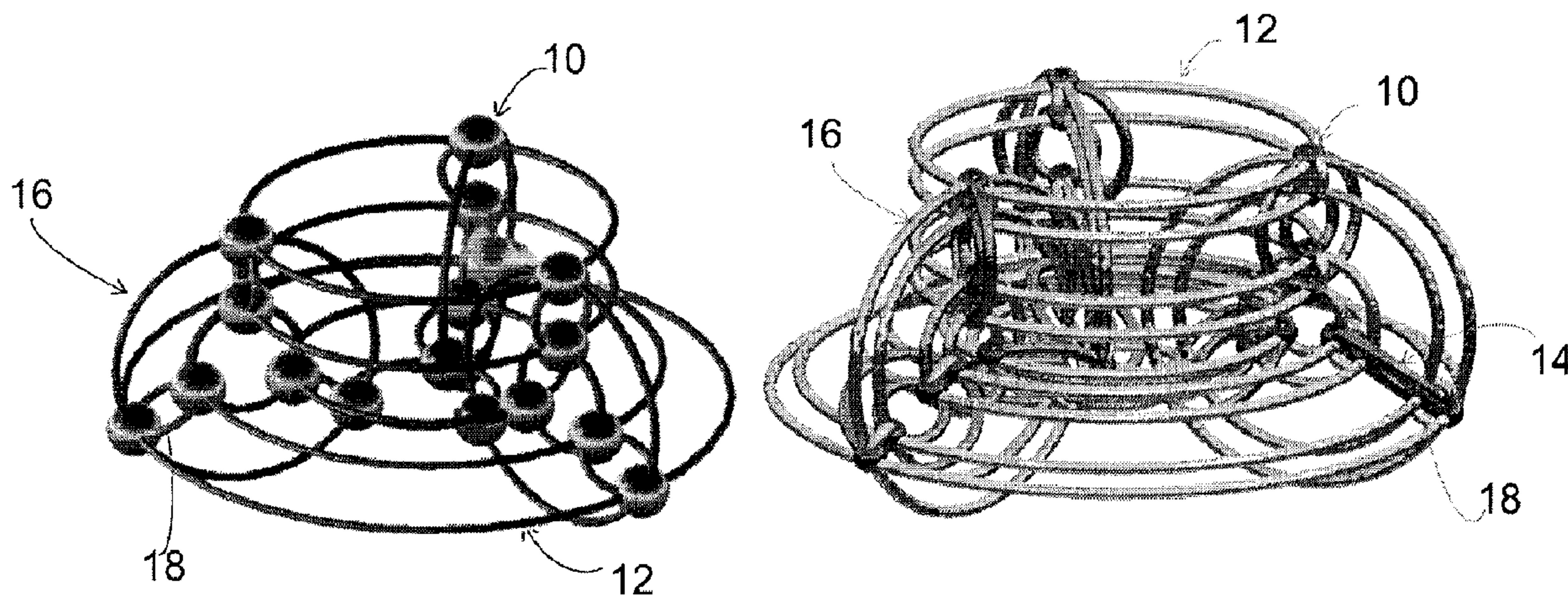




(22) **Date de dépôt/Filing Date:** 2014/08/29  
 (41) **Mise à la disp. pub./Open to Public Insp.:** 2015/03/05  
 (62) **Demande originale/Original Application:** 2 921 751  
 (30) **Priorité/Priority:** 2013/08/29 (US61/871,721)

(51) **Cl.Int./Int.Cl. H04Q 1/16** (2006.01),  
**H04L 12/935** (2013.01)  
 (71) **Demandeur/Applicant:**  
ROCKPORT NETWORKS INC., CA  
 (72) **Inventeur/Inventor:**  
OPREA, DAN, CA  
 (74) **Agent:** ANTICIPATE LAW

(54) **Titre : PROCÉDE ET APPAREIL DE GESTION DU CABLAGE ET DE L'EXTENSION DE COMMUTATEURS  
D'INTERCONNEXION DIRECTE DANS DES RESEAUX INFORMATIQUES**  
 (54) **Title: METHOD AND APPARATUS TO MANAGE THE DIRECT INTERCONNECT SWITCH WIRING AND GROWTH IN  
COMPUTER NETWORKS**



(57) **Abrégé/Abstract:**

The present invention provides a method for managing the wiring and growth of a direct interconnect network implemented on a torus or higher radix interconnect structure based on an architecture that replaces the Network Interface Card (NIC) with PCIe switching cards housed in the server. Also provided is a passive patch panel for use in the implementation of the interconnect, comprising: a passive printed circuit board that houses node to node connectivity for the interconnect; and at least one connector board plugged into the passive printed circuit board comprising multiple connectors. The multiple connectors are capable of receiving an interconnecting plug to maintain the continuity of the torus or higher radix topology when not fully enabled. The PCIe card for use in the implementation of the interconnect comprises: at least 4 electrical or optical ports for the interconnect; a local switch; a processor with RAM and ROM memory; and a PCI interface.

**ABSTRACT**

The present invention provides a method for managing the wiring and growth of a direct interconnect network implemented on a torus or higher radix interconnect structure based on an architecture that replaces the Network Interface Card (NIC) with PCIe switching cards housed in the server. Also provided is a passive patch panel for use in the implementation of the interconnect, comprising: a passive printed circuit board that houses node to node connectivity for the interconnect; and at least one connector board plugged into the passive printed circuit board comprising multiple connectors. The multiple connectors are capable of receiving an interconnecting plug to maintain the continuity of the torus or higher radix topology when not fully enabled. The PCIe card for use in the implementation of the interconnect comprises: at least 4 electrical or optical ports for the interconnect; a local switch; a processor with RAM and ROM memory; and a PCI interface.

**METHOD AND APPARATUS TO MANAGE THE DIRECT INTERCONNECT SWITCH  
WIRING AND GROWTH IN COMPUTER NETWORKS****5 FIELD OF THE INVENTION**

The present invention relates to computer network topology and architecture. In particular, the present invention relates to a method and apparatus for managing the wiring and growth of a  
10 direct interconnect switch implemented on, for example, a torus or higher radix wiring structure.

**BACKGROUND OF THE INVENTION**

15 The term Data Centers (DC) generally refers to facilities used to house large computer systems (often contained on racks that house the equipment) and their associated components, all connected by an enormous amount of structured cabling. Cloud Data Centers (CDC) is a term used to refer to large, generally  
20 off-premise facilities that similarly store an entity's data.

Network switches are computer networking apparatus that link network devices for communication/processing purposes. In other words, a switch is a telecommunication device that is capable of  
25 receiving a message from any device connected to it, and transmitting the message to a specific device for which the message was to be relayed. A network switch is also commonly referred to as a multi-port network bridge that processes and routes data. Here, by port, we are referring to an interface  
30 (outlet for a cable or plug) between the switch and the computer/server/CPU to which it is attached.

Today, DCs and CDCs generally implement data center networking using a set of layer two switches. Layer two switches process and route data at layer 2, the data link layer, which is the protocol layer that transfers data between nodes (e.g. servers) on the same local area network or adjacent nodes in a wide area network. A key problem to solve, however, is how to build a large capacity computer network that is able to carry a very large aggregate bandwidth (hundreds of TB) containing a very large number of ports (thousands), that requires minimal structure and space (i.e. minimizing the need for a large room to house numerous cabinets with racks of cards), that is easily scalable, and that may assist in minimizing power consumption.

The traditional network topology implementation is based on totally independent switches organized in a hierarchical tree structure as shown in **Figure 1**. Core switch **2** is a very high speed, low count port with a very large switching capacity. The second layer is implemented using Aggregation switch **4**, a medium capacity switch with a larger number of ports, while the third layer is implemented using lower speed, large port count (forty/forty-eight), low capacity Edge switches **6**. Typically the Edge switches are layer two, whereas the Aggregation ports are layer two and/or three, and the Core switch is typically layer three. This implementation provides any server **8** to server connectivity in a maximum of six hop links in the example provided (three hops up to the core switch **2** and three down to the destination server **8**). Such a hierarchical structure is also usually duplicated for redundancy-reliability purposes. For example, with reference to **Figure 1**, without duplication if the right-most Edge switch **6** fails, then there is no connectivity to the right-most servers **8**. In the least, core switch **2** is duplicated since the failure of the core switch **2** would generate a total data center connectivity failure. For reasons that are

apparent, this method has significant limitations in addressing the challenges of the future DC or CDC. For instance, because each switch is completely self-contained, this adds complexity, significant floor-space utilization, complex cabling and manual  
5 switches configuration/provisioning that is prone to human error, and increased energy costs.

Many attempts have been made, however, to improve switching scalability, reliability, capacity and latency in data centers.  
10 For instance, efforts have been made to implement more complex switching solutions by using a unified control plane, but such a system still uses and maintains the traditional hierarchical architecture. In addition, given the exponential increase in the number of system users and data to be stored, accessed, and  
15 processed, processing power has become the most important factor when determining the performance requirements of a computer network system. While server performance has continually improved, one server is not powerful enough to meet the needs. This is why the use of parallel processing has become of  
20 paramount importance. As a result, what was predominantly north-south traffic flows, has now primarily become east-west traffic flows, in many cases up to 80%. Despite this change in traffic flows, the network architectures haven't evolved to be optimal for this model. It is therefore still the topology of the  
25 communication network (which interconnects the computing nodes (servers)) that determines the speed of interactions between CPUs during parallel processing communication.

The need for increased east-west traffic communications led to  
30 the creation of newer, flatter network architectures, e.g. toroidal/torus networks. A torus interconnect system is a network topology for connecting network nodes (servers) in a

mesh-like manner in parallel computer systems. A torus topology can have nodes arranged in 2, 3, or more (N) dimensions that can be visualized as an array wherein processors/servers are connected to their nearest neighbor processors/servers, and wherein processors/servers on opposite edges of the array are connected. In this way, each node has  $2N$  connections in a N-dimensional torus configuration (**Figure 2** provides an example of a 3-D torus interconnect). Because each node in a torus topology is connected to adjacent ones via short cabling, there is low network latency during parallel processing. Indeed, a torus topology provides access to any node (server) with a minimum number of hops. For example, a four dimension torus implementing a  $3 \times 3 \times 3 \times 4$  structure (108 nodes) requires on average 2.5 hops in order to provide any to any connectivity. Unfortunately, large torus network implementations have not been practical for commercial deployment in DCs or CDCs because large implementations can take years to build, cabling can be complex ( $2N$  connections for each node), and they can be costly and cumbersome to modify if expansion is necessary. However, where the need for processing power has outweighed the commercial drawbacks, the implementation of torus topologies in supercomputers has been very successful. In this respect, IBM's Blue Gene supercomputer provides an example of a 3-D torus interconnect network wherein 64 cabinets house 65,536 nodes (131,072 CPUs) to provide petaFLOPs processing power (see **Figure 3** for an illustration), while Fujitsu's PRIMEHPC FX10 supercomputer system is an example of a 6-D torus interconnect housed in 1,024 racks comprising 98,304 nodes). While the above examples dealt with a torus topology, they are equally applicable to other flat network topologies.

The present invention seeks to overcome the deficiencies in such prior art network topologies by providing a system and architecture that is beneficial and practical for commercial deployment in DCs and CDCs.

5

#### **SUMMARY OF THE INVENTION**

In one aspect, the present invention provides a method for managing the wiring and growth of a direct interconnect network implemented on a torus or higher radix interconnect structure, comprising: populating a passive patch panel comprising at least one connector board having multiple connectors with an interconnect plug at each of said connectors; removing an interconnect plug from a connector and replacing said plug with a connecting cable attached to a PCIe card housed in a server to add said server to the interconnect structure; discovering connectivity of the server to the interconnect structure; and discovering topology of the interconnect structure based on the servers added to the interconnect structure.

20

In another aspect, the present invention provides a passive patch panel for use in the implementation of a torus or higher radix interconnect, comprising: a passive backplane that houses node to node connectivity for the torus or higher radix interconnect; and at least one connector board plugged into the passive backplane comprising multiple connectors. The passive patch panel may be electrical, optical, or a hybrid of electrical and optical. The optical passive patch panel is capable of combining multiple optical wavelengths on the same fiber. Each of the multiple connectors of the at least one connector board is capable of receiving an interconnecting plug

30

that may be electrical or optical, as appropriate, to maintain the continuity of the torus or higher radix topology.

In yet another aspect, the present invention provides a PCIe card for use in the implementation of a torus or higher radix interconnect, comprising: at least 4 electrical or optical ports for the torus or higher radix interconnect; a local switch; a processor with RAM and ROM memory; and a PCI interface. The local switch may be electrical or optical. The PCIe card is capable of supporting port to PCI traffic, hair pinning traffic, and transit with add/drop traffic. The PCIe card is further capable of combining multiple optical wavelengths on the same fiber.

#### 15 **BRIEF DESCRIPTION OF THE DRAWINGS**

The embodiment of the invention will now be described, by way of example, with reference to the accompanying drawings in which:

20 **FIGURE 1** displays a high level view of the traditional data center network implementation (Prior art);

**FIGURE 2** displays a diagram of a 3-dimensional torus interconnect having 8 nodes (Prior Art);

25

**FIGURE 3** displays a diagram showing the hierarchy of the IBM Blue Gene processing units employing a torus architecture (Prior Art);

30 **FIGURE 4** displays a high level diagram of a 3D and 4D torus structure according to an embodiment of the present invention;

**FIGURE 5** displays a diagram for a 36 node 2-D torus according to an embodiment of the present invention as an easy to follow example of the network interconnect;

5 **FIGURE 6** displays a three dimensional configuration of the 2-D configuration shown in Fig. 5 replicated three times and interconnected on the third dimension;

**FIGURE 7** displays a wiring diagram of the node connectivity for  
10 the 2-D torus shown in Fig. 5;

**FIGURE 8** displays a wiring diagram of the node connectivity for the 3-D torus shown in Fig. 6;

15 **FIGURE 9** displays a diagram of the passive backplane of the Top of the Rack Patch Panel (TPP) that implements the wiring for the direct interconnect network of the present invention;

**FIGURE 10** displays the TPP and interconnecting plug of the  
20 present invention;

**FIGURE 11** displays the rear view of the passive backplane of the TPP with the unpowered integrated circuits used to identify the connector ID and the patch panel ID, and the PCIe card connected  
25 to the TPP;

**FIGURE 12** displays an alternative embodiment of the passive backplane of the TPP;

30 **FIGURE 13** displays a high level view of an optical TPP implementation of the present invention;

**FIGURE 14** displays a high level view of a data center server rack with a TPP implementation in accordance with the present invention;

5 **FIGURE 15** displays a high level view of a hybrid implementation of a torus topology with nodes implemented by Top of the Rack switches and PCIe cards housed in the server;

**FIGURE 16** displays a block diagram of a PCIe card implementation  
10 in accordance with the present invention;

**FIGURE 17** displays the packet traffic flow supported by the PCIe card shown in Figure 16;

15 **FIGURE 18** displays a block diagram of a PCIe card with optical multiwavelengths in accordance with the present invention;

**FIGURE 19** displays a high level view of a TPP having a passive optical multiwavelengths implementation of the present  
20 invention;

**FIGURES 20a to 20c** displays the pseudocode to generate the netlist for the wiring of a 4D torus structure;

25 **FIGURE 21** displays the connectors installed on the TPP; and

**FIGURE 22** is the rear view of the connector board of the TPP with unpowered integrated circuits used to identify connector ID and patch panel ID.

