PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification [3]: H01L 29/06, 29/04, 23/48, 29/44, 29/46, 23/32	A1	(11) International Publication Number: WO 80/01220
		(43) International Publication Date: 12 June 1980 (12.06.80)

(21) International Application Number: PCT/US79/01012

(22) International Filing Date: 13 November 1979 (13.11.79)

(31) Priority Application Number: 964,547

(32) Priority Date: 29 November 1978 (29.11.78)

(33) Priority Country: US

(71) Applicant: HUGHES AIRCRAFT COMPANY [US/US]; Centinela Avenue and Teale Street, Culver City, CA 90230 (US).

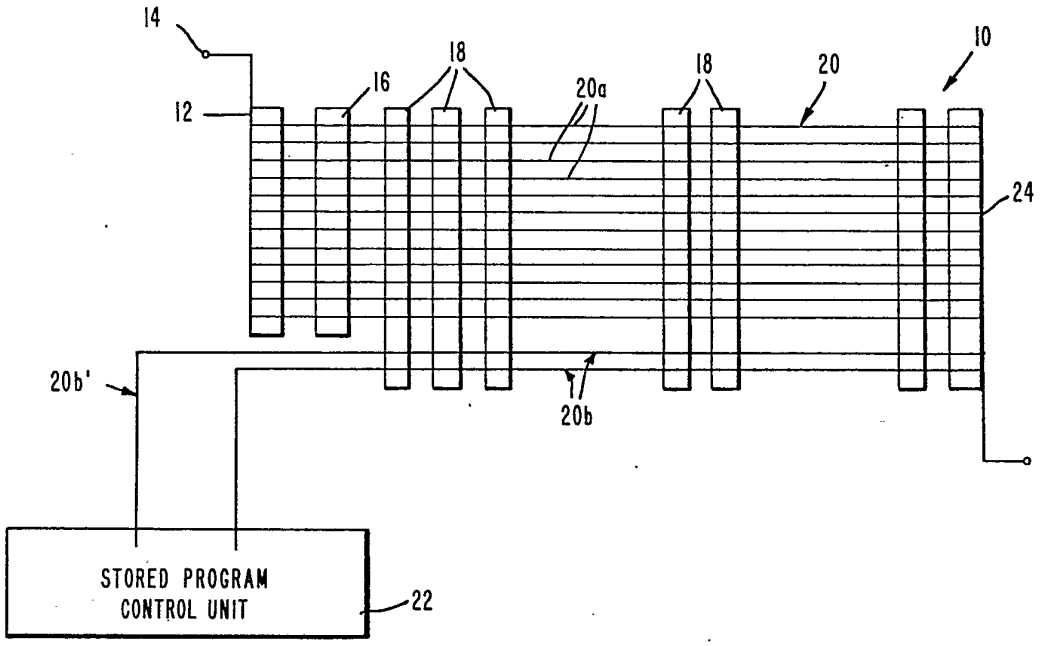
(72) Inventors: GRINBERG, Jan; 1141 Carmelina Avenue, Los Angeles, CA 90049 (US). JACOBSON, Alexander, D.; 12256 Canna Road, Los Angeles, CA 90049 (US). CHOW, Kuen; 1684 Calle Yucca, Thousand Oaks, CA 91360 (US).

(74) Agents: MacALLISTER, W.H. et al.; P.O. Box 90515, Los Angeles, CA 90009 (US).

(81) Designated States: DE (European patent), FR (European patent), GB, GB (European patent), JP.

Published
With international search report

(54) Title: THREE-DIMENSIONALLY STRUCTURED MICROELECTRONIC DEVICE



(57) Abstract

A large scale parallel architecture in which many parallel channels operate simultaneously to create a natural and efficient organization for processing two-dimensional arrays of data. The architecture comprises a plurality of stack integrated circuit wafers (16, 18) having top and bottom surfaces, electric signal paths (20) extending through each of the wafers between the surfaces, and micro-interconnects (50) on the surfaces of adjacent wafers interconnecting the respective electric signal paths with a topographical one-to-one correspondence.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT

AT	Austria	**LU**	Luxembourg
BR	Brazil	**MC**	Monaco
CF	Central African Republic	**MG**	Madagascar
CG	Congo	**MW**	Malawi
CH	Switzerland	**NL**	Netherlands
CM	Cameroon	**RO**	Romania
DE	Germany, Federal Republic of	**SE**	Sweden
DK	Denmark	**SN**	Senegal
FR	France	**SU**	Soviet Union
GA	Gabon	**TD**	Chad
GB	United Kingdom	**TG**	Togo
JP	Japan	**US**	United States of America

THREE-DIMENSIONALLY STRUCTURED MICROELECTRONIC DEVICE

TECHNICAL FIELD

The present invention relates to a three-dimensional micro-electronic device structure and, particularly, to such a device in which information is transferred through stacked wafers interconnected with adjacent wafers with topological one-to-one correspondence.

