The new centre for submicron IC technology

W. G. Gelling and F. Valster

Taking part in the race to produce integrated circuits with diminishing details and expanding areas calls for enormous investment in people and buildings, not to mention computers and production equipment. People are necessary for their knowledge of processes such as plasma etching, chemical vapour deposition, implantation, oxidation and diffusion, and for combining these processes to produce patterns that form useful electronic circuits on a silicon slice. The equipment is necessary for projecting these patterns in minute detail on to the slice and creating the right conditions for the processes. Computers are needed for calculating and simulating the integrated circuits and for evaluating test data. Buildings are necessary to accommodate the people and hardware. In the new centre for submicron IC technology, now virtually completed at the Philips Research Laboratories site in Eindhoven, these buildings include areas where vibration and the dust content of the air are held to extremely low levels.

Introduction

In the sixties Philips brought out a hearing aid, for behind-the-ear wear, which contained an integrated circuit [1]. This IC consisted of an amplifier with three transistors. The 1-Mbit static RAM (RAM = random-access memory), which will go into pilot production in a few years in the new centre for IC submicron technology, will contain nearly ten million transistors. The hearing aid circuit was produced on a silicon chip with an area of about 0.6 mm², the static RAM will be produced on an area of about 90 mm². The old circuit was used for processing analog information, the new one will be used for storing immense amounts of digital information. These two examples typify the developments in electronics: integrated circuits are growing and contain more and more functions, the area required for each function continues to decrease, and digital technology is becoming firmly established.

If we look at the developments in the IC industry over a number of years, we see that the manufacturing cost per electronic function has decreased by a factor of ten in every five or six years. And this trend appears to be continuing. So it is more than likely that the number of applications of electronics in our daily life will continue to increase. We have digital electronics very much in mind here. The digitization of audio and video equipment, for example, is already under way, and the computer in all its manifestations will occupy an increasingly important place.

Keeping up with all of these developments requires much effort from a company such as ours. To meet this challenge, Philips decided to build a new centre for IC technology at the Research Laboratories site in Eindhoven. In this centre staff from the research laboratories will cooperate closely with development engineers from the Elcoma division, who are the producers of integrated circuits in the Philips group.

The new IC centre will mainly be housed in two buildings; see fig. 1 and 2. In one building there will be research and development on submicron technology — the technology that will give us integrated circuits with details smaller than 1 μm. This building will also contain the pilot production of advanced ICs before the Elcoma division take over the technology and start to manufacture the circuits. The design of this building is very largely determined by the requirements for dust-free production and insensitivity to vibration — vital conditions for the manufacture of high-technology integrated circuits. In the other building, now completed, the electronic circuits are designed. The layout of this building is mainly determined by the

Drs W. G. Gelling is a Deputy Director and Ir F. Valster a Director of Philips Research Laboratories, Eindhoven.
large computers and the test equipment it will have to accommodate: advanced ICs can only be designed and simulated with the help of computers (CAD, or Computer-Aided Design). For example, the computers in use at the moment require about ten hours to calculate how a large IC will respond to a signal at the input; the computers to be installed in the new building will perform calculations of this type in less than an hour.

The mastery of modern IC technology can be assessed by the ability to produce large memory circuits with a high yield. The largest integrated circuits now being made are RAMs. Since 1971, when the first IC memory with a capacity of 1 kbit ($2^{10} = 1024 \approx 10^3$ bit) was introduced, the capacity per IC has increased roughly sixfold every five years. Philips, for example, are well on the way with the design of a static RAM with a capacity of 256 kbit; see fig. 3. Philips plan to start pilot production of this memory, on 6-inch silicon slices, at the new centre in about a year. Samples of a 1-Mbit static RAM will be available in 1988. (1 Mbit = $2^{20} = 1048576 \approx 10^6$ bits.)

The 1-Mbit memory will be made in CMOS technology, and the individual transistors will have a channel length of 0.7 μm. The extremely detailed patterns required for these circuits will be projected on the silicon slice by photolithography. For this work a new high-resolution wafer stepper (or repeater projector) will be developed as a successor to the well-known Silicon Repeater [2]. The submicron technology required for manufacturing the static 1-Mbit memory will be used later for a wide variety of other types of integrated circuit. The advanced memory will thus act as the 'locomotive' that will provide the motive power for other Philips IC technology in the coming years.