30

**DETAILED DESCRIPTION OF THE INVENTION**

The present invention uses a torus mesh or higher radix wiring to implement direct interconnect switching for data center applications. Such architecture is capable of providing a high performance flat layer 2/3 network to interconnect tens of thousands of servers in a single switching domain.

With reference to **Figure 4**, the torus used is multidimensional (i.e. 3D, 4D, etc.), in order to promote efficiency of routing packets across the structure (although even a single dimensional torus can be used in certain deployments). In this respect, there is a minimum number of hops for any to any connectivity (e.g. a four dimension torus implementing a 3 x 3 x 3 x 4 structure (108 nodes) requires on average only 2.5 hops in order to provide any to any connectivity). Each node **10** (server) can be visualized as being connected on each dimension in a ring connection (**12**, **14**, **16**, and **18**) because the nodes **10** (servers) are connected to their nearest neighbor nodes **10** (servers), as well as nodes **10** (servers) on opposite edges of the structure. Each node **10** thereby has 2N connections in the N-dimensional torus configuration. The ring connection itself can be implemented as an electrical interconnect or as an optical interconnect, or a combination of both electrical and optical interconnect.

One problem to be addressed in such a topology, however, is how to reduce deployment complexity by promoting wiring simplification and simplicity when adding new nodes in the network without impacting the existing implementation. This is one aspect of the present invention, and this disclosure

addresses the wiring issues when implementing large torus or higher radix structures.

**Figure 5** displays a simple 2D torus wiring diagram for a 6 x 6 thirty-six node configuration for ease of explanation. As shown, the structure is a folded 2D torus wherein the length of each connection (**12**, **13**) is equivalent throughout. Each node **10** in this diagram represents a server interconnected via a PCIe switch card **41** (shown in **Figure 16** for instance) that is housed in the server.

**Figure 6** displays a three dimensional configuration build using the 2D configuration of **Figure 5**, but replicated three times and interconnected on the third dimension.

15

**Figure 7** displays the wiring diagram for the two dimensional torus structure shown in **Figure 5**. In the implementation shown, each of the 36 nodes **10** has connectors **21** (which can, for instance, be a Very High Density Cable Interconnect VHDCI connector supplied by Molex or National Instruments, etc.) with four connections (north (N), south (S), east (E), west (W)) that provide the switch wiring when the cable from the PCIe card **41** (not shown) is plugged in. In order to simplify the wiring, the connectors **21** are interconnected in a passive backplane **200** (as shown in **Figure 9**) that is housed by a Top of the rack Patch Panel (TPP) **31** (as shown in **Figures 10** and **14**). The passive backplane **200** presented in **Figure 9** shows three fields: the main field (as shown in the middle of the diagram in dashed lines) populated with the 42 connectors **21** implementing a 2D 7 x 6 torus configuration, the field on the left (in dashed lines) populated with the 2 groups of 6 connectors **21** for expansion on the third dimension, and the field on the right (in dashed

25

30

lines) with 2 groups of 6 connectors **21** to allow for expansion on the fourth dimension. The 3D expansion is implemented by connecting the 6 cables (same type as the cables connecting the PCIe card **41** to the TPP connector **21**) from the TPP to the TPP on a different rack **33** of servers. The TPP patch panel backplane implementation can even be modified if desired, and with a simple printed circuit board replacement (backplane **200**) a person skilled in the art can change the wiring as required to implement different torus structures (e.g. 5D, 6D, etc.). In order to provide the ability to grow the structure without any restrictions or rules to follow when adding new servers in the rack **33**, a small interconnecting plug **25** may be utilized. This plug **25** can be populated at TPP manufacture for every connector **21**. This way, every ring connection is initially closed and by replacing the plug **25** as needed with PCIe cable from the server the torus interconnect is built.

**Figure 8** presents the wiring diagram for a three dimensional torus structure. Note for instance the 6 connections shown at the nodes at the top left of the diagram to attach the PCIe cables to the 3D structure: +X, -X, +Y, -Y, +Z and -Z. The TPP implementation to accommodate the 3D torus cabling is designed to connect any connector **21** to every other connector **21** following the wiring diagram shown in **Figure 8**.

The novel method of generating a netlist of the connectivity of the TPP is explained with the aid of pseudocode as shown at **Figures 20a** to **20c** for a 4D torus wiring implementation (that can easily be modified for a 3D, 5D, etc. implementation or otherwise). For the 3D torus (Z, Y, X) each node **10** will be at the intersection of the three rings - ringZ, ringY and ringX. If a person skilled in the art of network architecture desires to interconnect all the servers in a rack **33** (up to 42 servers; see the middle section of **Figure 9** as discussed above) at once,

there are no restrictions - the servers can be wired in random fashion. This approach greatly simplifies the deployment - you add the server, connect the cable to the TPP without any special connectivity rules, and the integrity of the torus structure is maintained. The network management system that a person skilled in the art would know how to implement will maintain a complete image of the data center network including the TPP and all the interconnected servers, which provides connectivity status and all the information required for each node.

10 As shown in **Figure 11**, each PCIe card **41** (housed in a node server) has connectivity by cable **36** to the TPP. The cable **36** connecting the PCIe card **41** to the TPP provides connectivity to the 8 ports **40** (see **Figure 16**) and also provides connectivity to the TPP for management purposes. The backplane **200** includes  
15 unpowered electronic devices / integrated circuit (IC) **230** attached to every connector **21**. Devices **230** are interrogated by the software running on the PCIe card **41** in order to get the connector ID where it is connected. Every device **230** attached to the connector uses a passive resistor combination that uniquely  
20 identifies every connector.

The TPP identification mechanism (patch panel ID) is also implemented using the electronic device **240** which may be programmed at installation. The local persistent memory of device **240** may also hold other information - such as  
25 manufacturing date, version, configuration and ID. The connectivity of device **240** to the PCIe cards permits the transfer of this information at software request.

At the card initialization the software applies power to the IC  
30 **230** and reads the connector **21** ID. A practical implementation requires wire connectivity - two for power and ground and the third to read the connector **21** ID using "1-Wire" technology.

In a similar fashion, the patch panel ID, programmed at installation with the management software, can be read using the same wiring as with IC **230**. The unpowered device **240** has non-volatile memory with the ability to support read/write transactions under software control. IC **240** may hold manufacturing information, TPP version, and TPP ID.

**Figure 12** displays another passive patch panel implementation option using a separate printed circuit board **26** as a backplane. This implementation can increase significantly the number of servers in the rack and also provides flexibility in connector/wiring selection.

The printed circuit board **23** supporting the connectors **21** is plugged via high capacity connectors **22** to the backplane **26**. The printed circuit board **24** also has high capacity connectors **22** and is also plugged into the backplane **26** to provide connectivity to the connector board **23**.

The high capacity connectors **21** on the board **24** can be used to interconnect the TPPs rack **33** to rack **33**.

The direct interconnect wiring is implemented on the backplane **26**. Any time the wiring changes (for different reasons) the only device to change is the backplane **26**. For example, where a very large torus implementation needs to change (e.g. for a 10,000 server configuration the most efficient 4D torus would be a 10 x 10 x 10 configuration as opposed to trying to use a 6 x 7 x 16 x 15; and for a 160,000 server deployment the most efficient configuration would be a 20 x 20 x 20 x 20), you can accommodate these configurations by simply changing the backplane **26** while maintaining the connector boards **23** and **24** the same.

**Figure 13** displays an optical patch panel implementation. Such implementation assumes port to port fiber interconnect as per

the wiring diagram presented in **Figures 5** or **6** (2D or 3D torus). The optical connectors on boards **28** and **29** are interconnected using optical fiber **27** (e.g. high density FlexPlane optical circuitry from Molex, which provides high density optical routing on PCBs or backplanes). The optical TPP is preferably fibered at manufacturing time and the optical plugs **250** should populate the TPP during manufacturing. The connectors and the optical plugs **250** are preferably low loss. The connector's optical loss is determined by the connector type (e.g. whether or not it uses micro optical lenses for collimation) and the wavelength (e.g. single mod fiber in C band introduces lower optical loss than multimode fiber at 1340 nm).

Another implementation option for the optical TPP is presented in **Figure 19**. This implementation drastically reduces the number of physical connections (fibers) using optical wavelength multiplexing. The new component added to the TPP is the passive optical mux-demux **220** that combines multiple optical wavelengths on the same fiber. The fibers **27** interconnects the outputs of the mux-demux **220** to implement the optical direct interconnect torus structure. To connect two different racks (TPP to TPP), connector **222** is used. This implementation requires a modified version of the PCIe card **41** as shown in **Figure 18**. The card **41** includes the optical mux-demux **220**, optical transmitters **225** on different wavelengths, and optical receivers **224**.

The TPP can also be deployed as an electrical/optical hybrid implementation. In such a case, the torus nodes would have optical ports and electrical ports. A hybrid implementation would usually be used to provide connectivity to very large data centers. You could use the electrical connectivity at the rack level and optical connectivity in all rack to rack or geographical distributed data center interconnects. The

electrical cables are frequently used for low rate connectivity (e.g. 1Gbps or lower rate 10/100Mbps). Special electrical cables can be used at higher rate connectivity (e.g. 10 Gbps). The higher rate interconnect network may use optical transmission,  
5 as it can offer longer reach and can support very high rates (e.g. 100 Gbps or 400 Gbps).

**Figure 15** shows a combined deployment using a Top of the Rack (ToR) switch **38** and a PCIe card **41** based server interconnect in  
10 a torus structure that is suited to implement hybrid compute servers and storage server configurations. The PCIe **41** based implementation has the advantage of increased add/drop bandwidth since the PCI port in a server can accommodate substantially more bandwidth than a fixed switch port bandwidth (e.g. 1 Gbps  
15 or 10 Gbps). The PCIe card **41** supporting the 4D torus implementation can accommodate up to 8 times the interconnect bandwidth of the torus links.

The ToR switch **38** is an ordinary layer 2 Ethernet switch. The switch provides connectivity to the servers and connectivity to  
20 other ToR switches in a torus configuration where the ToR switch is a torus node. According to this embodiment of the invention the ToR switches **38** and the PCIe cards **41** are interconnected further using a modified version of the TPP **31**.

**Figure 16** displays the block diagram of the PCIe card  
25 implementation for the present invention. This card can be seen as a multiport Network Interface Card (NIC). The PCIe card **41** includes a processor **46** with RAM **47** and ROM **48** memory, a packet switch **44** and the Ethernet PHY interface devices **45**. The card **41**  
30 as shown has a PCIe connection **42** and 8 interface ports **40**, meaning the card as shown can provide for the implementation of up to a four dimension torus direct interconnect network.

**Figure 17** displays the packet traffic flows supported by the card **41**. Each port **40** has access to the PCI port **42**. Therefore, in the case of port to PCI traffic (as shown by **400**), the total bandwidth is eight times the port capacity given that the total number of ports **40** is 8. The number of ports determines the torus mesh connectivity. An eight port PCIe Card implementation enables up to a four dimension torus (x+, x-, y+, y-, z+, z- and w+, w-).

A second type of traffic supported by the card **41** is the hair pinning traffic (as shown by **410**). This occurs where traffic is switched from one port to another port; the traffic is simply transiting the node. A third type of traffic supported by the card **41** is transit with add/drop traffic (as shown at **420**). This occurs when incoming traffic from one port is partially dropped to the PCI port and partially redirected to another port, or where the incoming traffic is merged with the traffic from the PCI port and redirected to another port.

The transit and add/drop traffic capability implements the direct interconnect network, whereby each node can be a traffic add/drop node.

CLAIMS:

We claim:

- 5        1. A passive patch panel for use in the implementation of a  
torus or higher radix interconnect, comprising:  
  
a passive printed circuit board that houses node to node  
connectivity for the torus or higher radix interconnect;  
10        and  
  
at least one connector board plugged into the passive  
printed circuit board comprising multiple connectors.
- 15        2. The passive patch panel of claim 1 wherein the passive  
patch panel is electrical, optical, or a hybrid of  
electrical and optical.
- 20        3. The passive patch panel of claim 2 wherein the passive  
patch panel is optical.
- 25        4. The passive patch panel of claim 3 wherein the optical  
passive patch panel is capable of combining multiple  
optical wavelengths on a same fiber.
- 30        5. The passive patch panel of claim 2 wherein the passive  
patch panel is a hybrid of electrical and optical, and  
wherein electrical connectivity is provided at the rack  
level and optical connectivity is provided in all rack to  
rack or geographically distributed data centre  
interconnects.

- 5 6. The passive patch panel of any one of claims 1 to 5 wherein each of the multiple connectors of the at least one connector board is capable of receiving an interconnecting plug to maintain the continuity of the torus or higher radix topology.
7. The passive patch panel of claim 6 wherein the interconnecting plug is electrical or optical.
- 10 8. The passive patch panel of any one of claims 1 to 7, wherein said multiple connectors comprise fields of multiple connectors, namely:
- 15 a main field of connectors for implementing a 2D torus interconnect network;
- a second field of connectors to allow for expansion of the network to a 3D torus interconnect network;
- 20 and a third field of connectors to allow for expansion of the network to a 4D torus interconnect network.
- 25 9. The passive patch panel of claim 8 wherein the main field comprises 42 connectors in a 7X6 torus configuration, the second field comprises 12 connectors in 2 groups of 6, and the third field comprises 12 connectors in 2 groups of 6.
- 30 10. The passive patch panel of claim 8 or 9 wherein each connector is initially populated by an interconnecting plug to initially close all connections, and wherein each of said plugs is capable of being replaced by a cable attached

to a PCIe card from a server to build an interconnect network.