BACKGROUND ART

Rapid advances in the science and technology of microelectronics have led to a rapid growth in the power of digital computers, for example, in their ability to calculate and process data. Even so, in recent years the increasing need for processing two-dimensionally organized data has out-stripped the power of the electronic digital computer in several important areas of application. These applications mainly involve near real-time machine processing of video, and wideband signals from radar or digitized infrared imagery, and solving multidimensional, non-linear parallel differential equations governing such physical systems as meteorology and aerodynamics. For example, the enormous quantity of information on a single continuous-tone high-resolution photograph challenges the capabilities of even the larger-present-day electronic digital computers. As a result, near real-time analysis of the multi-image input from video or infrared imaging systems exceeds the capabilities

of conventional serially organized electronic digital computers.

Elements of present-day electronic digital computers transfer information through leads on a surface of a chip or a wafer to its edge and then through external leads to another wafer. Such interconnections limit the number of channels of information, and result in very large physical size and power consumption with associated heat removal problems, high price, and relatively long interconnections. For example, this type of architecture precludes the use of low power logic, which is too slow to be used in serially structured machines for high data rate applications. Power requirements are relatively high, also because the high frequency, long lines have to be low impedance (50 ohm) matched transmission cables. The resultant system size is large because of its requirements for printed boards, connector cables, packages and the like. The price is high because of the high labor content of fabrication of and interconnections within printed circuit boards and packaging of the chips.

SUMMARY OF THE INVENTION

The present invention avoids or overcomes the above problems by transferring information through stacked wafers, each of which is interconnected to its adjacent wafer with topological one-to-one correspondence.

It is, therefore, an object of the present invention to provide a means by which massive parallel computing channels (10^2 or more) can be effected.

Another object of the present invention is to provide for equivalent processing throughput which is 100 times higher than the fastest pipeline processors presently known.

A further object of the present invention is to provide for significantly lower dissipated power than present computing devices.

Another object of the present invention is to permit the use of low power logic.

Still another object of the invention is to permit the use of high impedance, e.g., 10 kilo-ohms, short lines of less than one inch, for example, as compared to 50 ohms on conventional high speed computers.

Another object of the invention is to provide for smaller sizes, e.g., 100 times smaller, than present computers.

Another object of the present invention is to reduce the cost of microelectronic devices by reducing their labor content.

Other aims and objects as well as a more complete understanding of the present invention will appear from the following explanation of the exemplary embodiments and the accompanying drawings thereof.

BRIEF DESCRIPTION OF THE DRAWINGS

FIG. 1 is a view of a preferred embodiment of the present invention, depicting a basic block diagram of a three-dimensional microelectronic computer;

FIG. 2 illustrates one means by which a signal is fed through a wafer by using thermomigration technology;

FIGS. 3-5 depict one means by which interconnections between wafers may be made with topological one-to-one correspondence;

FIGS. 6 and 7 illustrate a modification of the interconnection means shown in FIGS. 3-5; and

FIGS. 8a-8c illustrate means by which the interconnections, embodied as microsprings, may be fabricated.

DESCRIPTION OF THE PREFERRED EMBODIMENT

While the present invention has general applicability to any three-dimensionally structured microelectronic device, it was developed primarily for use in

constructing microelectronic computers, to meet one of the most pressing needs in rapid data processing, i.e., the processing of two-dimensionally organized information.

The present invention, as applied to digital computing, is intended to reorganize the electronic digital processor from serial to parallel architecture. In the new computer, many parallel channels, 10^2 or more, operate simultaneously on the data. This organization is more natural, and therefore more efficient, for processing two-dimensional array data. To organize a digital computer in parallel, microelectronic wafers or chips, which are densely packed with circuits must be interconnected. Because each wafer requires a large number of connecting leads, the wafers are stacked against each other and the leads are interconnected from adjacent wafers with topological one-to-one correspondence.

A basic block diagram of a typical system is shown in FIG. 1 as a computer 10. An input wafer 12 is, for example, a charge coupled serial-parallel CCD array which accepts serial electrical analog signals from input 14. An analog-to-digital converter 16 converts the signal which is then supplied to a stack of logic wafers 18. Conductor lines 20 extend through all the wafers and comprise data, address and control bus lines. In this typical system, data lines 20a can number 10^2 to 10^6, for example, and only relatively few (10 or so) serve address and control functions as shown by 20b, for example.