In the following pages we shall explain why CMOS technology, photolithography and a static RAM have been chosen for the new centre. We shall also look at some of the technical aspects of the new buildings.

### IC technology

#### CMOS

Large digital ICs are now more and more likely to be made in MOS technology (MOS: metal-oxide semiconductor). Bipolar technology is mainly used in analog circuits and in high-speed digital circuits. Bipolar ICs are built up from npn transistors and in many cases pnp transistors. Bipolar transistors derive their name from the fact that both electrons and holes contribute to the conduction. In MOS transistors, on the other hand, the conduction is due to the transport

---

of holes or electrons: there is only one type of charge carrier.

The operation of a MOS transistor is based on the capacitive coupling of a conductor (the 'gate') to a p-doped or n-doped semiconducting material via a thin insulating oxide layer. A silicon NMOS transistor (on the left in fig. 4) contains two islands of heavily doped n-type silicon in a p-type silicon substrate. The two islands are called the source and drain respectively. The transistor operation is based on the conduction of electrons in the p-type silicon. These electrons are drawn from the source by a positive voltage on the gate to a thin layer of the p-type silicon (the 'channel') immediately below the insulating oxide. The gate may consist of metal or of heavily doped polycrystalline silicon. In a PMOS transistor (on the right in fig. 4) the source and drain take the form of islands of heavily doped p-type silicon in a larger island of n-type silicon. In this type of transistor a negative voltage on the gate produces hole conduction at the surface of the n-type silicon. The operation briefly described above applies to PMOS and NMOS transistors of the enhancement type. Transistors of the depletion type, on the other hand, conduct when there is no voltage applied to the gate.

In CMOS technology (CMOS: complementary metal-oxide semiconductor) PMOS and NMOS transistors are combined. As an example a diagram of a logic inverter made in CMOS technology is shown in fig. 4. With a positive supply voltage the output voltage $V_O$ is 0 when there is a positive input voltage $V_I$. If on the other hand $V_I$ is 0, then $V_O$ is equal to the supply voltage. As we shall presently see, similar transistor combinations are also found in static RAMs.

Compared with bipolar transistors, MOS transistors have a simple structure and make ideal switches. Another advantage is that simpler techniques are adequate for insulating MOS transistors from each other. The 'packing density' of MOS ICs is therefore about four times that of bipolar ICs. This high packing density and their operation as switches make MOS transistors particularly suitable for large digital integrated circuits. The first MOS circuits were produced in PMOS technology. It was the introduction of ion implantation in IC manufacture that made circuits in NMOS technology possible. ICS in NMOS technology are faster than those in PMOS technology, because electrons have a higher mobility than holes.

Raising the packing density in integrated circuits has increased the problem of heat removal. It is therefore a considerable advantage that the heat dissipation of ICs in CMOS technology is low. This is because the elements of logic circuits and memories in CMOS technology only conduct when there is a change in the information content of a switching element. The only currents flowing in the circuit in fig. 4, for example, in the steady state are the leakage current of the non-conducting transistor and the leakage currents in the reverse-biased p-n junctions. A disadvantage of CMOS technology is that the packing density is lower than in PMOS or NMOS technology, since n-type silicon islands are required. The problem of the lower packing density can be compensated to a considerable extent by ingenious design.

The level of 'sophistication' of the technology used to produce an integrated circuit can be characterized by the channel length ($l$ in fig. 4). In recent years this parameter has been dramatically reduced, and in the static 256-kbit memory we mentioned it is 1.2 μm. The transistors to be made in the new centre for IC submicron technology will have a channel length of 0.7 μm. This sharp reduction in scale will create new problems, of course, and we shall look at some of them here.

To produce a smaller MOS transistor with a useful characteristic it is necessary to satisfy certain rules for scaling. A smaller channel area, the product of the length and width of the channel, implies that the thickness of the channel oxide must also be proportionately smaller. At the same time, the implantation dose of boron or arsenic must be larger to achieve the required threshold voltage. The diffusion depth must also be proportionately smaller if the lateral source and drain diffusion is to be limited during the diffusion of boron or arsenic at high temperature. Shallow diffusion, however, increases the resistances in the transistor, and this can reduce its speed of response.