5 11. The passive patch panel of claim 10 wherein the printed circuit board includes unpowered electronic devices using a passive resistor combination attached to every connector, and wherein when a plug populating a connector is replaced by a cable attached to a PCIe card the device attached to the connector is interrogated by software running on the  
10 PCIe card in order to obtain an ID of the connector.

12. The passive patch panel of any one of claims 1 to 7, wherein said multiple connectors comprise fields of multiple connectors, namely:

15 a main field of connectors for implementing a torus interconnect network in N dimensions;

20 a second field of connectors for expanding the network to a N+1 dimension torus interconnect network;

and a third field of connectors for expanding the network to a N+2 dimension torus interconnect network, wherein N is at least 3.

25 13. The passive patch panel of claim 12 wherein each connector is initially populated by an interconnecting plug to initially close all connections, and wherein each of said plugs is capable of being replaced by a cable attached to a  
30 PCIe card from a server to build an interconnect network.

14. The passive patch panel of claim 13 wherein the printed circuit board includes unpowered electronic devices using a passive resistor combination attached to every connector, and wherein when a plug populating a connector is replaced  
5 by a cable attached to a PCIe card the device attached to the connector is interrogated by software running on the PCIe card in order to obtain an ID of the connector.
15. The passive patch panel of any one of claims 1 to 14  
10 wherein an additional printed circuit board with multiple connectors is used to increase a number of servers that may be used in a rack, wherein the printed circuit board is plugged via high capacity connectors to the backplane to provide connectivity to the connector board, and wherein  
15 high capacity connectors on the board may be used to patch panels rack to rack.