A stored program control unit 22 performs the control function and stores the software program. The software program consists of address and control commands. The program counter, included in the control unit, reads out the appropriate address and control information into the address and control bus lines. The address line 20b' addresses typically two wafers simultaneously, one from which the information is read out and the one to which

the information is written in. The control command determines the mode of operation of the written-in wafer. After the information has been processed on the wafer, it is similarly transferred to another wafer for processing or storage and so on. Each wafer is considered to be an independent station that can be addressed by means of the address lines. The NxN information array is transferred and processed in parallel. An output wafer 24 serves as a serial, digital input/output element. Processed information is read out through this element. This unit is also useful for local processing control. A mask of information can be introduced this unit, and then the mask can be transferred to any wafer by the bus lines. In this manner one can, for example, carry out locally different processing on different parts of the array information.

The output of the serial input/output unit can be transferred to a process evaluation unit (not shown in FIG. 1). This unit evaluates the process output and modifies the program stored in the control unit accordingly, or introduces a modified mask into the processor. In this case, there is a digital feedback system which in near real-time modifies the programming according to image and evaluation data.

The use of massive parallel organization and storage of the entire image on the wafer eliminates memory cycling and increases the computational speed of this system.

The illustrated system has been compared to the faster digital computers utilizing semiconductor memory. It is assumed words will be processed in parallel, and bits in a word will be processed serially. For 10 MHz clock rates for the three-dimensional computer, the system advantage for two-dimensional information analysis is high data-rate processing capability (ten to one hundred times faster), low power (by a factor of ten to one hundred

times), high computation capability, and small weight and volume (by a factor of ten to one thousand times). These features are mutually exclusive with present computer technology.

Processing speeds in excess of 10^9 instructions per second are required in typical applications for high altitudes surveillance, navigation, target identification or cueing, and tracking by model matching. Present computers cannot meet this requirement and appear unlikely to be able to come close to this speed in the foreseeable future. The extensive parallelism of the present invention can provide the required throughput (10^{10} instructions per second or more).

There are two basic requirements for implementing the three-dimensional computer system. It requires a feedthrough of electrical signals from one side of the wafer to the other while preserving the spatial modulation of the signal. It must be fast and work equally well in both directions, and must provide a means of interconnecting the wafers in a stack.

While it is possible to etch holes through wafers to overcoat the holes with an insulating layer and electroplate conductive materials on the wall of the bore of holes, the preferable approach employs thermomigration technology, as depicted in FIG. 2, to transfer the electrical signal from one side of the wafer to the other.

Prior work in other laboratories has shown the effects of thermomigration. Thermomigration is the phenomenon of a liquid zone in the form of a droplet, a sheet, or a rod migrating in a solid along a thermal gradient. The thermomigration of liquid droplets in solids can be understood as follows. A liquid droplet is embedded in a solid subjected to a thermal gradient. This thermal gradient causes the temperature on the

forward droplet interface to be higher than the temperature on the rear droplet interface. Since the solubility of a solid in a liquid increases with temperature, the concentration of dissolved solid atoms is higher in the liquid at the hotter forward interface of the droplet than in the liquid at the cooler rear interface of the droplet. This inequality in concentration produces a concentration gradient of dissolved solid atoms across the liquid droplet. This concentration gradient, in turn, generates a flux of dissolved solid atoms down the concentration gradient from the front to the rear interface of the liquid droplet. To feed this diffusion flux, additional solid atoms dissolve into the liquid at the forward hot face of the droplet while dissolved solid atoms deposit on the cold rear face of the droplet. As a consequence, the liquid droplet moves forward up the thermal gradient in the solid by dissolving the solid at its front, passing the atoms or molecules of the solid through itself to its rear, and redepositing these atoms behind itself.

This dissolution process was used to form one device of present invention utilizing liquid-aluminum droplets migrating up a thermal gradient in a crystal of n-type silicon. The liquid droplets migrated as explained above: solid silicon dissolved into the liquid aluminum at the forward, hotter interface of the droplet, the dissolved silicon atoms were transported through the intervening liquid, and the same silicon atoms were redeposited on the rear, cooler interface of the liquid aluminum droplet. At the rear face of the droplet, trace amounts of aluminum up to the solid solubility limit of aluminum in solid silicon at the migration temperature were incorporated in the redeposited silicon. The dissolved aluminum atoms doped the redeposited silicon in the trail behind the droplet with a p-type conductivity, thereby forming a

highly conductive channel which is insulated from the bulk by a p-n junction.

Such wafer feedthroughs are based on developments by Thomas R. Anthony and Harvey E. Cline of the General Electric Company. Known patents of these two parties include U.S. Patents, 3,895,967; 3,898,106; 3,899,361; 3,899,362; 3,901,736; 3,902,925; and 3,904,442, all dated in the months of July-September, 1975. Publications include the following: Journal of Applied Physics, Vol. 43, No. 11, November 1972, pages 4391-4395; IEEE Transactions on Electron Devices, Vol. ED-23, No. 8, August 1976, pages 818-823; Journal of Applied Physics, Vol. 47, No. 6, June 1976, pages 2316-2336 and 2550-2557; Journal of Applied Physics, Vol. 48, No. 6, June 1977, pages 2196-2201; Applied Physics Letters, Vol. 31, No. 2, 15 July 1977, pages 125-126; and Journal of Applied Physics, Vol. 48, No. 9, September 1977, pages 3943-3949.

