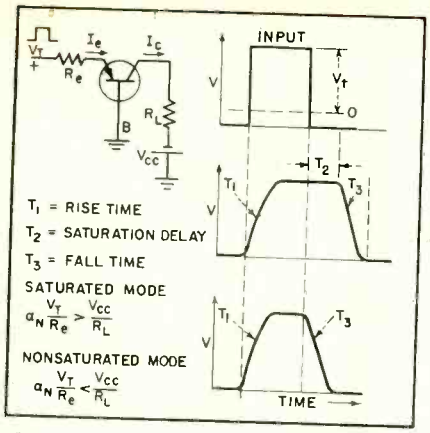

FIG. 1—Response time with storage delay (middle curve) and no delay (lower curve)

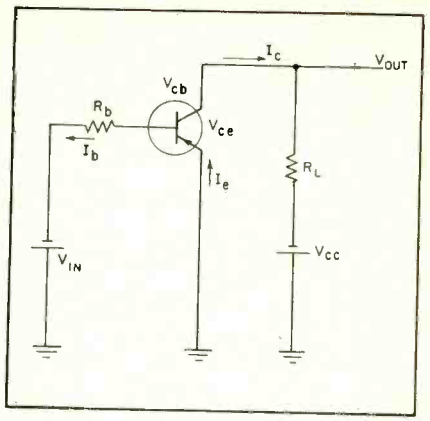

FIG. 2—Back-clamping technique giving voltage gain along with good efficiency

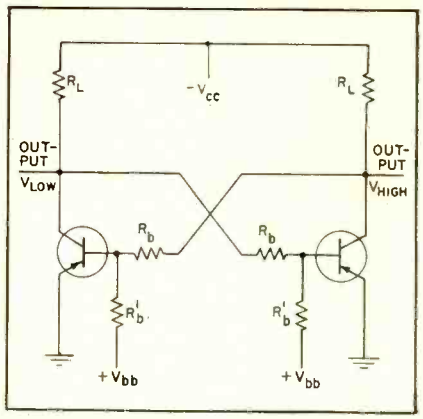

FIG. 3—Bistable saturated circuit is less efficient at low output power

BOOSTING TRANSISTOR

By RICHARD H. BAKER

Staff Member
Massachusetts Institute of Technology
Lincoln Laboratory, Cambridge, Mass.

SUMMARY — Transistor properties affecting response time in switching circuits are summarized and basic circuits given for obtaining maximum energy conversion efficiency. Combined use of pnp and npn transistors gives circuit symmetry that utilizes inherent advantages of transistors. Other circuits include saturated and nonsaturated current-demand flip-flops with single or double triggering, designed for maximum reliability despite normal variations in circuit constants and input pulses

THE NORMAL three-region junction transistor (excluding graded-base or drift types) is a slow device when compared to a vacuum-tube triode. In a tube, the movement of electrons from cathode to plate is aided by strong electric fields, whereas in a transistor the transport of carriers (electrons or holes) is only by diffusion.

In designing transistor circuits for high-speed switching, the designer must consider normal integrative effects due to shunt capacitances as well as the delay or carrier transit time between emitter and collector. When the transistor is operated in the saturated mode, there exists an additional effect, that of hole storage or saturation delay.

Response Times

There may be as many as three separate response times, depending upon the mode of operation, associated with a single-stage transistor network. These are rise time, storage or saturation delay and fall time, all shown in Fig. 1. If R_L, R_e, V_{cc} and V_T are chosen so that the voltage polarity across the collector junction maintains the collector junction under reverse bias at the peak of the output pulse, the saturation delay vanishes.

The magnitudes of response times T_1, T_2 and T_3 are different for each of the three basic connections. In all modes of operation, however, the transistor switching time is dependent on the constants of the device and the amount of overdrive supplied at the input.

The single most important factor affecting the switching time is the frequency response of the device itself. Also, minimum response time occurs when current gain α_N is 1.