1/16

PRIOR ART

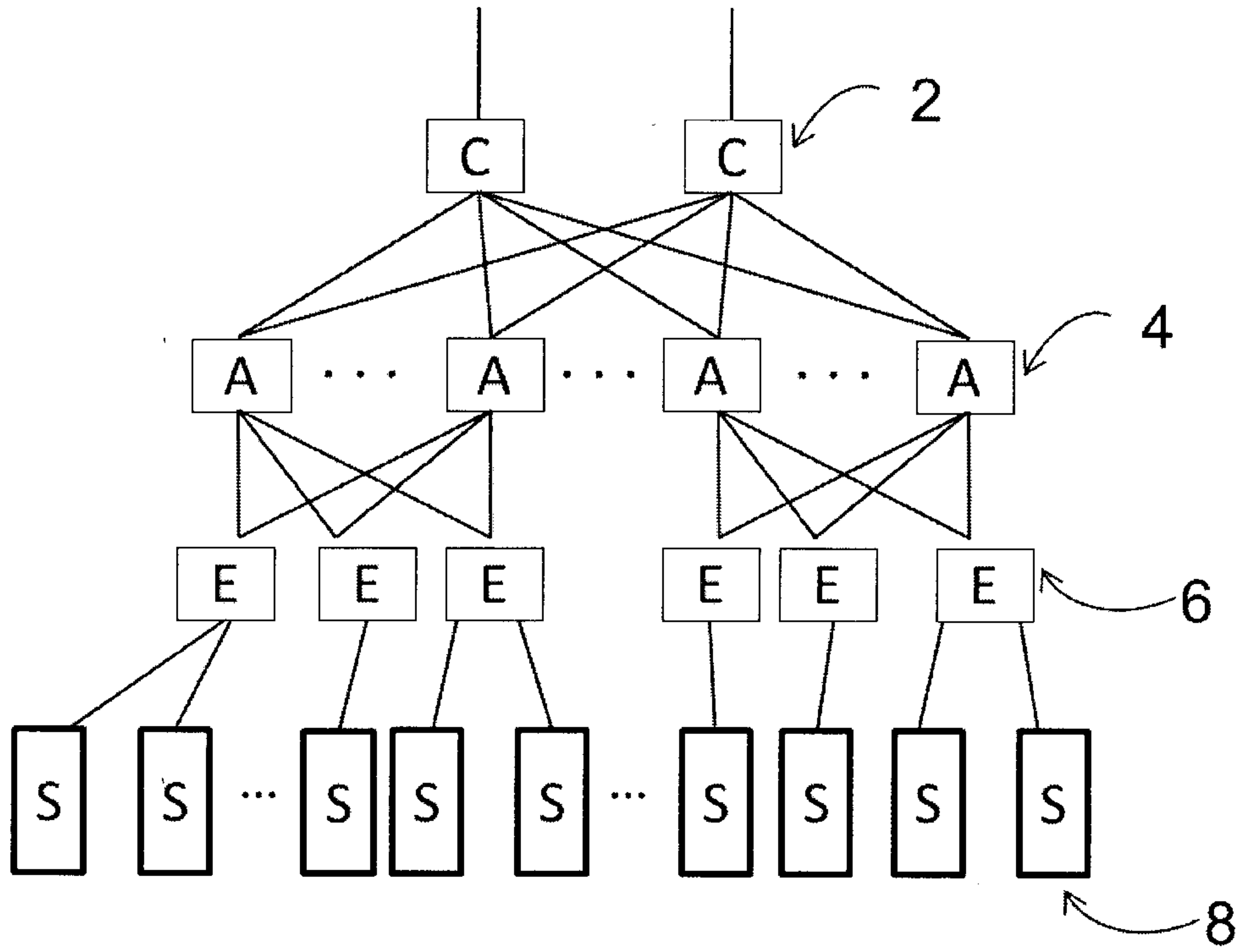


FIGURE 1

2/16

PRIOR ART

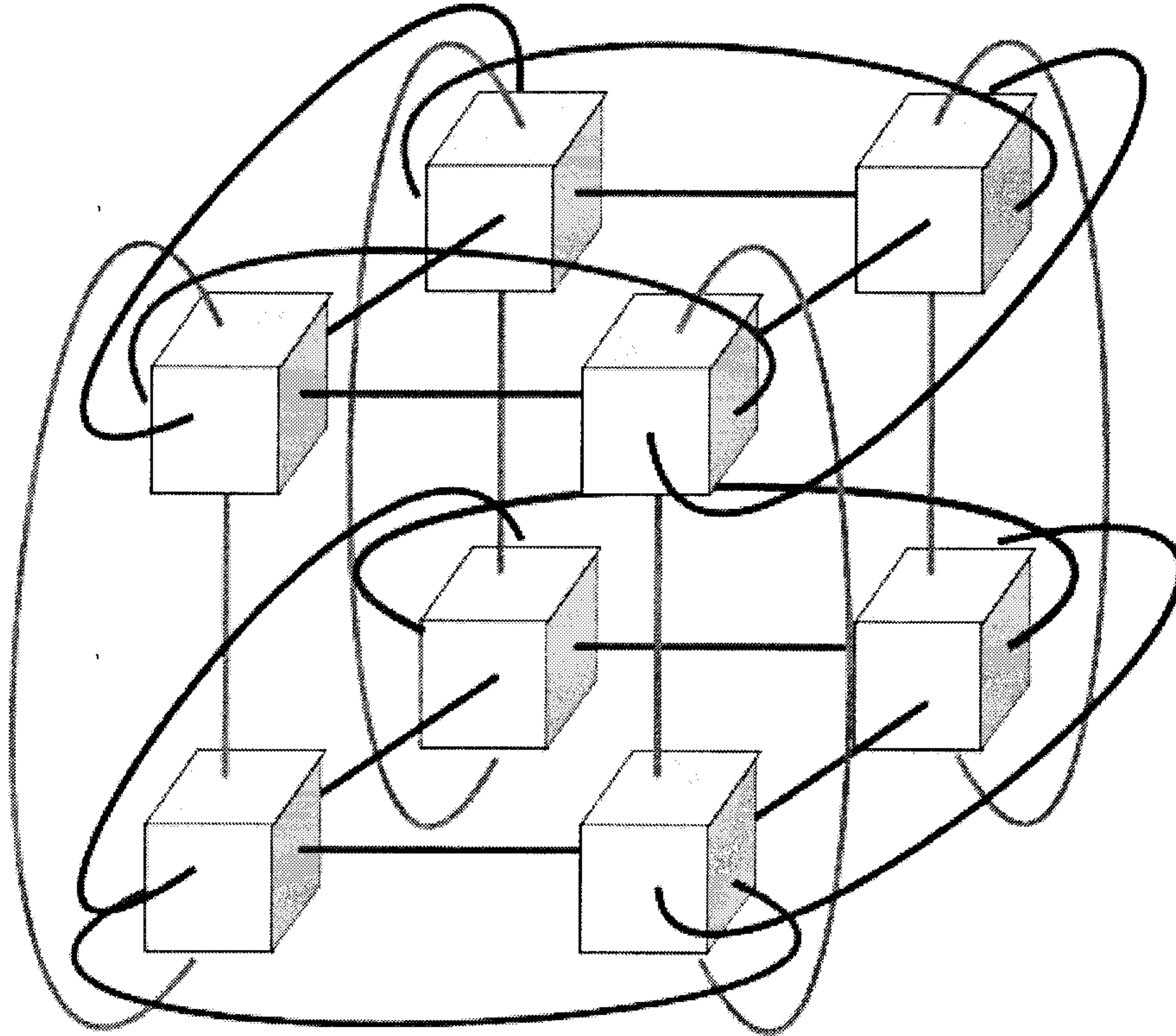


FIGURE 2

3/16

PRIOR ART

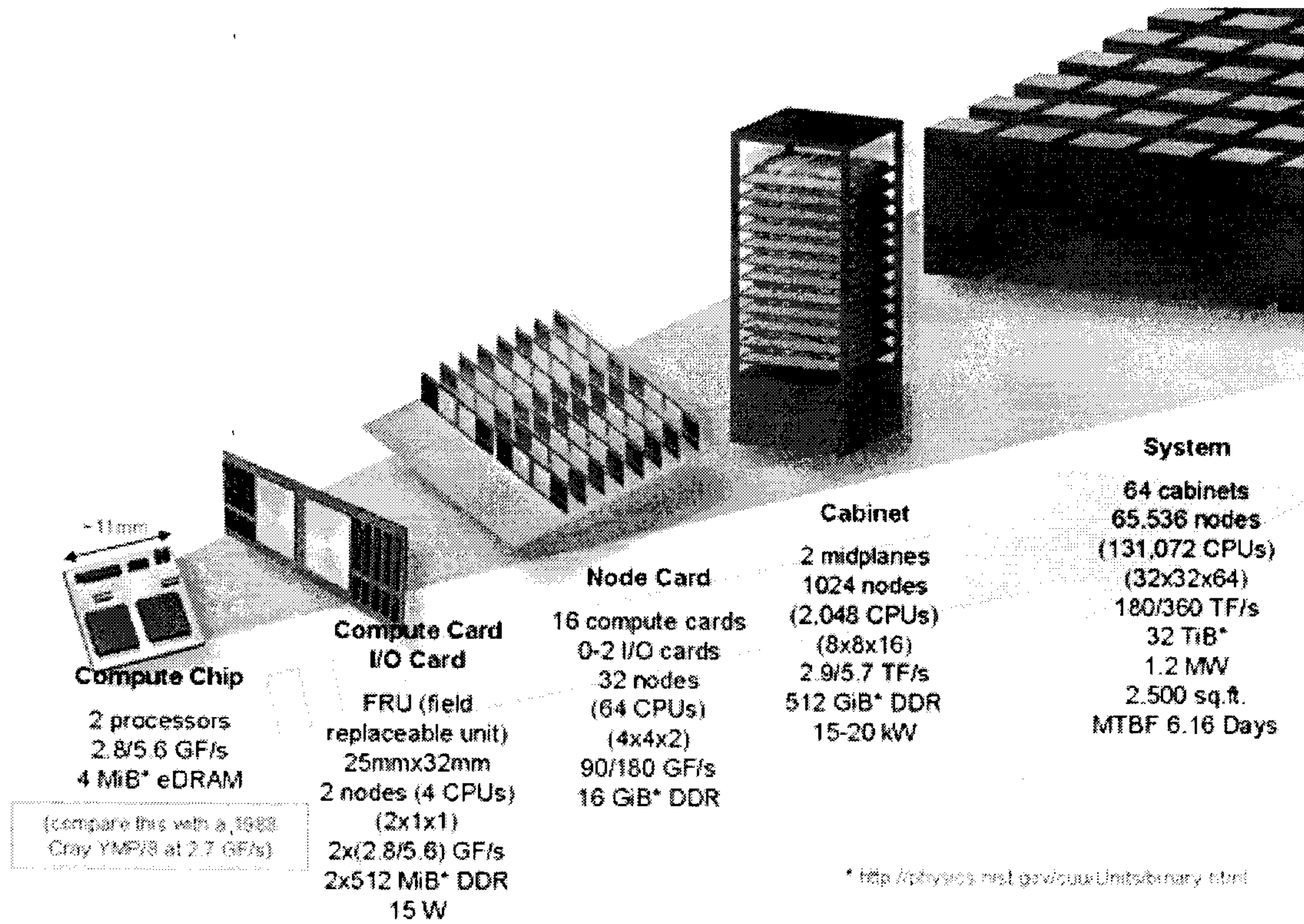


FIGURE 3

4/16

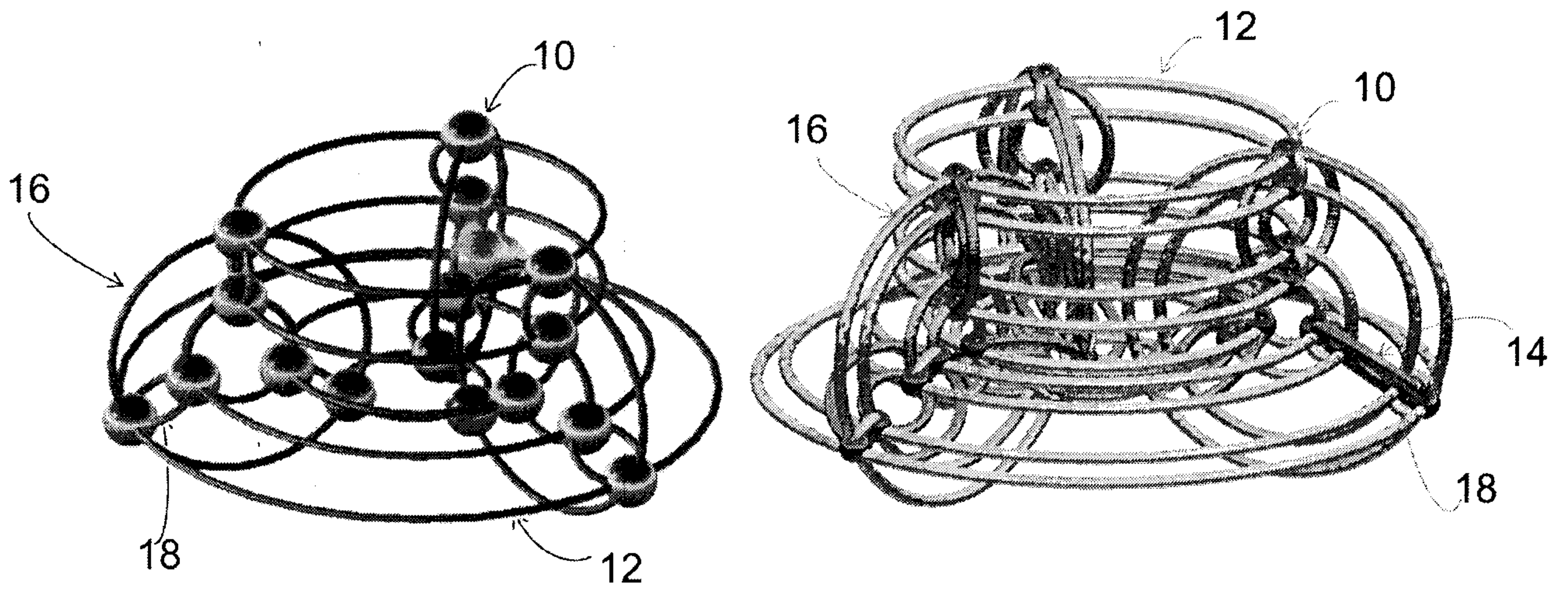


FIGURE 4

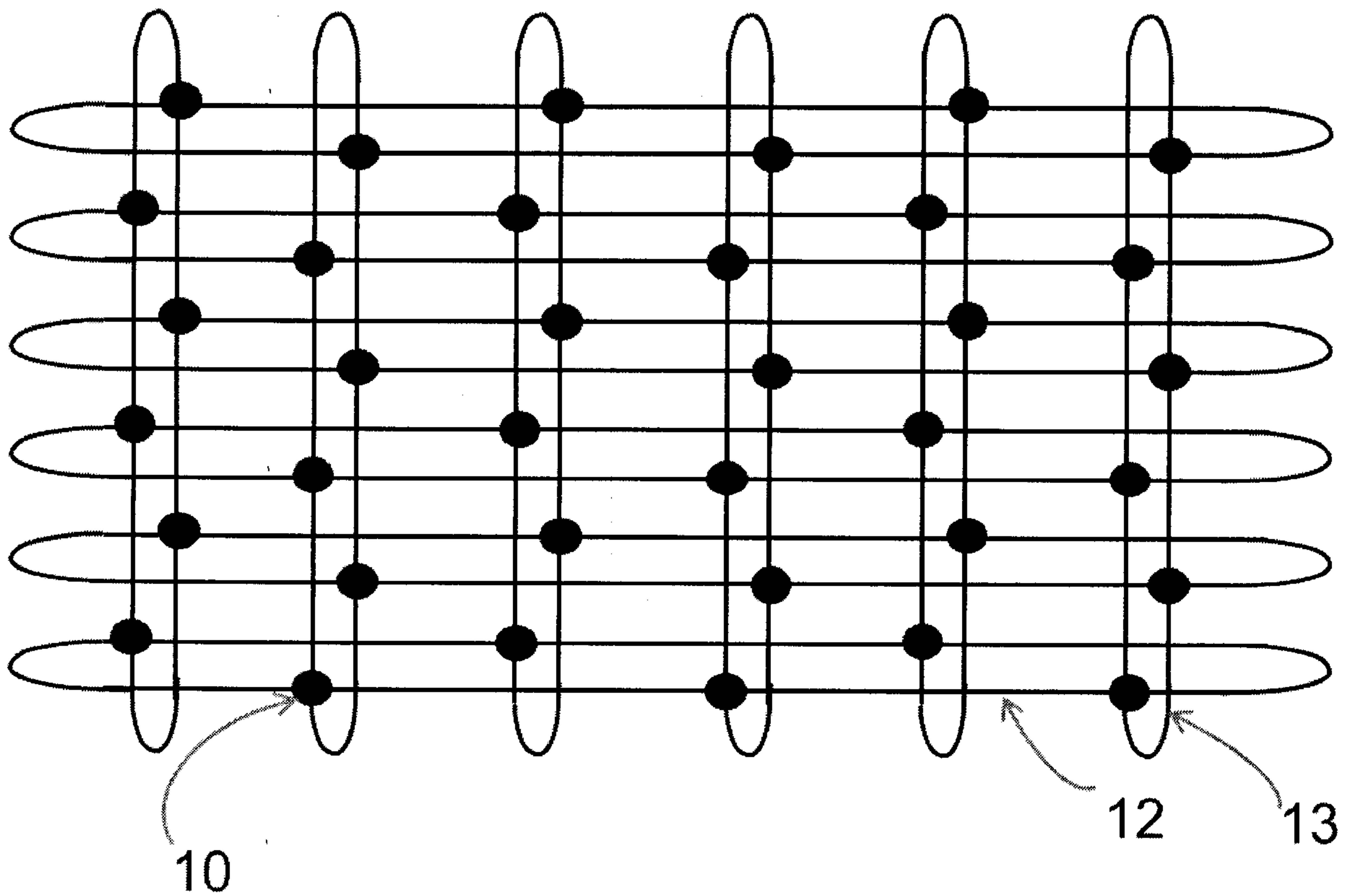


FIGURE 5

5/16

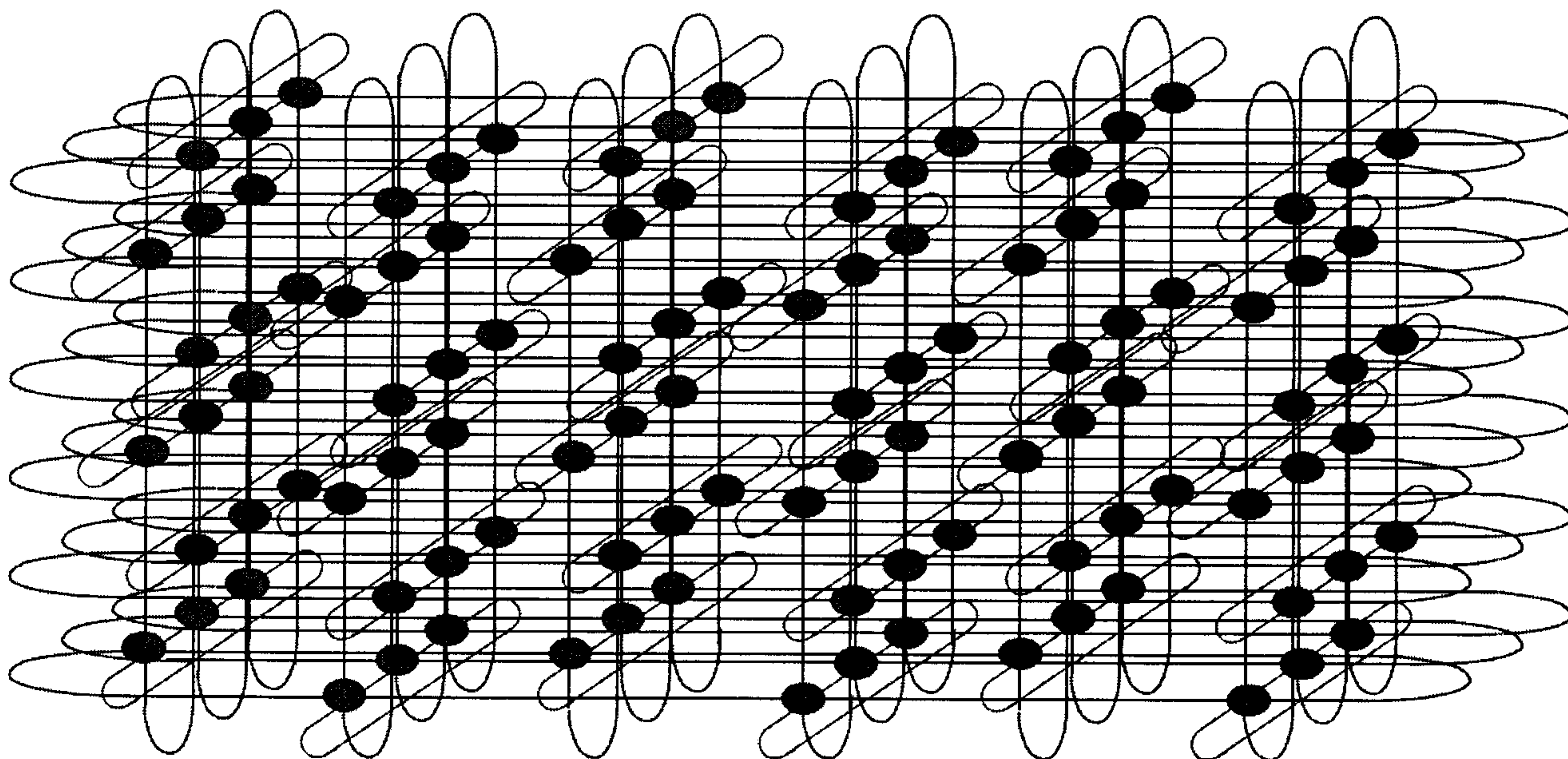


FIGURE 6

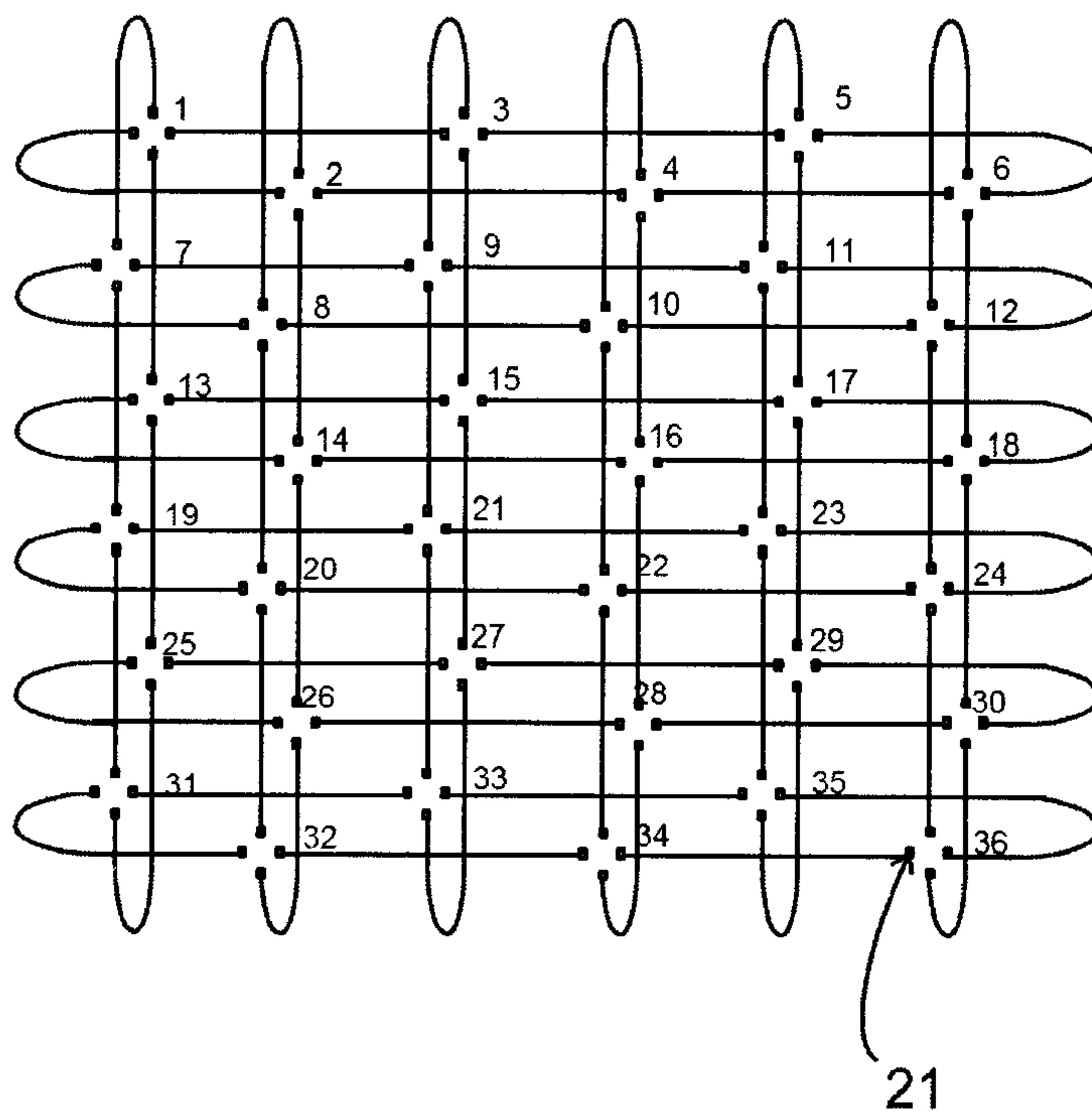


FIGURE 7

6/16

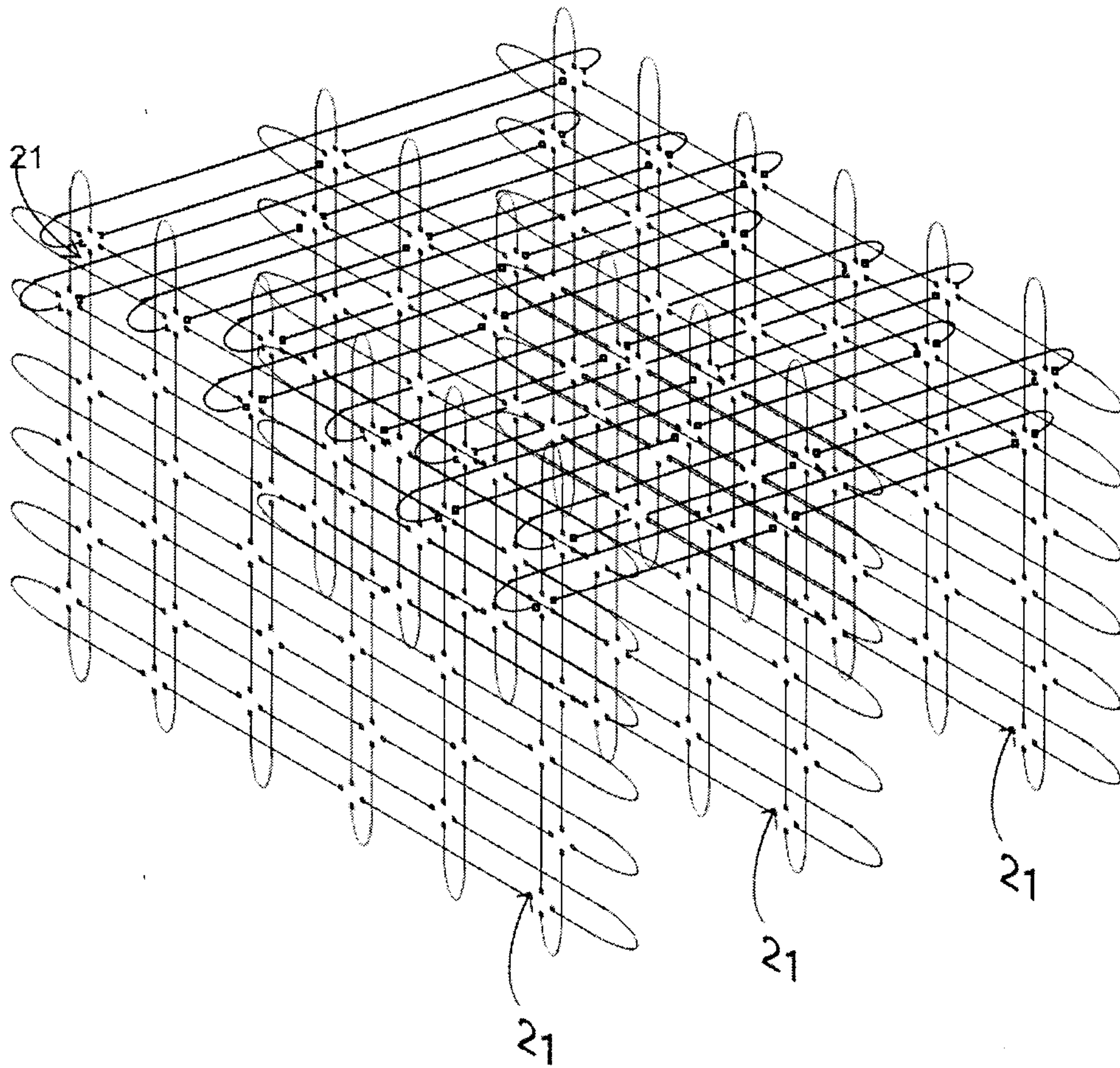


FIGURE 8

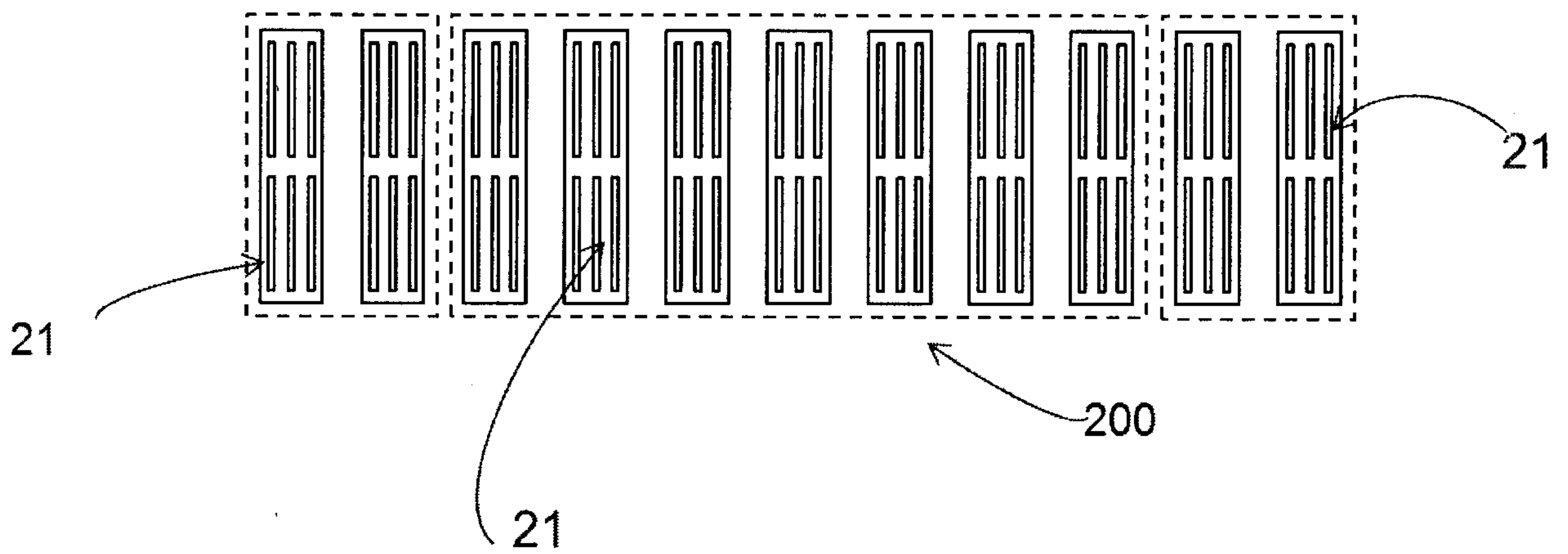


FIGURE 9

7/16

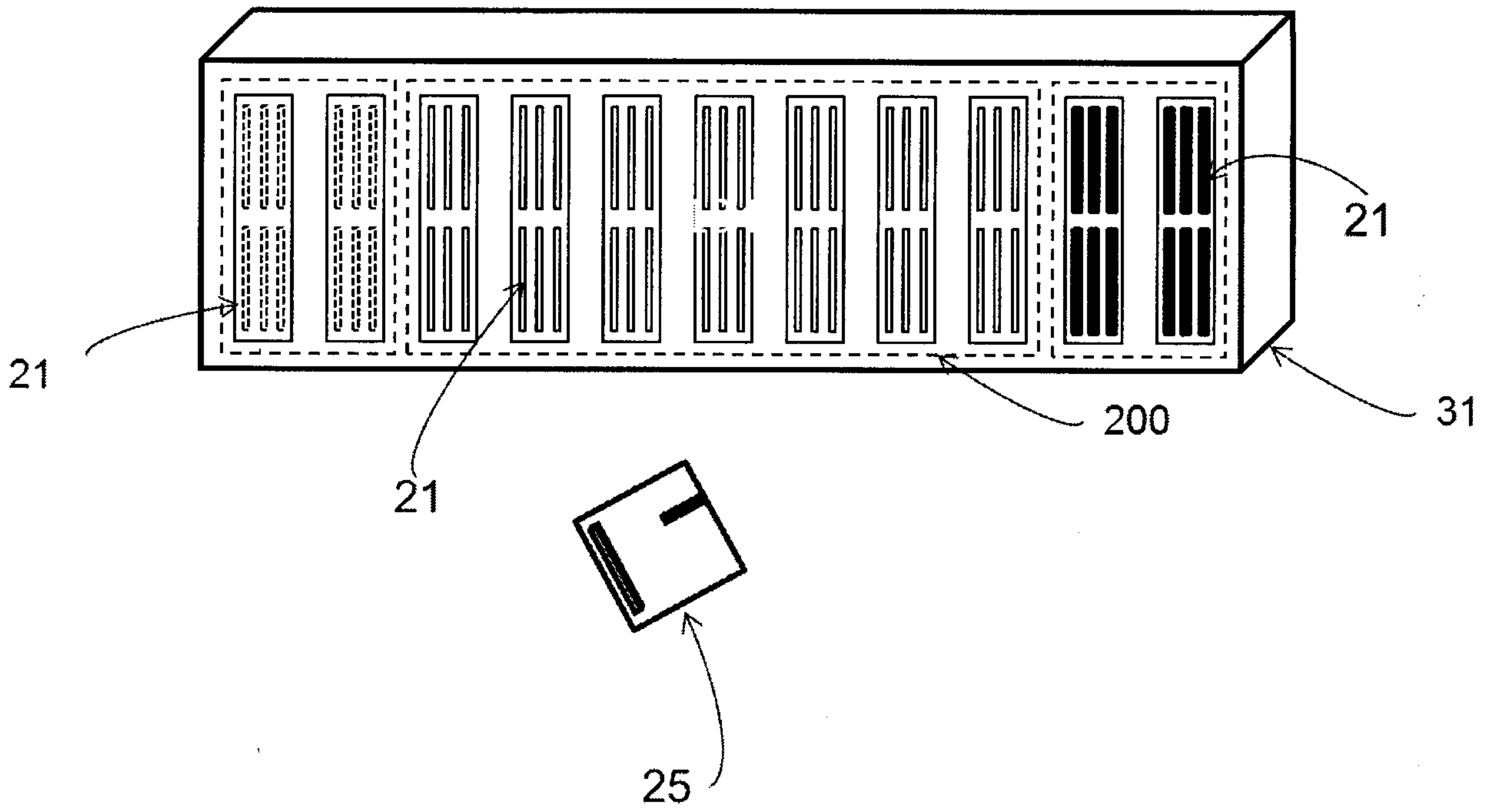


FIGURE 10

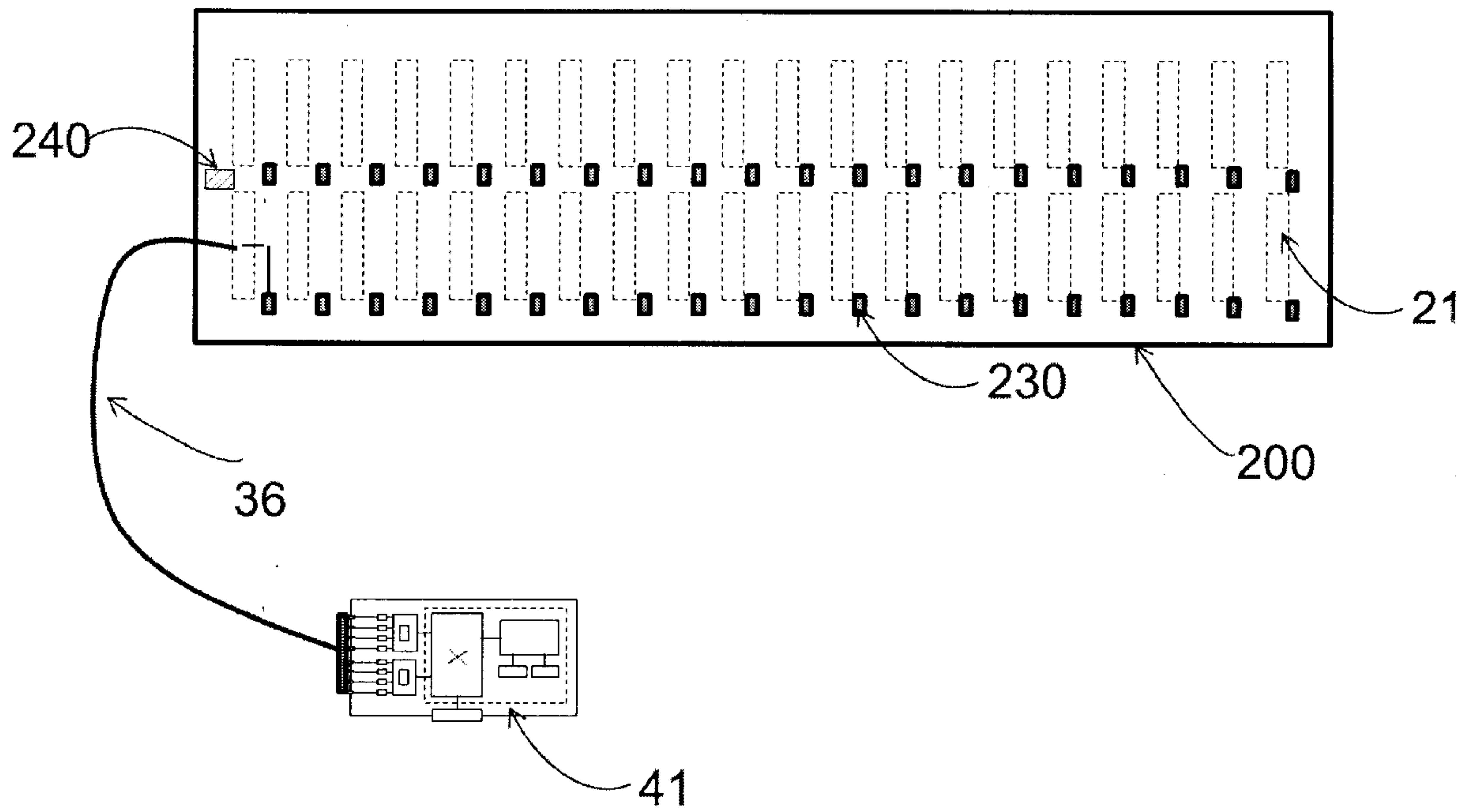


FIGURE 11

8/16

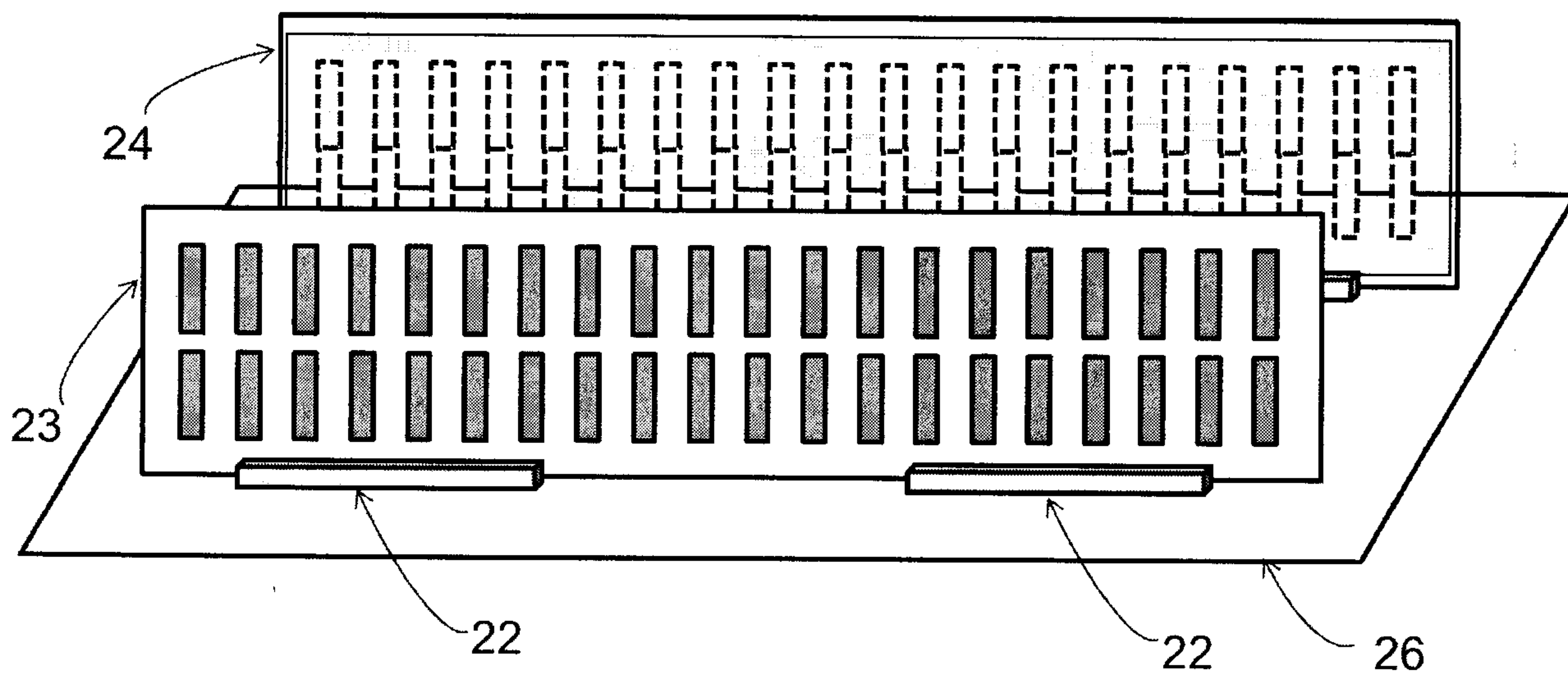


FIGURE 12

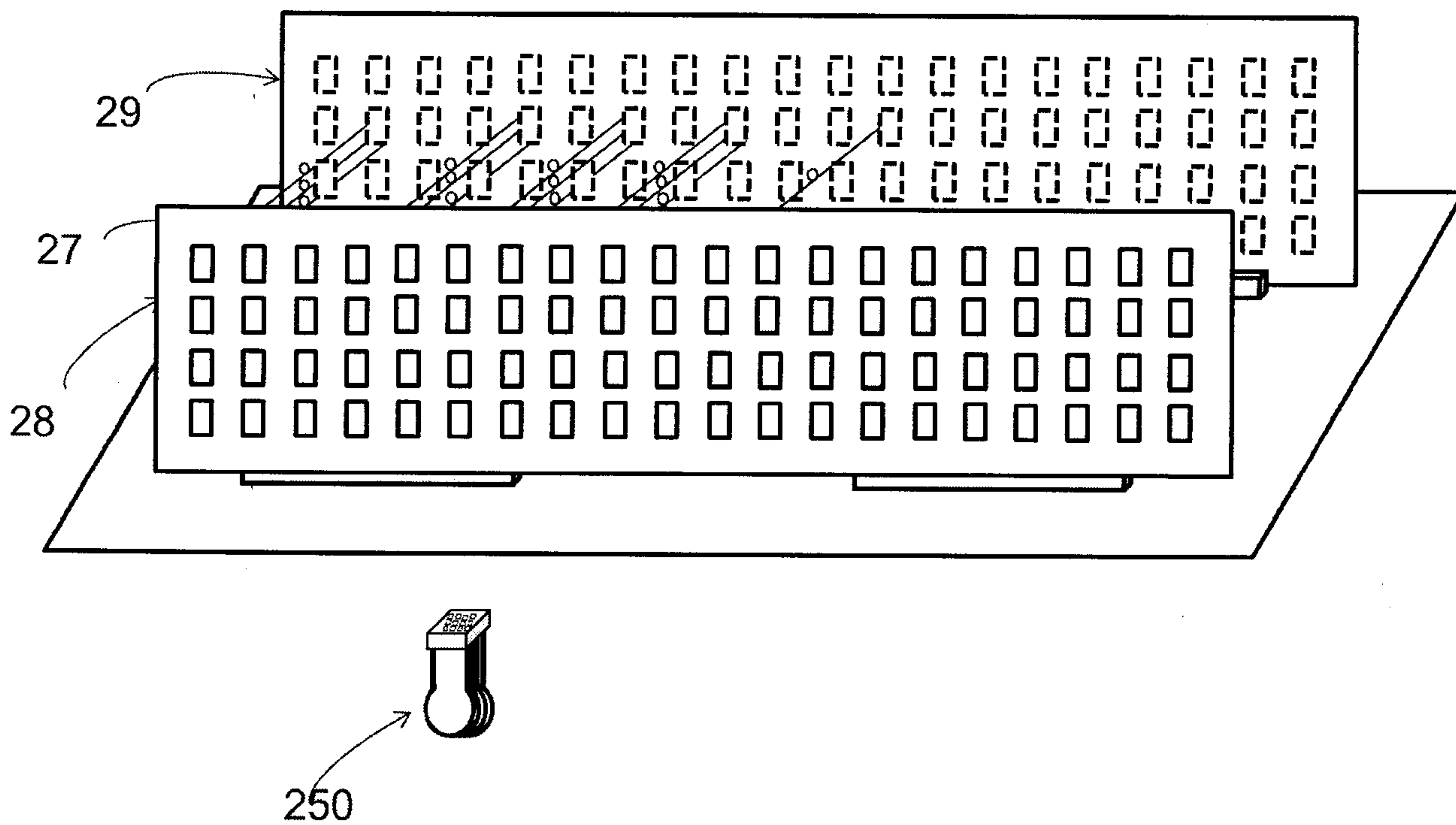


FIGURE 13

9/16

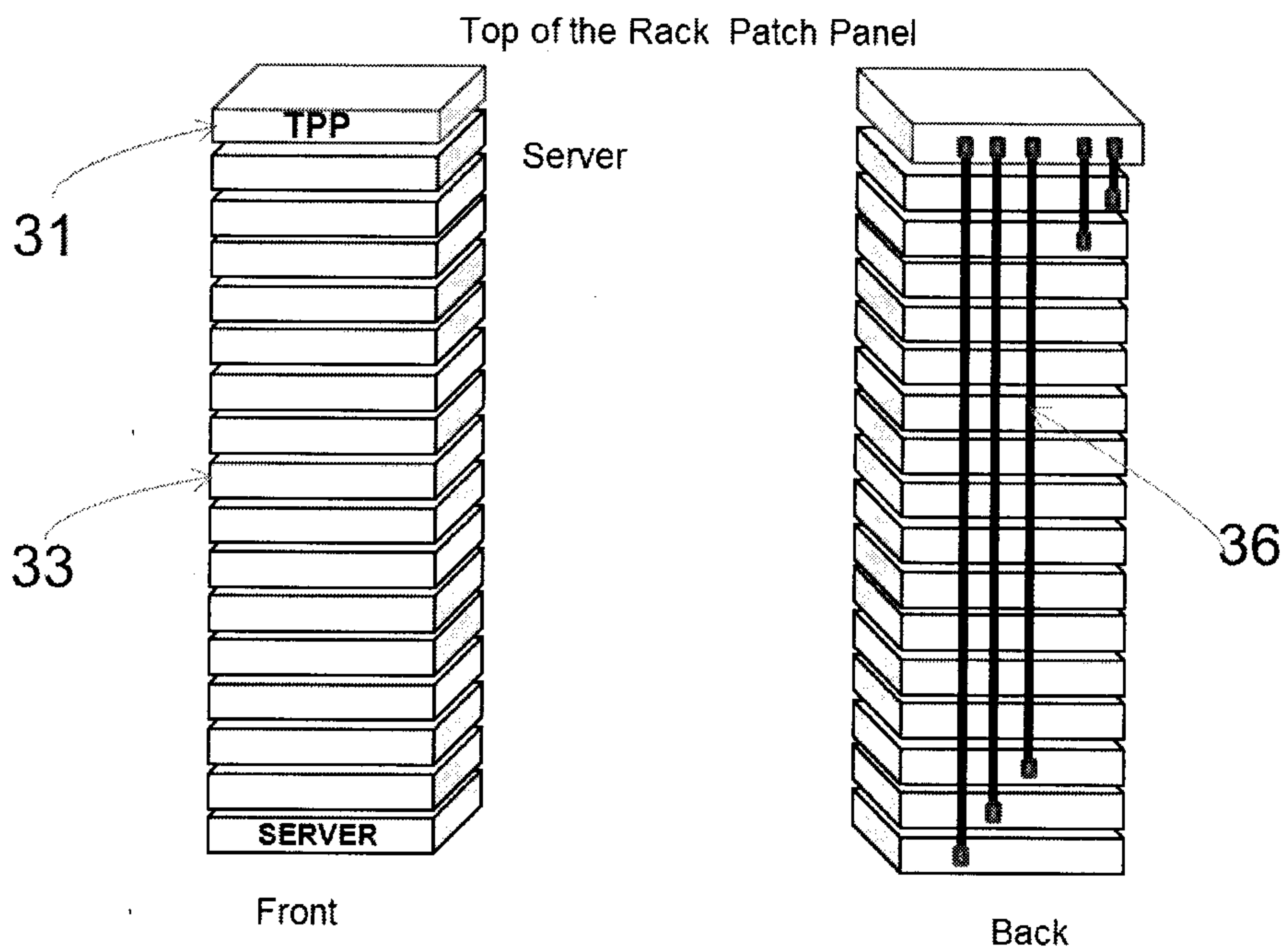


FIGURE 14

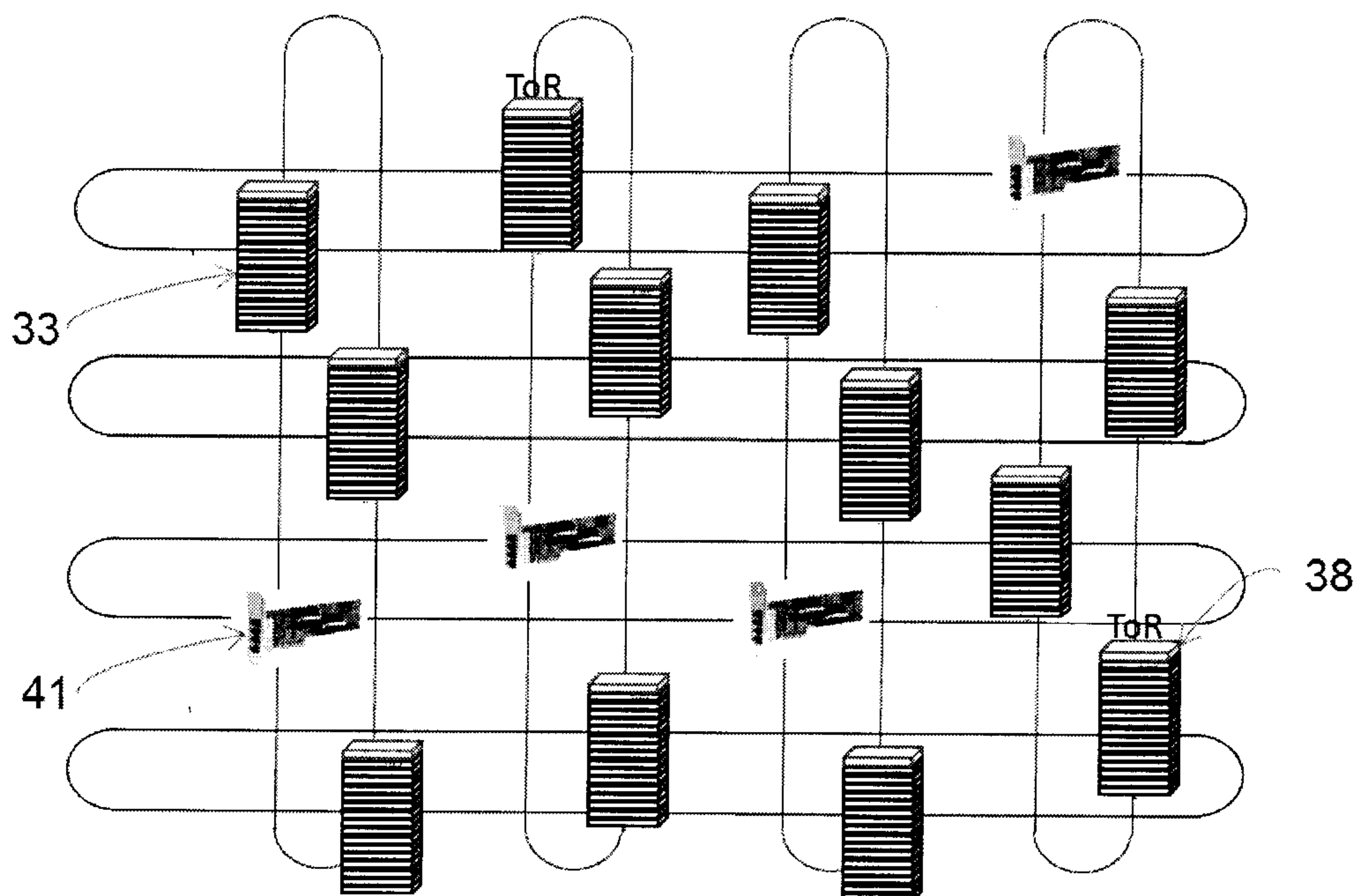


FIGURE 15

10/16

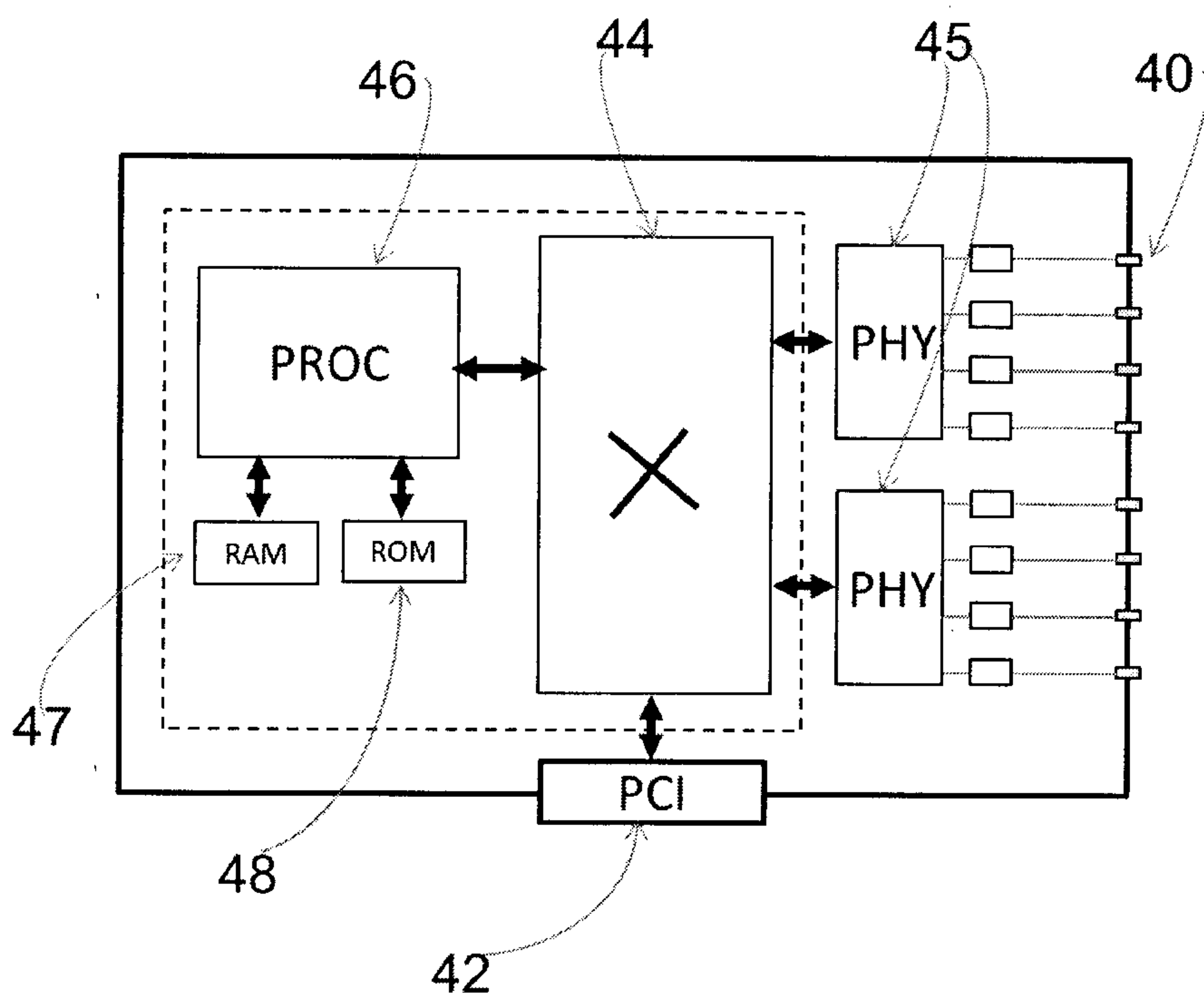


FIGURE 16

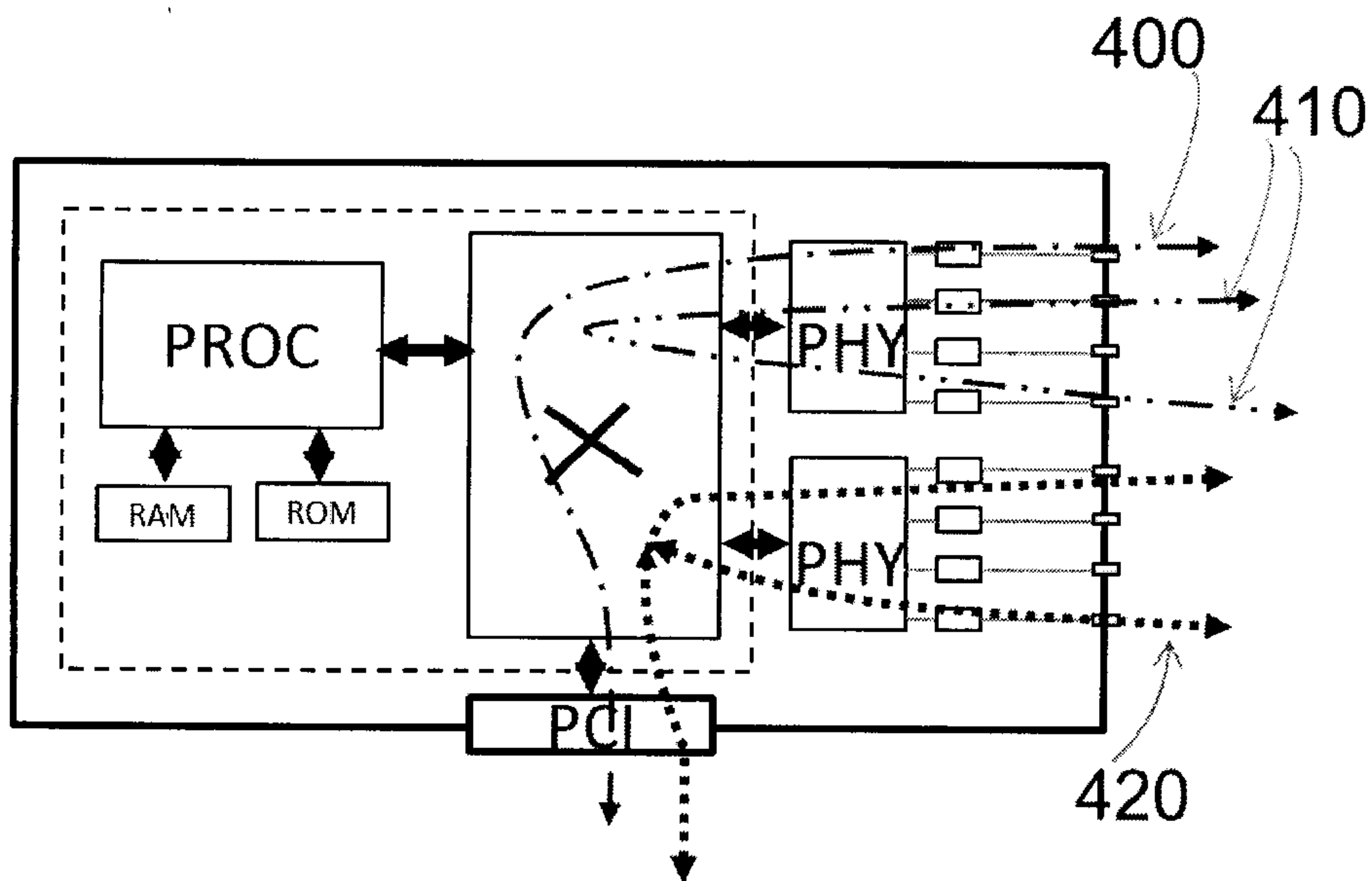


FIGURE 17

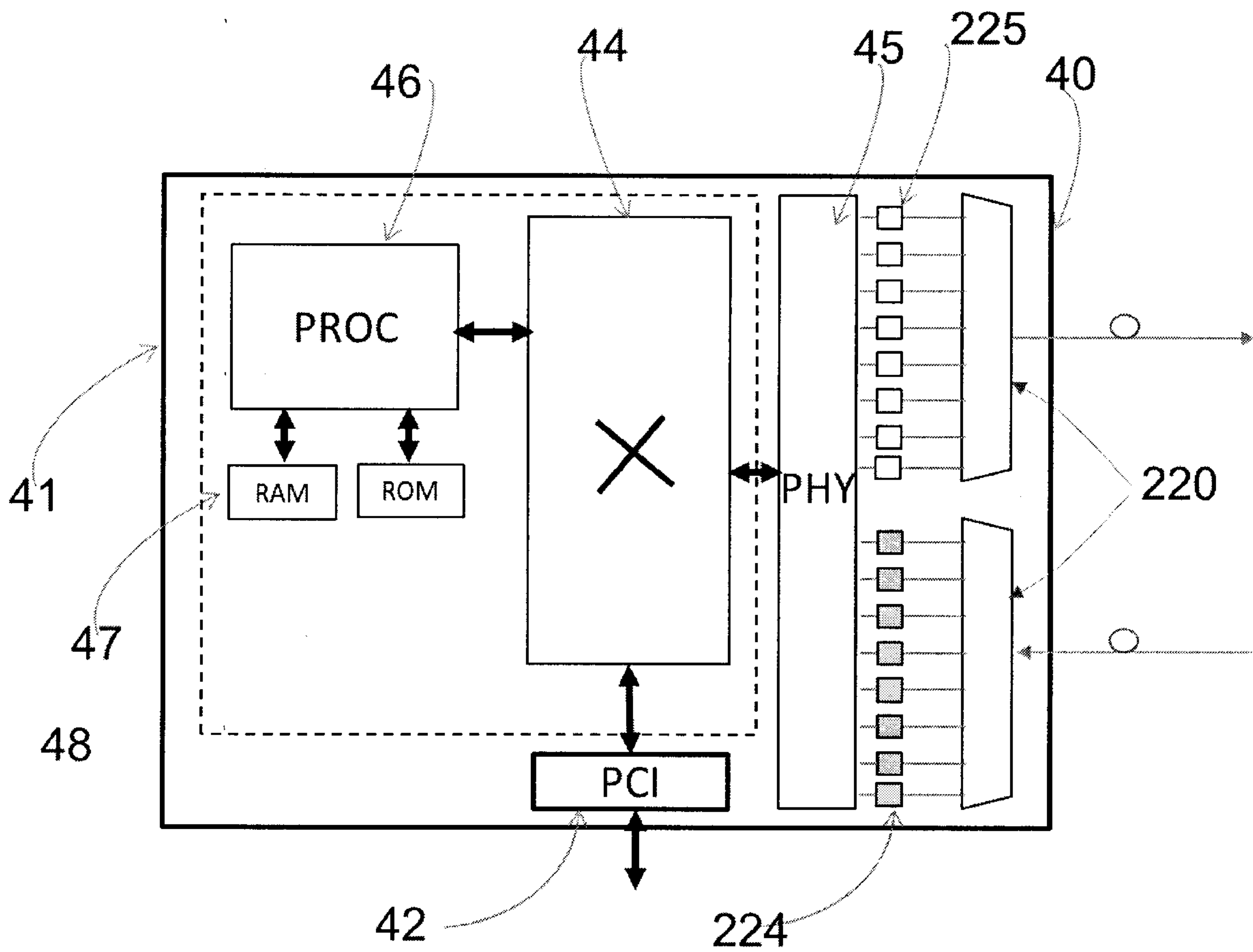


FIGURE 18

12/16

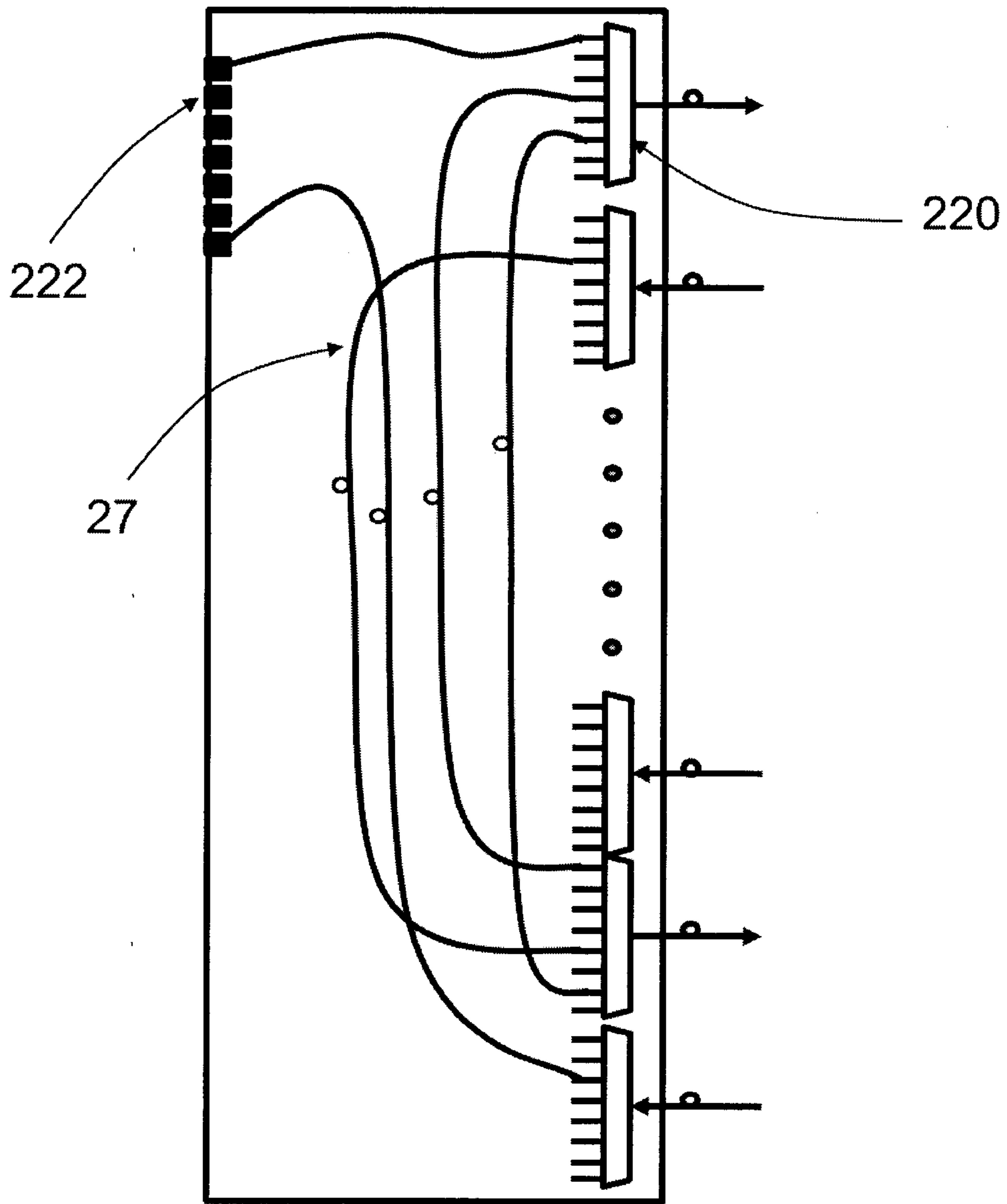


FIGURE 19

13/16

```

int main (int argc, char* argv[])
{ std::pair<int,int> rstart, pstart;