The specific experimental setup used for thermomigration processing and the important physical parameters of the process such as thermomigration temperature, droplet size, effect of crystallographic orientation, random walk displacement, and required thermal gradient are as follows.

The basic processing sequence of thermomigration for the implementation of bus lines through a thin silicon wafer is illustrated in FIGS. 2a-2e. First, by using conventional photolithography and etching techniques, an array of holes 28 through a layer 32 of silicon dioxide (10 to 25 μm deep) was etched into a n-type silicon wafer 30 (FIG. 2). These holes were used to prevent the subsequent liquid zones from moving laterally on the silicon surface, and thus serve to preserve the registration and pattern of the deposited array. Next, the wafer was placed in the chamber of a metal evaporator where a layer of aluminum 34 was deposited (FIG. 2b). The thickness of this aluminum layer is preferably approximately equal to

the depth of the etched holes (FIG. 2). Next, excess aluminum was removed by grinding (FIG. 2c). The array of 34a of aluminum dots, half buried in the surface of the silicon, was then ready for thermomigration processing (FIG. 2d).

A slight modification of this process is presently preferred. The steps depicted in FIGS. 2a and 2b have been eliminated. Instead, aluminum dots are placed on a base silicon wafer using standard lithographic techniques. No holes are etched.

The wafer was placed in an electron beam thermomigration apparatus, which was designed to produce a very uniform vertical temperature gradient. The radiation block was made of molybdenum, which is heated by an incident electron-beam. Radiation from the hot molybdenum block produces uniform heating of one face of the silicon wafer to a temperature of 1000 to 1200°C. The wafer was suspended in a vacuum of 5×10^{-6} Torr and seated in a silicon sample holder. The aluminum-evaporated side of the wafer faced the water-cooled copper block. A temperature gradient of 10 to 100°C/cm was maintained between the two surfaces of the silicon bulk. Cylindrical radiation shields around the sample prevented unwanted radial gradients from developing in the wafer. It is also presently preferred that, instead of using the vacuum thermomigration system, a silicon epireactor with helium or nitrogen cooling gas is now utilized. As shown in FIG. 2d, the buried aluminum melted and alloyed with the silicon and then migrated up the thermal gradient, leaving behind an array 38 of heavily doped regions of p-type silicon forming the desired feedthrough part of the bus line. P-type dots were then diffused as spots 40 at the end of feedthrough lines 38. These dots 40 are sufficiently large to make contact with feedthroughs 38, which exhibit some random, rather than straight-line, migration through the wafer.

The fabrication of chip interconnections is the second major technological development required for the 3-dimensional computer. The very large number of contacts (10^4 to 10^8) imposes the need for high reliability and small contact dimensions. Also, any proposed interconnection technique must be compatible with wafers that are somewhat distorted. The concept of a bus that passes through all wafers in the stack requires that the contacts have low capacitance, have low resistivity, and be independent of polarity and orientation.

The preferred interconnect, as depicted in FIGS. 3-8, uses spring contacts, and is an outgrowth of the technology for the fabrication of miniature audio frequency tuning forks and beam lead crossovers. These crossover-type spring contacts are fabricated so that the height of the tunnel underneath is sufficient to accommodate for the distortion across a wafer and permit complete interconnection of all the contacts in parallel. The contacts may be batch-fabricated by either electroplating or vacuum evaporation, and the process is compatible with silicon technology.

The basic structure of a spring contact 50 is shown in FIG. 3. When fabricated on silicon chips, the contact on one chip is oriented at right angles to the contact on the other chip. Thus, when juxtaposed wafers 52 and 54 are mated, the springs make contact at right angles, as shown in FIG. 4, forming a cross 56. This juxtaposition increases the probability of contact and ensures a more secure interconnection. Also, this arrangement accommodates maximum chip displacement while occupying minimum space on the silicon chip. This design satisfies the need for reliable miniature contacts, even when the chips are distorted. Another advantage of this type of contact is that the stack may be disassembled and the individual chips demounted. If desired, spacers 58 (FIG. 5) of photoresist, for example, may be secured to

and between adjacent wafers 52 and 54 to limit spring deformation and prevent crossing contacts from being squashed. Greater numbers of contacts may be used, as illustrated in FIGS. 6 and 7, showing contact structures 60 and 62.