There is promise of obtaining high-frequency transistors by using graded-base structures and other configurations. However, the interim solution of transistor manufacturers has been to build transistors with very narrow base widths to increase the frequency response. This approach is fruitful to a degree, but there is an op-

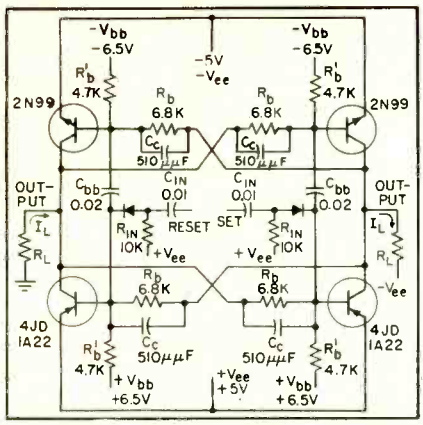

FIG. 4—Saturated current-demand single-triggering circuit with 1-mc transistors

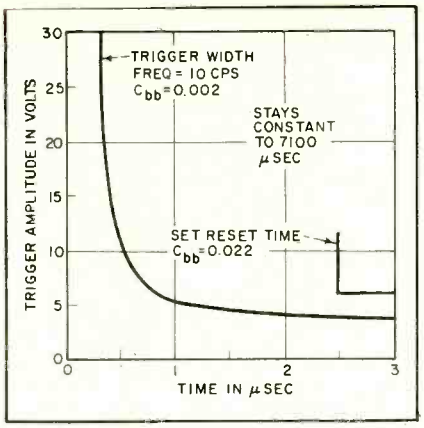

FIG. 5—Characteristics of single-triggering circuit of Fig. 4

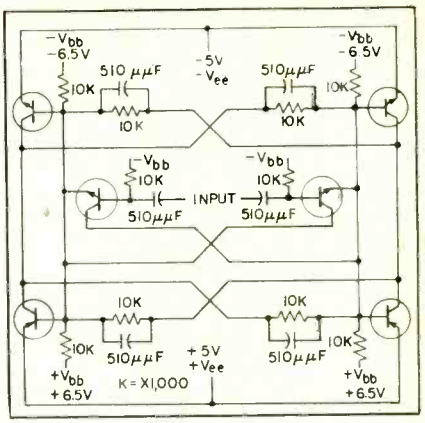

FIG. 6—Saturated current-demand flip-flop double-triggering with 1-mc transistors

SWITCHING SPEED

timum base width that yields minimum switching time for practical switching circuits. This optimum base width is generally different for each of the three basic connections.

In transistor circuits the transit time of the carrier across the base region imposes an absolute minimum input pulse width. This in turn sets rather large minimum capacitance values in a given circuit, creating recovery time problems that may be more serious than actual rise time considerations.

Signal Levels

Because transistors are extremely efficient voltagewise, the system levels are usually set by a combination of system and transistor considerations.

The low voltage limit is automatically set if the transistors are allowed to saturate, this being primarily determined by speed considerations.

The upper voltage limit is set by the total power consumption of the system and by the punch-through and avalanche phenomena in the transistor.

The signal voltage swing in an all-transistor system is usually chosen as a compromise between two inherent opposing effects. As the signal level increases (total swing), the amount of energy dissipated in charging and discharging capacitance increases. This effect indicates that the signal level should be low. On the other hand, for convenience of circuit design the signal level should be large compared with the transistor off-on uncertainty region, which is about 0.2 volt for germanium transistors and about 1 volt for silicon transistors.

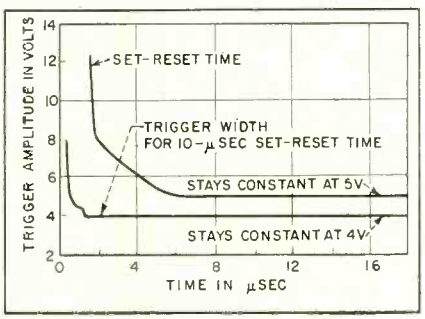

FIG. 7—Characteristics of double-triggering circuit of Fig. 6

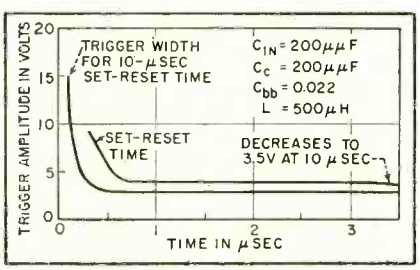

FIG. 8—Characteristics of double-triggering circuit using 5-mc transistors

Energy Conversion Efficiency

Fundamentally the transistor, like the vacuum tube, has gain by virtue of dissipation changes. Unlike the vacuum tube, the input impedance is much lower than the output impedance. In the design of realistic transistor systems, then, a serious problem arises in the available power to drive succeeding stages. This situation is aggravated still further in the design of high-speed systems, since it is necessary in the transient state to overdrive the stages to obtain fast switching. This fact, more than any other, accounts for the large number of transistors required to build transistor systems compared with equivalent vacuum-tube systems.