  pstart.first=0;
  static const int xDim=6;
  static const int yDim=7;
  static const int zDim=6;
  static const int aDim=7;

  std::cout << "// Patch: incrFolded - Xrings!!!\n";
  pstart.first=0;
  for (int a=0; a < aDim; a++) {
    pstart.second=0;
    for (int z=0; z < zDim; z++) {
      rstart.first=0;
      for (int y = 0; y < yDim; ++y) {
        mkRing(pstart, rstart, aDim, zDim, yDim, xDim, ringX);
        rstart.first++;}
      pstart.second++;}
    pstart.first++;}
  std::cout << "// Patch: incrFolded - Yrings!!!\n";
  pstart.first=0;
  for (int a=0; a < aDim; a++) {
    pstart.second=0;
    for (int z=0; z < zDim; z++) {
      rstart.second=0;
      for (int x = 0; x < xDim; ++x) {
        mkRing(pstart, rstart, aDim, zDim, yDim, xDim, ringY);
        rstart.second++;}
      pstart.second++;}
    pstart.first++;}
  std::cout << "// Patch: incrFolded - Zrings!!!\n";
  pstart.first=0;
  for (int a=0; a < aDim; a++) {
    rstart.first=0;
    for (int y=0; y < yDim; y++){
      rstart.second=0;
      for (int x = 0; x < xDim; ++x) {
        mkRing(pstart, rstart, aDim, zDim, yDim, xDim, ringZ);
        rstart.second++;}
      rstart.first++;}
    pstart.first++;}
