Semiconductor Technology from A to Z

Reliability

ESD Protection Structures

ESD Models and Damage Mechanisms

Electrostatic discharge (ESD) describes the abrupt equalization of an electrostatic charge between two objects at different potentials. When a charged human body, a component, or a machine touches an integrated circuit, a current of several amperes flows through the connection pins into the device for a few nanoseconds to microseconds. This brief but high energy density can break down gate oxides, thermally destroy p-n junctions, or melt metallization traces.

Three models have become established for evaluating ESD robustness, each reproducing a different discharge source, and they differ markedly in pulse shape, rise time, and peak current:

Model	Equivalent circuit	Rise time	Pulse duration	Typical peak current
HBM (Human Body Model)	100 pF through 1.5 kΩ	2–10 ns	approx. 150 ns	approx. 1.3 A at 2 kV
MM (Machine Model)	200 pF, no series resistance	<1 ns (ringing)	approx. 150–300 ns	several A at 200 V
CDM (Charged Device Model)	device self-capacitance (a few pF)	<1 ns	<1 ns	up to 10–15 A at 500 V

Qualification of a device typically requires minimum robustness levels, such as 2 kV under the HBM model and 500 V to 1 kV under the CDM model for consumer-oriented applications; automotive devices often demand considerably higher levels. Because the entire charge stored in the component discharges within under a nanosecond under CDM, this model produces by far the highest current peaks despite the lower voltage, and the most demanding requirements on the response speed of the protection structure.

Protection Diodes and Snapback Behavior

The simplest ESD protection structure consists of two clamping diodes at every input or output pad: one blocks toward the supply voltage V_{DD}, the other toward ground V_{SS}. During a positive ESD pulse, the upper diode conducts and diverts the current into the V_{DD} rail; during a negative pulse, the lower diode takes over the discharge path to ground. The clamping voltage of the diodes limits the voltage that reaches the gate oxide of the downstream core circuit to a non-critical level – for a typical silicon diode the forward voltage is about 0.7 V, so the voltage at the protected node is clamped to roughly V_{DD} + 0.7 V or V_{SS} –

0.7 V.

For higher current robustness, snapback structures are frequently used instead of simple diodes. Their characteristic corresponds to a parasitic thyristor (Silicon Controlled Rectifier, SCR). Once the voltage across the device exceeds the breakdown voltage – typically in the range of 8 to 15 V in modern CMOS technologies –, a parasitic npn bipolar transistor turns on, whose base current in turn activates a parasitic pnp transistor – the structure latches into a low-impedance state with a significantly reduced holding voltage of often only 1 to 2 V. This snapback behavior allows large ESD currents to be discharged with low power dissipation, but requires careful tuning of the holding voltage above the supply voltage to avoid unintended latchup during normal operation.

Grounded-Gate NMOS and Power Clamps

A widely used snapback protection structure is the grounded-gate NMOS (GGNMOS), in which the gate and source of an NMOS transistor are permanently tied to ground while the drain is connected to the pad to be protected. An ESD pulse drives the drain-bulk junction into avalanche breakdown; the resulting holes flow through the bulk region and raise its potential until the parasitic npn bipolar transistor between source, bulk, and drain turns on and discharges the ESD current at low impedance. The maximum current-carrying capability of a GGNMOS structure before thermal failure (failure current I_{t2}) scales roughly with the finger width of the transistor and is typically on the order of a few milliamperes per micrometer of gate width.

To protect the internal supply rails, a power clamp is additionally placed between V_{DD} and V_{SS}. It remains high-impedance during normal operation and switches to a low-impedance state only for the duration of a fast ESD voltage rise, detected through an RC network with a typical time constant of 0.1 to 1 µs, tuned to be much shorter than the millisecond-scale rise time of a regular power-up event. The design of every ESD structure involves a fundamental trade-off: larger protection structures increase current robustness but also increase the parasitic capacitance at the pad and therefore the signal delay, which must be weighed carefully for high-frequency and high-speed interfaces.