The spring-contact concept originated in the technology for producing miniature tuning forks (tunistors) and resonant gate transistors used in monolithic audio oscillators. A similar technology, known as the beam-lead crossover, was also developed at about the same time. The basic fabrication process is the same in each case. The principal steps are shown in FIGS. 8a-8c. First, a spacer 70 10 μM or more, thick is evaporated or electroplated onto a substrate 72 (FIG. 8a). Then the spring contact 74 is evaporated or electroplated on top of the spacer (FIG. 8). Finally, the spacer is etched away, and a flexible microspring bridge 74a results (FIG. 8). To secure the contact between two microsprings, each microspring is also coated with a thin layer of indium. During the assembly of the wafer stack, the wafers are heated to near the melting point of indium and the two microsprings will be bonded.

Although the invention has been described with reference to particular embodiments thereof, it should be realized that various changes and different embodiments or modifications may be made therein without departing from the spirit and scope of the invention.

CLAIMS

What is Claimed is:

1. A three-dimensional microelectronic device comprising a plurality of stacked integrated circuit wafers having top and bottom surfaces, electric signal paths extending through each of said wafers between the surfaces, and means on the surfaces of adjacent wafers interconnecting their respective electric signal paths with a topological one-to-one correspondence into an electronic function.

2. A three-dimensional microelectronic device according to Claim 1 in which said electric signal paths comprise data bus lines interconnecting individual electronic functions of said wafers into the electronic function.

3. A three-dimensional microelectronic device according to Claim 2 in which said electric signal paths comprise trails defined by thermomigration of metal droplets through each of said wafers.

4. The three-dimensional microelectronic device according to Claim 3 in which said wafers comprise silicon and said metal trails comprise aluminum.

5. The three-dimensional microelectronic device according to Claim 2 in which the individual electronic functions comprise two dimensionally arrayed computing cells and said data bus lines comprise 10^2 to 10^6 channels operating simultaneously on the data for processing thereof, and said contacts are paired in one equivalent 10^2 to 10^6 number.

6. The three-dimensional microelectronic device according to Claim 5 in which all computing cells and one of said wafers are identified for processing data bit planes.

7. The three-dimensional microelectronic device according to Claim 6 in which said bus lines comprise three types including address, control and data bus lines in which the address and control bus lines are activated by a control unit external to said stacked wafers, the address determining all computing cells on one of said wafers, in which the control lines determine the mode of operation of said addressed wafer, an array of data being transferred forth and back, from wafer-to-wafer within said stack to accomplish each time an elementary computing step.