  std::cout << "// Patch: incrFolded - Arings!!!\n";
  pstart.second=0;
  for (int z=0; z < zDim; z++) {
    rstart.first=0;
    for (int y=0; y < yDim; y++) {
      rstart.second=0;
      for (int x = 0; x < xDim; ++x) {
        mkRing(pstart, rstart, aDim, zDim, yDim, xDim, ringA);
        rstart.second++;}
      rstart.first++;}
    pstart.second++;}}

```

FIGURE 20a

14/16

```

enum Ringtype {ringA,ringZ,ringY,ringX};
std::pair<int,int> swapPair (std::pair<int,int>p)
{return std::pair<int,int> (p.second, p.first);}
std::pair<int,int> incrFolded(std::pair<int,int> cur, int rowDim, int colDim)
{std::pair <int, int> res(cur);
  if (colDim >1) {int nextCol;
    if (cur.second &1) {nextCol=cur.second -2;
      if (nextCol <0) nextCol=0;} else {nextCol = cur.second+2;
        if (nextCol >=colDim) nextCol = 2*colDim-1-nextCol;}
    res.second = nextCol;} else if (colDim > 0) {res.second += (cur.second &1)?-1:1;}
  return res;}

void mkRing (std::pair<int,int> & ppair, std::pair<int,int> & rpair, int dimA, int dimZ, int dimY, int
dimX, Ringtype rt)
{
  std::pair<int,int> pstart=ppair, rstart=rpair;
  switch (rt) {
  case ringX:
    {if (dimX > 1) {
      std::string sstart= "\", 0";
      std::string sres = "\", 1";