These considerations indicate that circuitry should be designed to deliver maximum output power and that a high percentage of the available output power should be available to drive other transistors. Further, since currently available high-frequency transistors are ex-

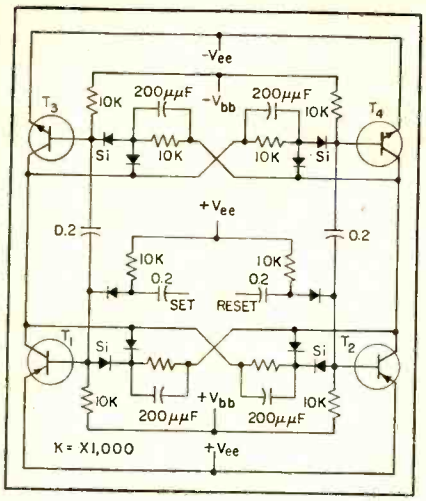

FIG. 9—Nonsaturated current-demand single-triggering flip-flop

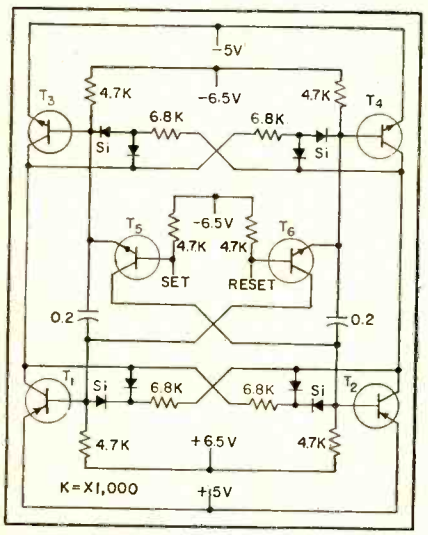

FIG. 10—High-speed nonsaturated current-demand double-triggering flip-flop

tremely low-power devices (on the order of 0.5 mc-watt as a figure of merit for SBT-100 at 50 mc and 10 mw), circuitry should be designed to give highest possible energy conversion efficiency. The ratio n_1 of useful signal output power to transistor dissipation should therefore approach infinity.

For optimum circuit design of a minimum-power-drain system, the ratio n_2 of useful signal output power to power supply drain should approach 1.

The product of n_1 and n_2 should be made as large as possible. The ratio represented by n_1 can be made large by allowing the transistor to saturate or by controlling the voltage from collector to base through the use of clamping diodes. However, minimum-power systems can be built only by making n_2 close to unity (this must be true if there exists a minimum power level to process intelligence).

In most present transistor circuit designs, a high percentage of useful output power from the transistor is dissipated in the load resistors. This is especially true for direct-coupled logic. Therefore, the value of n_2 may be increased significantly by removing the standby power dissipated in this area.

The circuit design techniques in the following sections show how the values of n_1 and n_2 may be increased to give minimum power dissipation, maximum speed and minimum sensitivity to component and transistor drift circuits.

Maximum-Efficiency Circuits

Aside from eliminating the power dissipated in load resistors, an additional gain in system power efficiency may be obtained by using circuits that draw power from the supplies according to the power and demand at the output. This process always involves feedback. Cathode-follower and emitter-follower (grounded-collector) circuits do this, but unfortunately have no voltage gain.

A transistor circuit involving voltage gain, along with an ability to convert d-c power into signal power as required by the load, is shown in Fig. 2. The transistor dissipation is low and the output power is high for collector currents less than the maximum output current. The major portion of the power drawn from the supplies is available at the output for dissipation in the load resistor, so that n_2 approaches 1 and the circuit draws from the supplies only the power dissipated in R_L and R_b (neglecting transistor dissipation). The only transistor parameter of importance in the conducting state is the minimum base-to-collector current gain β_N.