The dimensions can also be scaled down by means of 'three-dimensional' cell structures. Research in the new centre will therefore include work on techniques such as the 'stacking' of elements and the fabrication process.
of 'vertical' structures in a groove in the slice. This introduces the additional difficulty of restoring the surface flatness of the slices during the production.

With smaller details it is obvious that the alignment should be more accurate and the resolution better when the mask patterns are projected on the slice. The accuracy of the application and partial removal of the various layers must also be improved. These are problems of lithography, and we shall now consider them more closely.

**Lithography**

It might seem obvious that X-ray lithography or electron lithography would be used in the new IC centre. It has been decided, however, to use photolithography, although this does not exclude the possibility of other techniques in the future, and there is research on new lithographic techniques in Philips laboratories [3][4].

X-ray and electron lithography have the advantage of a higher resolution because the radiation has a shorter wavelength (0.5 to 3 nm for X-rays, 0.05 nm for electron lithography and about 400 nm for photolithography). The disadvantage of X-ray lithography is that the masks are as yet very difficult to produce and it is necessary to use synchrotron radiation [4]. At present the large and expensive storage ring that is required as an X-ray source is not suitable for use in an IC factory. The disadvantage of electron lithography is that it takes a long time to write a pattern — about an hour for a complete 6-inch silicon slice. Although a more ingenious method of scanning the pattern saves time, the method would still not be economic even if the scanning duration were reduced by a factor of ten (if such a reduction were possible). Electron-optical pattern generators are however used for producing accurate masks for photolithography.

The wafer stepper that was developed at Philips Research Laboratories (the Silicon Repeater [2]) has a highly refined alignment system, with a sensitivity of 0.02 μm, which makes it suitable in principle for sub-micron technology. It is necessary, however, to improve the resolution of the projection optics, and a new optical system is therefore now being developed in cooperation with a specialist manufacturer.

The etching operations following exposure in the repeater projector in the new IC centre will in general be 'dry'. With conventional wet-etching techniques the highly detailed patterns cannot be transferred with the required accuracy and not all materials can be etched. The dry-etching techniques of plasma etching and (reactive) ion etching are less subject to these difficulties [6].

**Memory circuits**

As noted earlier, semiconductor memories may be regarded as the 'locomotive' of submicron IC technology, since the most advanced techniques are required in order to accommodate as many memory cells as possible in unit area. With good test procedures faults can be traced and localized very accurately in many cases. Minute examination of the faults then provides the vital clues that enable the technology to be improved. Logic circuits are not very suitable for our 'locomotive' function here, since it is much more difficult to localize the faults in a logic circuit. Memories are thus an important link in the learning process for an optimum mastery of IC technology, whether in research and development or in actual production. In turn, mastering the technology of standardized memories generates the skills required for the economic manufacture of other types of custom-made integrated circuit.

Semiconductor memories fall into various categories. Research on the applications of analog and digital CCD memories (CCD: charge-coupled device) has long been in progress at Philips [6]. Memories of this type work serially, on the shift-register principle, and the individual memory cells do not have addresses. The cells in ROM and RAM memories (ROM: readonly memory, RAM: random-access memory) do have addresses. A ROM can only be read. The contents are fixed and have already been entered into it during manufacture. A RAM can be 'read from' and 'written to', so that its contents can be changed.

RAMs are 'volatile' memories: the information is lost if the supply voltage fails. RAMs can be subdivided into dynamic and static types. A cell of a dynamic RAM consists in principle of a capacitor and a transistor; see fig. 5a. The gate of the transistor is connected to the 'word line' (for addressing the rows of the memory matrix) and the source is connected to the 'bit line' (for addressing the columns). The capacitor of a memory cell is charged by applying a relatively high voltage to the word line and bit line correspond-
ming to the address of the cell. The charge of the capacitor slowly leaks away to earth. The contents of the cell must therefore be ‘refreshed’ periodically, (e.g. every 2 ms), by supplying sufficient charge to restore the voltage to its original value. A disadvantage of a dynamic RAM is therefore that it requires extra electronics for the refresh operation and that the contents of the memory are not available during this operation. An advantage is that a dynamic-RAM cell takes up extremely little space.

