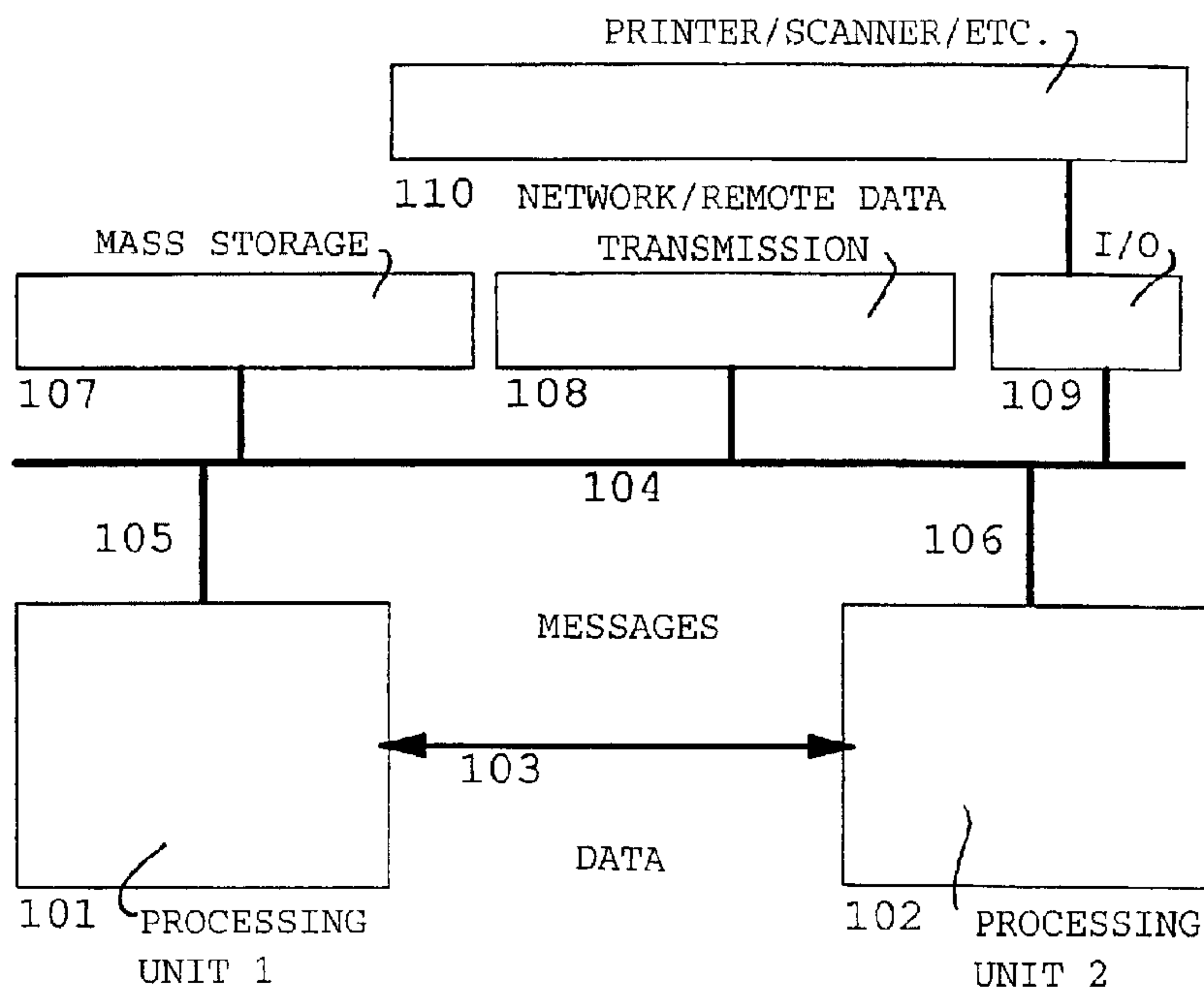




(86) Date de dépôt PCT/PCT Filing Date: 1994/07/22
 (87) Date publication PCT/PCT Publication Date: 1995/02/16
 (45) Date de délivrance/Issue Date: 2005/05/31
 (85) Entrée phase nationale/National Entry: 1996/02/02
 (86) N° demande PCT/PCT Application No.: DE 1994/000848
 (87) N° publication PCT/PCT Publication No.: 1995/004973
 (30) Priorité/Priority: 1993/08/09 (P 43 26 740.8) DE