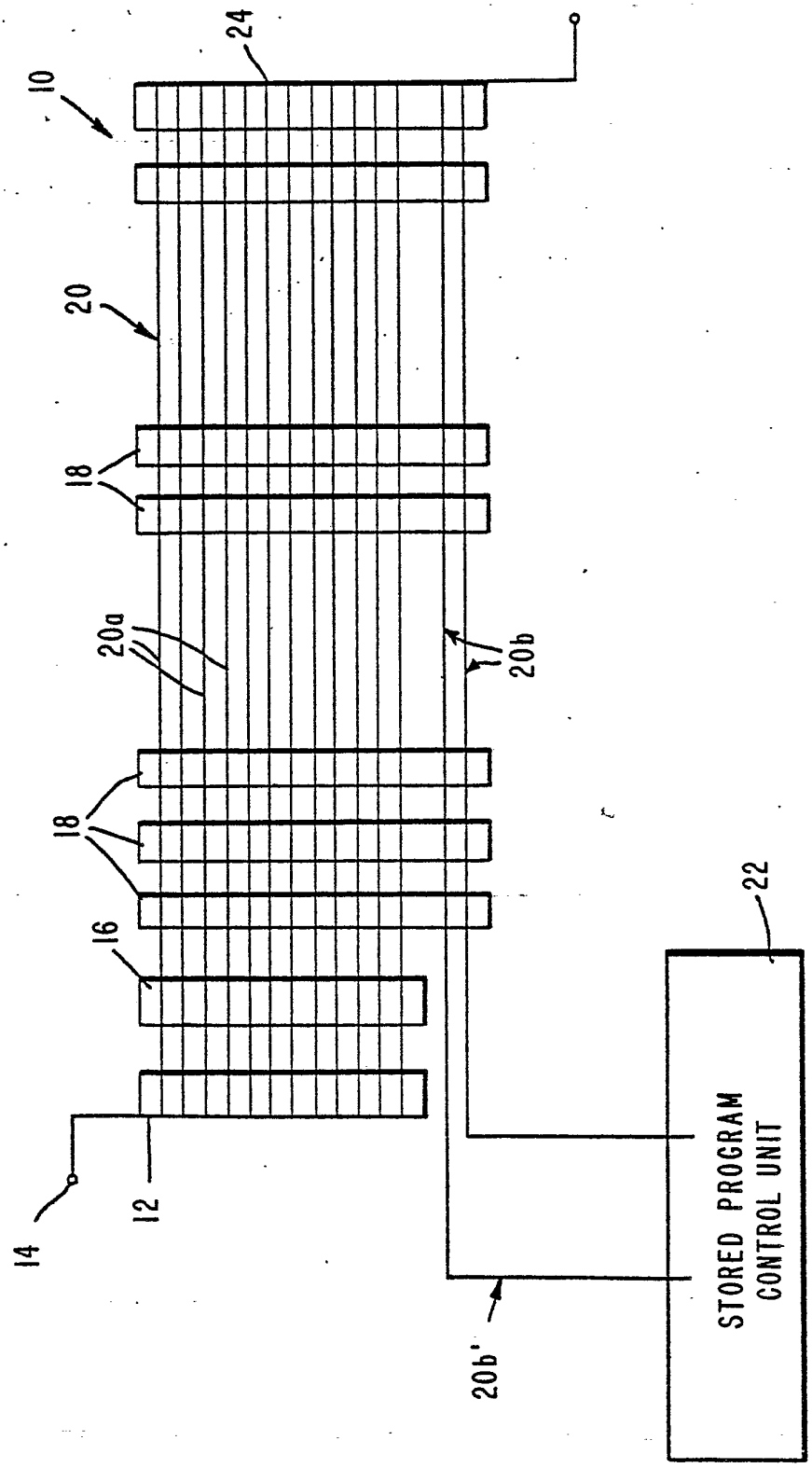


Fig. 1.

Fig. 2a.

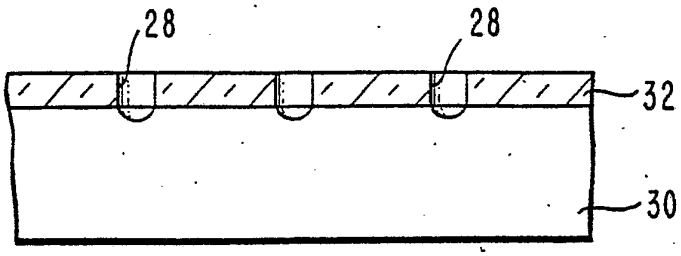


Fig. 2b.

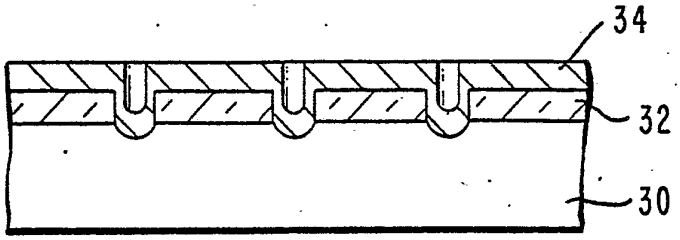


Fig. 2c.

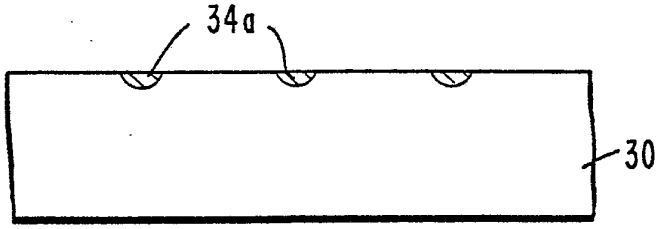


Fig. 2d.

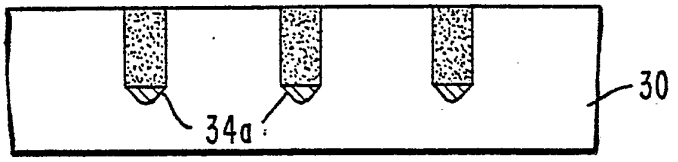


Fig. 2e.

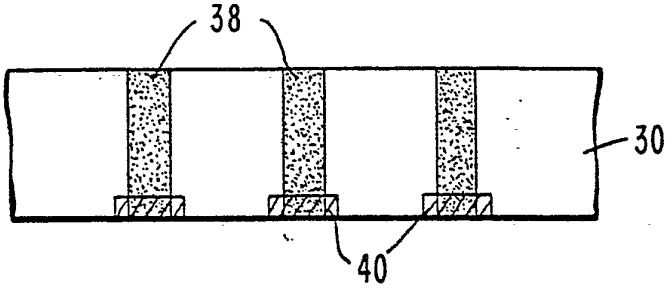


Fig. 3.

