \( \tau_t \) is greater than \( L/Z_0 \). From the figure, for \( \tau_t = 0.1 \text{ ns} \) and \( L = 1 \text{ nH} \), a 10 bit wide bus with impedance of 50 \( \Omega \) will result in a delay of 0.191 ns as compared to 0.1235 ns for a bus with impedance of...
100 Ω. Hence, choosing the right characteristic impedance for the signal lines becomes very important for minimizing SSN and timing error.

### 1.8.3 Relationship between SSN and Return Path Discontinuities

To illustrate the relationship between SSN and return path discontinuities for signal lines, two examples are discussed here. Figure 1-48(a) consists of a microstrip line over a voltage and ground plane. The signal line is referenced to the voltage plane, and the plane is continuous. A 1-A current source is applied between the input end of the signal line and the voltage plane that creates a forward current on the signal line, as shown in the figure. The return current flows on the voltage plane just beneath the signal line (at high frequencies), so the current loop is completed between the signal line and the voltage plane. This return current does not create any voltage disturbance between the voltage and ground plane, as
illustrated in Figure 1-49(a), where a two-dimensional plot of the voltage fluctuation between the two planes is shown as a function of position, indicating that there is no change in the voltage across the planes. Any voltage measured between the two planes, as in Figure 1-48(a), will show no coupling between the signal line and voltage/ground plane.

Now consider a slot on the voltage plane beneath the signal line, as shown in Figure 1-48(b). With a 1-A current source excitation between the signal line and voltage plane, the forward current on the signal line causes a return current on the voltage plane. At the slot, the return current flows on the ground plane (due to the absence of metal on the voltage plane). Therefore, the current loop is completed by the vertical currents shown in Figure 1-48(b) (also called displacement currents) and is a return path discontinuity that excites the voltage/ground plane and causes a voltage disturbance. The voltage disturbance is shown in Figure 1-49(b) as a two-dimensional plot for a 250 mm by 250 mm plane with a 50-mm slot at the center. The excitation of the voltage/ground plane causes the $f_{01}$ mode based on equation (1.29) at a frequency of about 750 MHz for a dielectric material with relative permittivity of 4.0 between the planes. Any voltage measured between the two planes, as in Figure 1-48(b), will now show significant coupling between the
PDN Modeling Methodology

The design process that captures the modeling methodology described in this book is shown in Figure 1-50. The entire methodology is centered around plane modeling because as packages and boards migrate toward gigahertz frequencies, planes play a very important role by reducing inductance and supporting

**Figure 1-49** (a) Voltage disturbance on power/ground plane for ideal microstrip. (b) Voltage disturbance on power/ground plane for microstrip above slot.

A capacitor with a resonant frequency of 750 MHz placed between the voltage and ground planes near the slot will provide a low impedance path for the current to return on the planes and therefore will reduce the voltage disturbance on the planes.

In summary, SSN caused by the switching of signal lines in the package and board is caused by the return path discontinuities. For high-frequency system applications, the referencing of signal lines to planes is very important. The return path discontinuity for signal lines can be evaluated by following the return current path on the planes. Any discontinuity will manifest itself as voltage fluctuations between voltage and ground. It is important to note that the connection of the drivers and terminations to the signal lines can alter the power supply noise, since the current loop can change at the input and output, as described in Chapter 3.

**1.9 PDN Modeling Methodology**

The design process that captures the modeling methodology described in this book is shown in Figure 1-50.
the return current of signal lines. Moreover, since the planes are electrically large along the lateral dimensions, an efficient numerical method is required from the modeling standpoint that reduces complexity (number of unknowns). So, in the methodology shown in Figure 1-50, the PDN consisting of planes, vias, capacitors, and solder balls is first modeled in the frequency domain. The goal of this analysis is to ensure that the target impedance of the PDN is met at the frequencies of interest, as dictated by the system application. The modeling of planes with vias and decoupling capacitors in the frequency domain is the subject of Chapter 2.

Assuming the target impedance is met in the frequency domain, the next step is to incorporate the signal lines along with the PDN. Doing so allows a designer to evaluate the coupling between the signal lines through the PDN and between the signal lines and PDN due to the return currents. Details on the incorporation of the signal lines into the PDN and coupling analysis in the frequency domain are found in Chapter 3.

Assuming the coupling meets the specifications in the frequency domain, the next step is to convert the frequency information into a time-domain signal to
compute SSN. Assuming no nonlinear circuits are connected to the signal lines, this process is straightforward, since inverse fast Fourier transform (IFFT) methods can be used for this purpose. However, this is not a true representation of SSN, since nonlinear circuits can have memory and feedback effect whereby excessive SSN can slow down the drivers, reducing or saturating the power supply noise [33]. The interaction between the nonlinear drivers, signal lines, and PDN can only be captured through a time-domain simulation. Chapter 4 describes time-domain simulation methods that convert the frequency-domain response of the signal lines and PDN into a modified nodal analysis (MNA) formulation, which is used by most circuit simulators such as Spice. The frequency response can therefore be converted into a Spice subcircuit to which nonlinear transistor circuits can be connected. During time-domain simulation, various PRBS patterns can be simulated to ensure that SSN is within specifications. If the specifications are violated for any particular bit stream, the impedance of the PDN or the routing of the transmission lines can be re-optimized around the frequencies corresponding to the bit stream before final tape out.

In Chapter 5, various applications are discussed using methods described in chapters 2, 3, and 4 to analyze impedances and scattering parameters in the frequency domain and SSN in the time domain. Advanced technologies such as embedded decoupling capacitors and electromagnetic bandgap (EBG) structures for controlling SSN are also discussed in Chapter 5. In addition, more focused examples are provided in each chapter to capture specific effects that can be used to evaluate commercial tools.

1.10 Summary

This chapter described the basic concepts in PDN design. The microprocessor and other ICs are sensitive to power supply fluctuations. An increase in the supply voltage causes reliability problems, while a decrease reduces the maximum operating frequency. By relating the impedance of the PDN and power supply noise, the network can be suitably designed in the frequency domain. By using the target impedance as a parameter for measuring the goodness of a PDN, the various components such as VRM, decoupling capacitors, vias, planes, and solder balls can be analyzed and optimized until the target impedance is met. The effect of the chip can also be included in the design process by representing it as either a lossless capacitor or a lossy capacitor. The relationship between impedance in the frequency domain and SSN in the time domain was also quantified in this chapter using simple models.
References


32. A. Peterson, Georgia Institute of Technology, Private Communication.