      rstart.second=0;
      for (int i=0; i< dimX; i++){
        std::pair<int,int> rres = incrFolded(rstart, dimY, dimX);
        std::pair<int,int> pres = pstart;
        std::cout << " { \" << pstart.first << " << pstart.second << " << rstart.first << " <<
<<rstart.second << sstart << ", \" << pres.first << " << pres.second << " << rres.first << " <<
rres.second << sres << "},\n";
        rstart=rres;}}}
    break;
  case ringY:
    {if (dimY > 1) {
      std::string sstart= "\", 2";
      std::string sres = "\", 3";
      rstart.first=0;
      for (int i=0; i< dimY; i++){
        std::pair<int,int>rres=

```

FIGURE 20b

15/16

```

swapPair(incrFolded(swapPair(rstart), dimX, dimY));
    std::pair<int,int> pres = pstart;
    std::cout << " {\\" << pstart.first << "\\" << pstart.second << "\\" << rstart.first
<< "\\" << rstart.second << sstart << "\\" << pres.first << "\\" << pres.second << "\\" <<
rres.first << "\\" << rres.second << sres << "\n";
    rstart=rres;}}
break;
case ringZ:
    {if (dimZ > 1) {
        std::string sstart= "\\", 4";
        std::string sres = "\\", 5";
        pstart.second=0;
        for (int i=0; i< dimZ; i++){
            std::pair<int,int> rres = rstart;