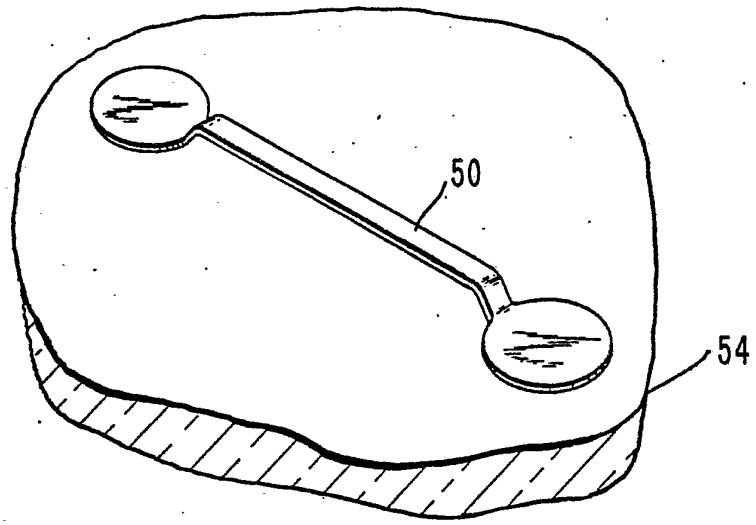
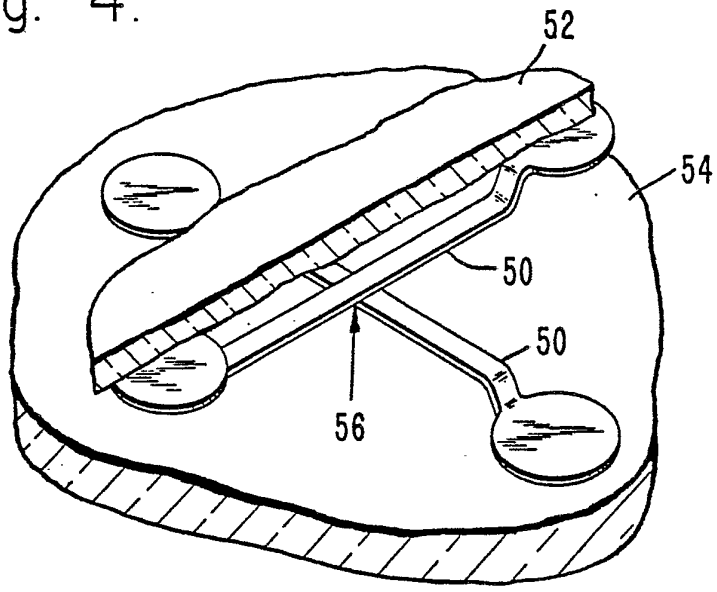


Fig. 4.



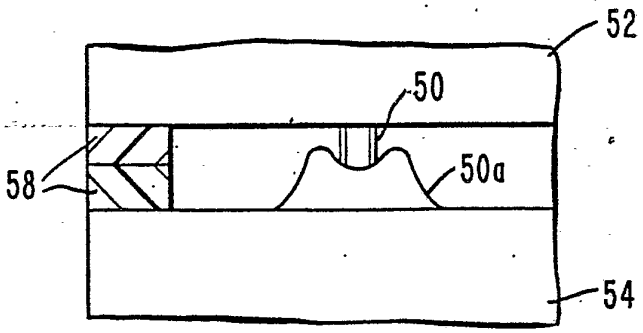


Fig. 5.

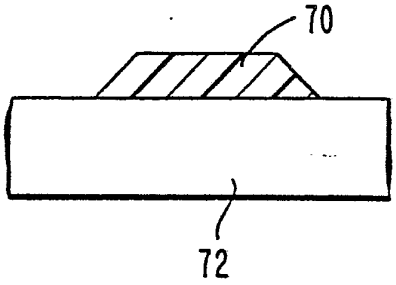


Fig. 8a.

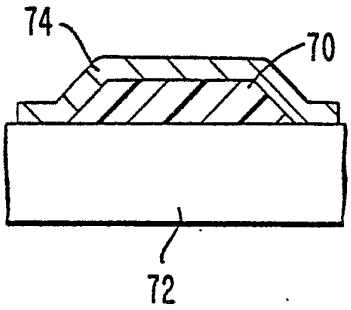


Fig. 8b.

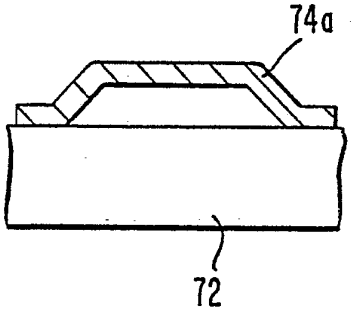


Fig. 8c.

Fig. 6.