To illustrate the design of this circuit, assume that β_N is 20, input d-c voltage V_{in} is 5 volts, transistor saturated base resistance r_b is 50 ohms, V_{cc} is 10 volts, R_b is 10,000 ohms and R_L is 1,000 ohms. Then I_b is about 0.5 ma, $I_{c\,max}$ is 10 ma and I_e is the sum of these or 10.5 ma. Useful signal output power is $\frac{1}{2}\ V_{cc}\ I_e$ or 50 mw and transistor dissipation is 10.5×0.2 mw, so that n_1 is about 24. Power supply drain is 50 mw + 2.5 mw, so that n_2 is about 0.95.

Two of the circuits of Fig. 2 may be coupled together, with only slight modification, to form the bistable circuit of Fig. 3. This has several drawbacks, however. The low-voltage level is not fixed, being dependent on I_{co} and other factors. The power dissipated in internal load resistor R_L (in shunt with the actual load) may be an appreciable percentage, particularly at low output power levels. For fast fall time (when the transistor is turned off), R_L must be made small.

Current-Demand Circuit

A circuit that circumvents these disadvantages is shown in Fig. 4. Here essentially all of the output current (collector current) is available to drive load R_L.

Standby power is low; when there is no load, the power taken from the supplies is approximately equal to the dissipation $2\ I_b^2 R_b$ in the base resistors. Both the high and low voltages are clamped (the transistors saturate). The

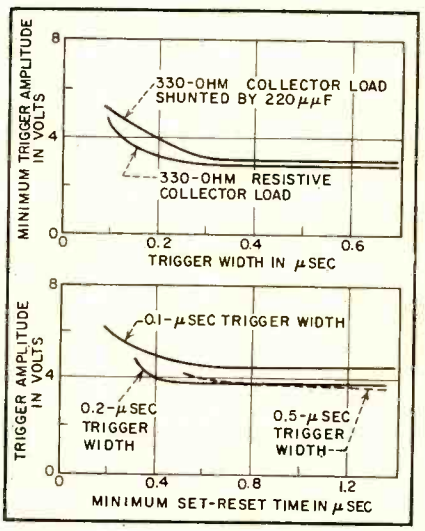

FIG. 11—Triggering characteristics of flip-flop of Fig. 10

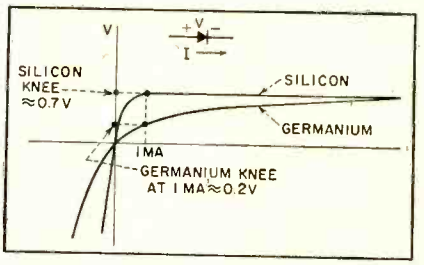

FIG. 12—Diode characteristic curves

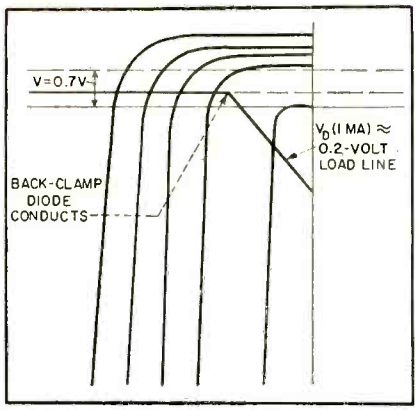

FIG. 13—Family of collector curves for silicon transistor

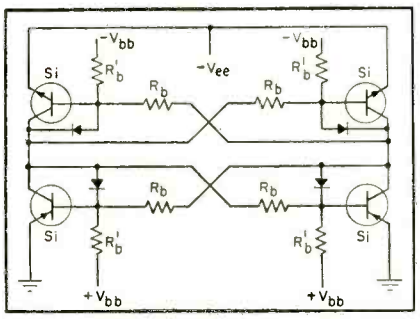

FIG. 14—Basic silicon transistor circuit for nonsaturated current-demand flip-flop

tolerance on all resistors may be large (on the order of 50 percent).