(51) Cl.Int.⁶/Int.Cl.⁶ G06F 15/80, G06F 13/42
 (72) Inventeur/Inventor:
KOPP, MARTIN, DE
 (73) Propriétaire/Owner:
KOPP, MARTIN, DE
 (74) Agent: ROBIC

(54) Titre : ARCHITECTURE POUR SYSTEME D'ORDINATEUR MODULAIRE DANS LEQUEL UN DES MODULES EST DEDIE A UNE TACHE USAGER-INTERFACE
 (54) Title: ARCHITECTURE FOR MODULAR COMPUTER SYSTEM IN WHICH ONE OF THE MODULES IS DEDICATED TO USER-INTERFACE TASK



(57) **Abrégé/Abstract:**

Very large amounts of data occur in the area of the user interface of computers. The main processor is thus permanently overloaded, so that the computing power available for the actual application is sharply reduced in a manner that cannot be predicted without limiting the user requests. The new computer architecture should allow a higher and more exactly predictable computing power to be achieved. In order to make the rest of the system independent from the user interface, the computer is subdivided in such a way that part 1 takes over the display and operation tasks of the user interface, part 2 the processing of the application programmes without their user surface and part 3 the remaining functional units. All parts have their own internal data paths. An interface between parts 1 and 2 supports the exchange of messages and data without affecting the independent process in the parts. Part 1 and 2 have their own access paths to common devices located in part 3.

PCT

WELTORGANISATION FÜR GEISTIGES EIGENTUM
Internationales Büro



INTERNATIONALE ANMELDUNG VERÖFFENTLICHT NACH DEM VERTRAG ÜBER DIE
INTERNATIONALE ZUSAMMENARBEIT AUF DEM GEBIET DES PATENTWESENS (PCT)

(51) Internationale Patentklassifikation ⁶ : G06F 15/80, 13/42 2168741	A1	(11) Internationale Veröffentlichungsnummer: WO 95/04973 (43) Internationales Veröffentlichungsdatum: 16. Februar 1995 (16.02.95)
---	----	---

(21) Internationales Aktenzeichen: PCT/DE94/00848 (22) Internationales Anmeldedatum: 22. Juli 1994 (22.07.94) (30) Prioritätsdaten: P 43 26 740.8 9. August 1993 (09.08.93) DE (71)(72) Anmelder und Erfinder: KOPP, Martin [DE/DE]; Frankenstrasse 34, D-68259 Mannheim (DE).	(81) Bestimmungsstaaten: AM, AT, AU, BB, BG, BR, BY, CA, CH, CN, CZ, DE, DK, ES, FI, GB, GE, HU, JP, KE, KG, KP, KR, KZ, LK, LT, LU, LV, MD, MG, MN, MW, NL, NO, NZ, PL, PT, RO, RU, SD, SE, SI, SK, TJ, TT, UA, US, UZ, VN, europäisches Patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI Patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG), ARIPO Patent (KE, MW, SD).
---	---