A memory cell of a static RAM needs much more space, because each cell consists of a flip-flop with four transistors. Furthermore, two additional transistors are required to connect the cell to the two bit lines in this case; see fig. 5b. The word line is connected to the gates of these coupling transistors, and each bit line is connected to a source. The contents of a cell are changed by applying the higher voltage to a bit line and the word line, and the lower voltage to the other.

Fig. 5. a) Diagram of a cell in a dynamic RAM. T MOS transistor. W word line. B bit line. C capacitor. b) Diagram of a cell in a static RAM. TA-4 MOS transistors that form a flip-flop. (The upper two transistors are PMOS transistors, the lower two are NMOS transistors.) TS coupling transistors. S+ inverse bit line. Vdd voltage on drain electrodes. Vss voltage on source electrodes. c) Photograph of some memory cells on the graphic display during CAD of the 256-kbit SRAM; see also fig. 3. The various coloured areas represent masks used in the fabrication process. For clarity not all of the masks have been shown. The vertical blue stripes are bit lines (B and B in b); the horizontal red stripe is a word line (W in b). d) Scanning electron micrograph of a few cells from a very similar memory at an intermediate stage of the fabrication process: the first metal conductor lines have been applied, but not the bit lines. One ‘dash’ on the scale on the left corresponds to 10 μm on the slice. The minimum pitch of the lines is 2.6 μm. The linewidth is 1.2 μm. The actual lines are obscured by layers above them, so that the lines seem to be wider. The circular patterns at the bottom of the picture correspond to the squares that can be seen at the bottom of the screen in (c). These patterns are the bit-line contacts on the extreme left and right in (b).
faults are found in one or more cells, they can be disconnected. This is done by interrupting conductor lines with a focused laser beam. Some of the spare cells then take over the function of the failed cells. This is done by interrupting lines to the gates of depletion-type transistors, so that these transistors then conduct.

A memory cell of a dynamic RAM takes up about a quarter of the space occupied by a cell of a static RAM. However, the control electronics of a static RAM occupies far less space and the stored information is continuously available at all times. Because of their specific advantages and disadvantages, both types of memory have their preferred areas of application. As we mentioned earlier, Philips will devote a substantial part of the effort in the new IC centre to the development of a 1-Mbit SRAM. In terms of packing density a 1-Mbit SRAM is comparable with a 4-Mbit DRAM.

Some technical aspects of the buildings

As more and more transistors of ever smaller dimensions are accommodated on a chip, the chance of an error occurring in any transistor will have to diminish correspondingly. Otherwise the probability of a fault-free circuit will become so small that a satisfactory production yield will no longer be possible. It is found that dust is one of the most common causes of faults. One of the principal requirements to be met by the processing rooms in the submicron-technology building (fig. 1) is therefore an extremely low dust content in the air.

Fig. 5 is a diagram showing the connection between the number of bits of a dynamic RAM, the width of the lines in it and the required minimum dust concentration in the air as specified by extrapolation of the American Federal Standard 209B [7]. In this standard the highest degree of cleanliness is referred to as class 100. In a class-100 clean room the air must contain no more than 3500 particles of 0.5 μm diameter or larger per m³ (100 such particles per cubic foot). The corresponding line in the figure relates to cumulative values for other particle sizes. It is now normal practice to extrapolate standards for classes 10 and 1 from this standard; lines for these classes are also shown in the figure. These degrees of cleanliness will probably be standardized as well. With the initial assumption that particles whose diameter is 10% of the linewidth will not cause chip rejects, it is further assumed that for manufacture of 64-kbit DRAMs the air must be cleaner than air of class 10. For 256-kbit DRAMs the air cleanliness should be class 1; see fig. 6.

As mentioned earlier, the new 1-Mbit SRAM has a packing density comparable to that of a 4-Mbit DRAM. Further extrapolation of Federal Standard 209B shows that the air in the clean room for fabricating memories of this type must contain no more than 350 particles of 0.1 μm diameter or larger per m³ (10 per cubic foot). This is a much more difficult require-

actual processes. The less-clean tunnels contain equipment requiring regular maintenance. For example, the ion-implantation equipment will be accommodated in a less-clean tunnel that opens into a superclean tunnel.