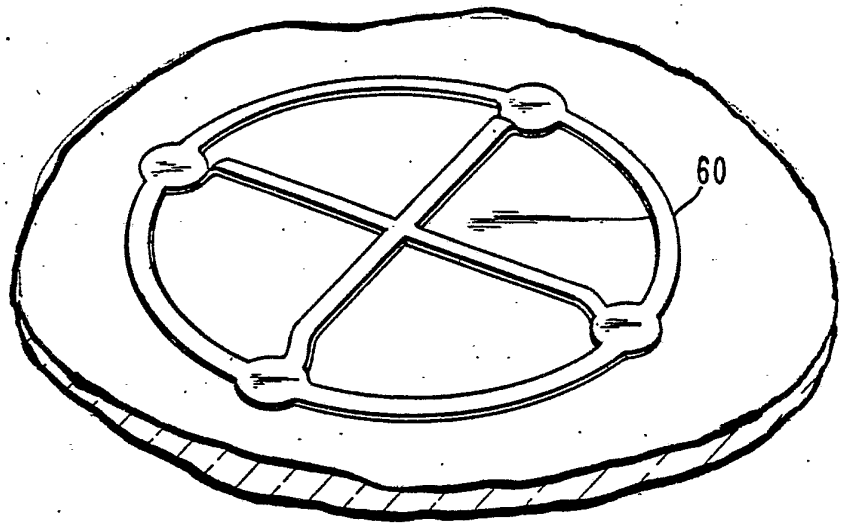
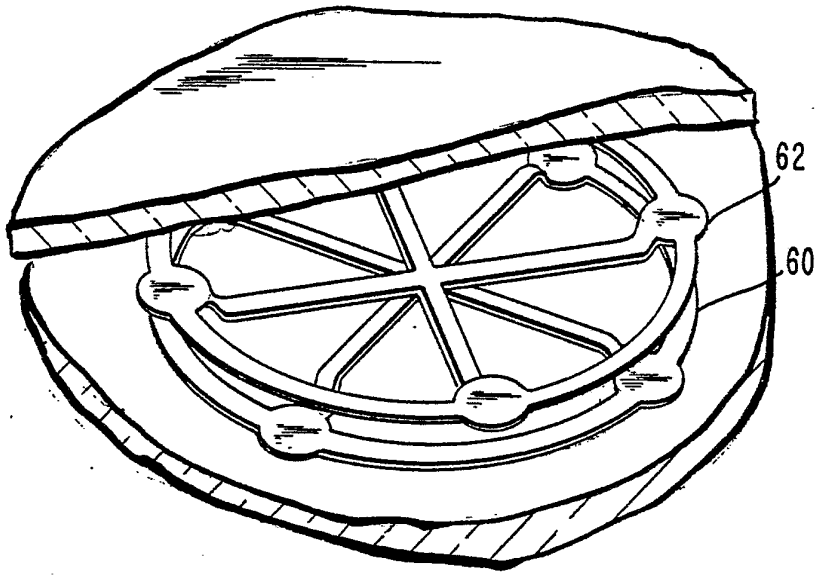


Fig. 7.



INTERNATIONAL SEARCH REPORT

International Application No PCT/US 79/01012

I. CLASSIFICATION OF SUBJECT MATTER (if several classification symbols apply, indicate all) [3]

According to International Patent Classification (IPC) or to both National Classification and IPC

U.S. CL. 357/55, 60, 68, 71, 76, 90

INT. CL.[3] H01L 29/06, 29/04, 23/48, 29/44, 29/46, 23/32,

II. FIELDS SEARCHED

Minimum Documentation Searched [4]

Classification System	Classification Symbols
U.S.	357/55, 60, 68, 71, 76, 90

Documentation Searched other than Minimum Documentation to the Extent that such Documents are Included in the Fields Searched [5]

III. DOCUMENTS CONSIDERED TO BE RELEVANT [14]

Category *	Citation of Document, [16] with indication, where appropriate, of the relevant passages [17]	Relevant to Claim No. [18]
X	US, A, 3,648,131, Published 7 MARCH 1972, Stuby	1-7
X	US, A, 3,982,268, Published 21 SEPTEMBER 1976, Anthony et al.	1-7
X	N, IEEE TRANSACTIONS ON COMPUTERS, Vol. C-20, No. 5 Issued 5 MAY 1971 (New York, N.Y.) Parks, BATCH-FABRICATED THREE-DIMENSIONAL PLANAR COAXIAL INTERCONNECTIONS FOR MICROELECTRONIC SYSTEMS	1-7
X	US, A, 4,117,588, Published 3 OCTOBER 1978, Johnson	1-7
X	CA, A, 981,799, Published 13 JANUARY 1976,	1-7
X	N, IBM TECH. DISC. BUL., Vol. 17, No. 3 Issued AUGUST, 1974 (Armonk, N. Y.) Noth	1-7

* Special categories of cited documents: [15]

"A" document defining the general state of the art

"E" earlier document but published on or after the international filing date

"L" document cited for special reason other than those referred to in the other categories

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but on or after the priority date claimed

"T" later document published on or after the international filing date or priority date and not in conflict with the application, but cited to understand the principle or theory underlying the invention

"X" document of particular relevance

IV. CERTIFICATION

Date of the Actual Completion of the International Search [2]	Date of Mailing of this International Search Report [2]
11 MARCH 1980	20 MAR 1980
International Searching Authority [1]	**Signature of Authorized Officer [20]**
ISA/US	E. Wojciechowicz E. Wojciechowicz

Form PCT/ISA/210 (second sheet) (October 1977)