            std::pair<int,int> pres = incrFolded(pstart, dimA, dimZ);
            std::cout << " {\\" << pstart.first << "\\" << pstart.second << "\\" << rstart.first
<< "\\" << rstart.second << sstart << "\\" << pres.first << "\\" << pres.second << "\\" <<
rres.first << "\\" << rres.second << sres << "\n";
            rstart=rres;
            pstart=pres;}}
        break;
    case ringA:
        {if (dimA > 1) {
            std::string sstart= "\\", 6";
            std::string sres = "\\", 7";
            pstart.first=0;
            for (int i=0; i< dimA; i++){
                std::pair<int,int> rres = rstart;
                std::pair<int,int> pres = swapPair(incrFolded(swapPair(pstart), dimZ, dimA));
                std::cout << " {\\" << pstart.first << "\\" << pstart.second << "\\" << rstart.first
<< "\\" << rstart.second << sstart << "\\" << pres.first << "\\" << pres.second << "\\" <<
rres.first << "\\" << rres.second << sres << "\n";
                rstart=rres;
                pstart=pres;}}
            break;}}
    break;}}

```

FIGURE 20c

16/16

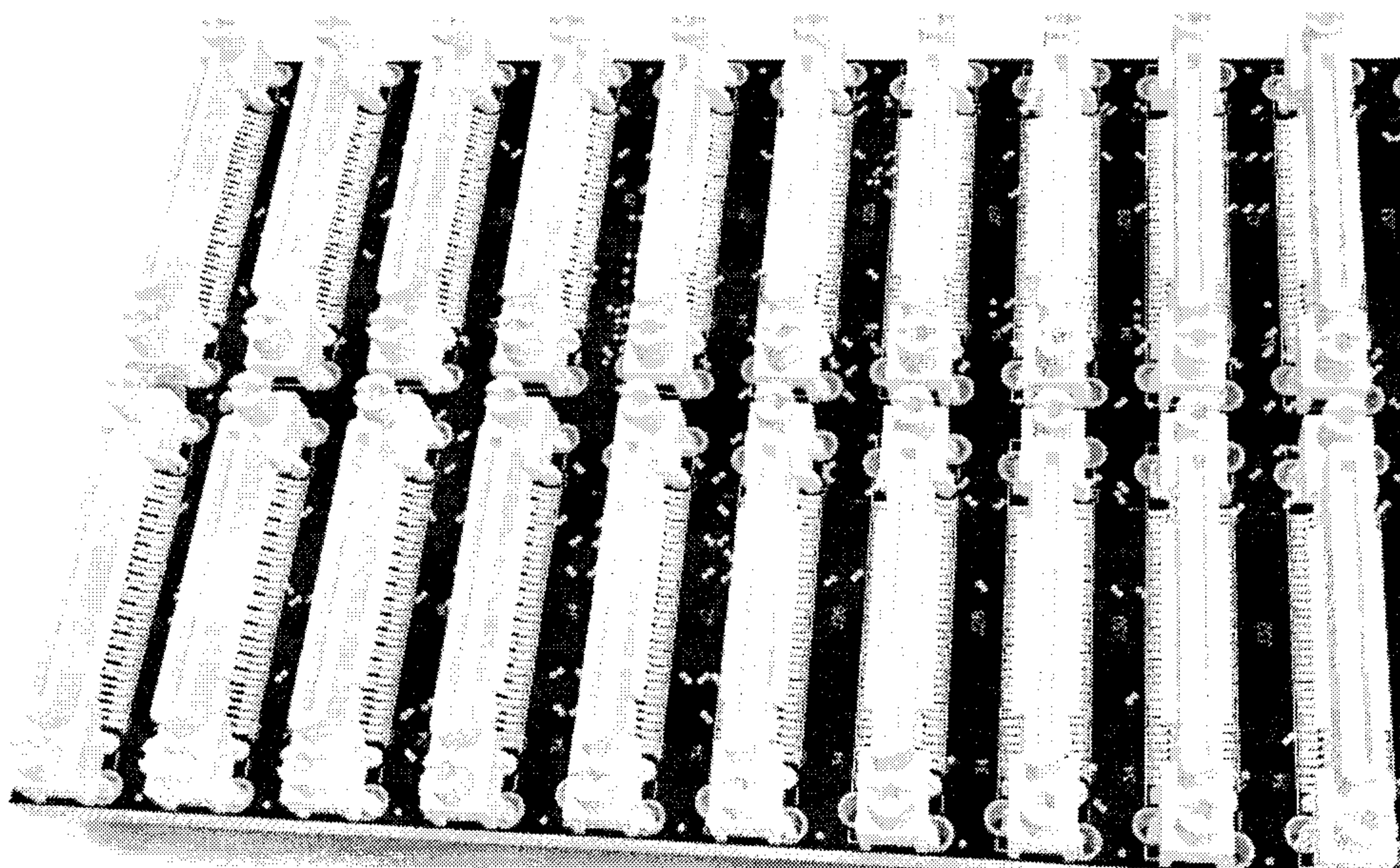


FIGURE 21

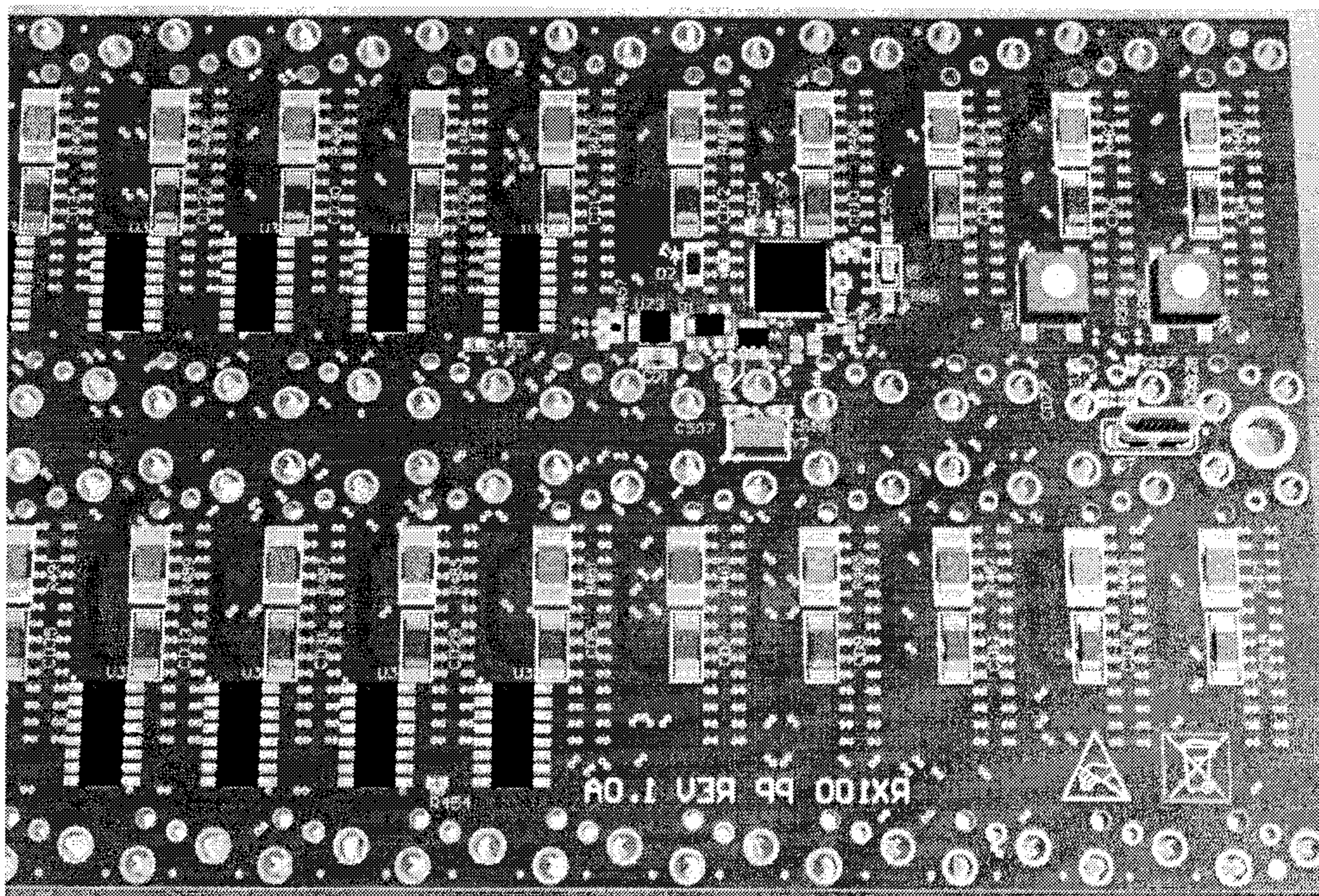


FIGURE 22