Circuit operation is substantially independent of transistor parameters. The stability of the configuration is insensitive to supply voltages. The configuration leads to fast rise and fall time since large transistor overdrive is inherent. The last three advantages accrue because the configuration allows the transistors to set their own levels. Some of the operating characteristics of the circuit are shown in Fig. 5.

One difficulty with the circuit of Fig. 4 is that there is an appreciable delay around the loop because the conducting transistors are saturated. This difficulty may be minimized by double triggering (triggering all four transistors simultaneously). The current-demand flip-flop circuit of Fig. 6, employing this feature, gives the characteristics shown in Fig. 7. By using 5-mc transistors in this circuit and changing all 510-$\mu\mu$f capacitors to 200 $\mu\mu$f, the characteristics of Fig. 8 can be obtained.

The circuit techniques described may be extended to nonsaturating circuits. The primary gain in designing the circuits to operate in the nonsaturating mode is decreased switching time. Figure 9 shows a typical design using the nonsaturated configuration with single triggering. Figure 10 shows a higher-speed version using double triggering, and Fig. 11 gives triggering conditions for the circuit. The diode characteristics in Fig. 12 show why these back-clamped circuits do not allow the transistors to be saturated.

If silicon transistors are used, the nonsaturated circuits do not require the four silicon diodes. This may be seen from the silicon collector curves in Fig. 13. The basic circuit using silicon transistors is shown in Fig. 14.

The salient features of the saturated back-clamping current-demand technique are low transistor dissipation, high conversion efficiency, insensitivity to component and transistor parameters (standby load resistors not needed), insensitivity to voltage supply drift, maximum system efficiency (power drawn from supplies according to needs of load), fast rise and fall time (inherent overdrive) and loop delay (caused by saturation time). Nonsaturated circuits give increased operating speed because they have no saturation delay, but are otherwise identical.

Gating Circuits

The design of maximum-reliability switching systems depends heavily upon the reliability of the voltage-pulse voltage-level gate. To assure maximum system reliability (assure positive action and suppress superfluous triggering), the gate circuits should be independent of pulse width, pulse amplitude, pulse repetition frequency and pulse level (within given limits), and should have fast response to pulse and level changes. The circuit design should also be insensitive to component values and transistor parameters, require minimum standby power, have high output power, present a constant load to the pulse source (driver) and deliver standardized output pulse and level amplitudes.

A circuit configuration that fulfills to a high degree the above reliability characteristics is shown in Fig. 15 along with its gating waveforms.

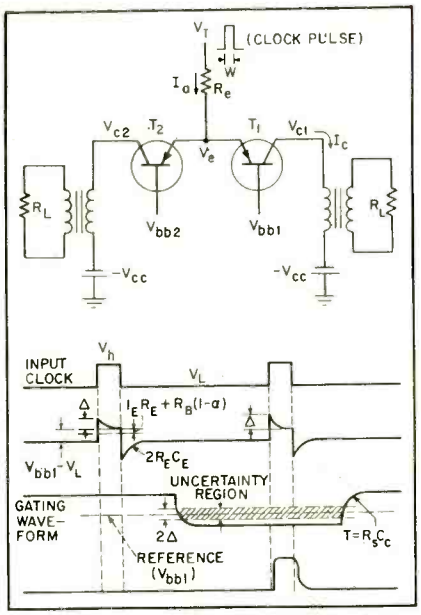

FIG. 15—Circuit and waveforms of pulse-level gating circuit

Conclusions

The reliability of transistor switching systems is closely related to the design of circuits. The circuit designer must consider the drift of operating points caused by aging and ambient self-generated temperature changes. For high-speed networks, due to the lack of high-speed transistors, overdrive must be used to speed up the circuit response.

Transistors are inherently efficient devices (both voltagewise and powerwise). This, along with the fact that two types of transistors are available (*npn* and *pnp*), allows circuit design that is extremely efficient in terms of power supply drain for a given signal power output.

The transistor, being an efficient, reliable and small device, may be soldered into systems much as are ordinary resistors and capacitors. This, plus the fact that it is basically a three-terminal passive device which can produce power gain, makes its use attractive in networks where feedback techniques are widely employed.

The research work herein described was supported jointly by the Army, Navy and Air Force under contract with Massachusetts Institute of Technology.