To avoid manipulations with bottles or containers in the actual processes, the less-clean tunnels contain equipment. All liquids and gases are transported along pipelines. Containers for liquids and gases are kept in separate compartments on the outside of the building. Alongside the actual technology building, adding 10% of outside air. The Sankey flow diagram in fig. 7 shows the extent to which particles are trapped by the mechanical filters in the superclean tunnels [8]. The air entering from outside passes through a filter that traps 99.8% of the particles of diameter 0.1 μm or larger. The air then passes through a 73% prefilter and then, before entering the tunnel, it goes through a 99.99995% filter — clearly a filter of very special design.

![Sankey flow diagram](https://example.com/sankey-diagram.png)

**Fig. 7.** Sankey flow diagram [8] for the dust content of the air flowing through the technology building. **T** (superclean) tunnel. The numbers in the ‘flow paths’ indicate the number of particles larger than 0.1 μm per m³ volume. The width of the flow paths is approximately proportional to this number. The filters are indicated by a zig-zag line; the percentages give the degree of filtering for particles of 0.1 μm and larger. **F₁** outside air filter. **F₂** prefilter for the superclean process tunnels. **F₃** final filter for these tunnels. **A** air drawn in from outside. The outside air contains 10¹¹ particles per m³. The fresh air amounts to 10% of the circulating air. **B** 1% of leakage air in the outside-air suction line; this leakage air contains 10⁶ particles per m³. **C** 1% of leakage air in the channels of the circulation system, with the same particle content. **D** 2% of leakage air in the fans of the circulation system, also with the same particle content. **E** exit air from the process rooms; a content of 2 × 10⁶ particles per m³ is assumed for **E**. This diagram shows that the air in the cleanest process tunnels contains 7 particles of 0.1 μm diameter or more per m³. In calculating the filters a safety factor of 50 was applied to the specification of a maximum of 350 of these particles per m³.

![Horizontal cross-section](https://example.com/horizontal-cross-section.png)

**Fig. 8.** Horizontal cross-section through the columns and walls that connect the foundation floor and the floor for the wafer steppers. The two floors and the walls form a stiff concrete box construction.

Other important aspects of the design of the technology building are the precautions for suppressing vibration at certain locations. The wafer steppers are the units most sensitive to vibration, so that these precautions relate mainly to the floor on which they stand. These instruments always have their own vibration isolation in the form of weak undamped springs. It can be shown theoretically [9] that the spring-mounted instrument should have a low natural frequency, since floor vibrations with a frequency of at least ½ times the natural frequency are attenuated in amplitude. The natural frequency of our wafer steppers on their weak springs is low: 2 to 4 Hz.

The floor and foundations for the wafer steppers are designed to give the whole structure the highest possible natural frequency, in any case higher than 4/2 Hz. In addition the combined mass of floor and foundations must be high, so that the impacts of collisions with moving objects will only produce vibrations of low amplitude. A large mass in combination with a high natural frequency results in a high stiffness. In the design of low-vibration buildings it is generally only the stiffness in the vertical direction that is kept

---

[8] This diagram is based on calculations made by Meissner & Wurst GmbH, Stuttgart, the firm that supplied the air-conditioning equipment.

high. Since impact can also operate in the horizontal direction, the floors for the repeater projectors also have a high stiffness in the horizontal direction.

Calculations have been made for two models of the combined construction of concrete piles, foundation floor and the actual floor for the wafer steppers. In one model both floors are connected by columns, in the other by walls and columns, giving a sort of box construction. The calculations on the box-construction model showed that this design had the highest natural frequency in the horizontal direction. Fig. 8 shows a horizontal cross-section through the concrete construction that has been used. The upper floor is 45 cm thick and supports the wafer steppers. The measured natural frequency in the horizontal direction is 55 Hz — much higher than the natural frequency of the repeater projectors.

Summary. The pilot production of 256-kbit static RAMs will start in about a year at the new centre for submicron technology. Use will be made of CMOS technology and photolithography with wafer steppers. In certain parts of the technology building the dust content of the air will be extremely low and the floors will be extremely insensitive to vibration. The low dust content will be achieved by using separate air-circulation systems for each "process tunnel" with filters that trap 99.99995% of particles down to a diameter of 0.1 μm. The vibration insensitivity of the floors that support the wafer steppers will be obtained by using a concrete box construction with high vertical and horizontal stiffness.