Veröffentlicht

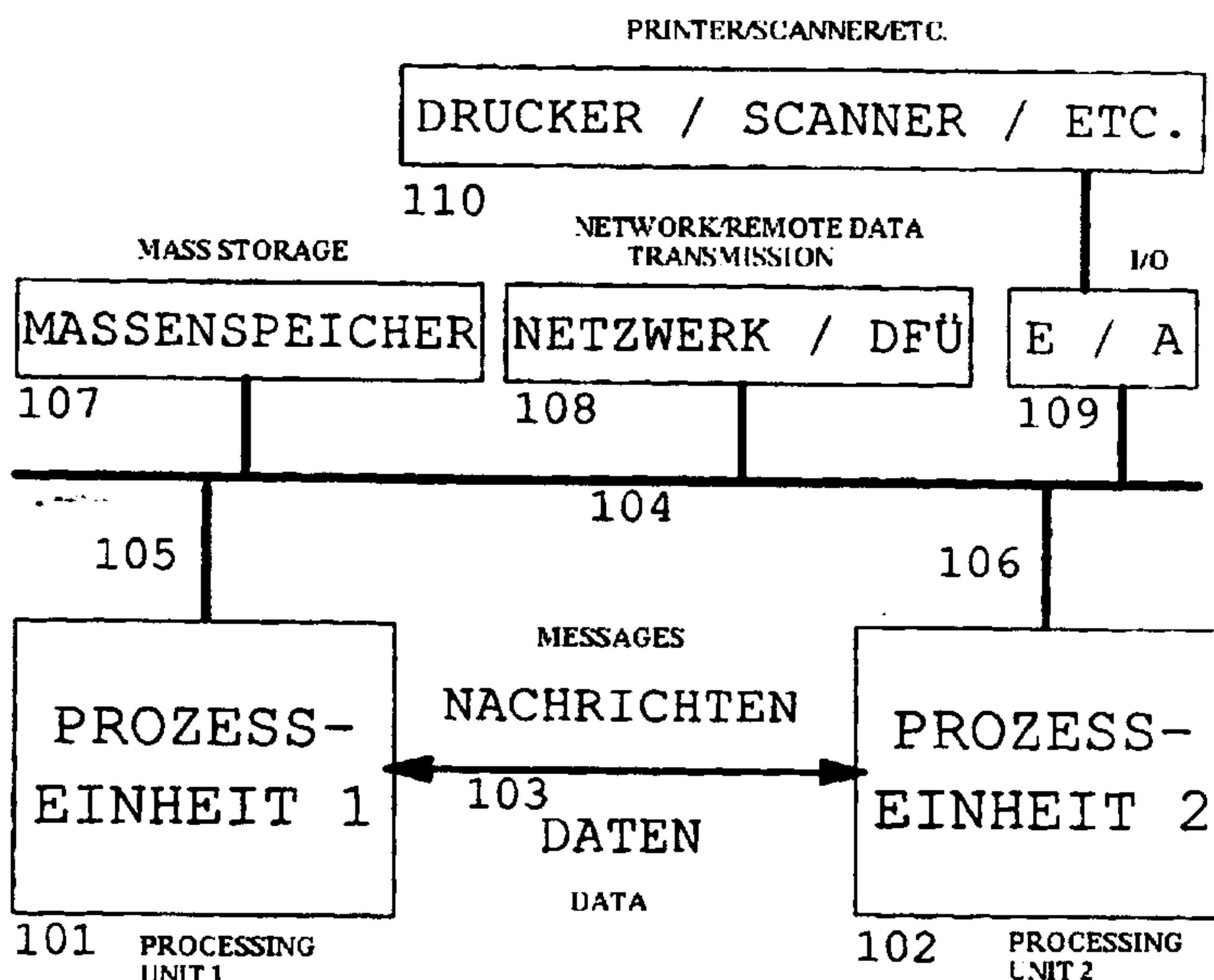
*Mit internationalem Recherchenbericht.
Vor Ablauf der für Änderungen der Ansprüche zugelassenen
Frist. Veröffentlichung wird wiederholt falls Änderungen
eintreffen.*

(54) Title: COMPUTER ARCHITECTURE

(54) Bezeichnung: ARCHITEKTUR FÜR EINE RECHENANLAGE

(57) Abstract

Very large amounts of data occur in the area of the user interface of computers. The main processor is thus permanently overloaded, so that the computing power available for the actual application is sharply reduced in a manner that cannot be predicted without limiting the user requests. The new computer architecture should allow a higher and more exactly predictable computing power to be achieved. In order to make the rest of the system independent from the user interface, the computer is subdivided in such a way that part 1 takes over the display and operation tasks of the user interface, part 2 the processing of the application programmes without their user surface and part 3 the remaining functional units. All parts have their own internal data paths. An interface between parts 1 and 2 supports the exchange of messages and data without affecting the independent process in the parts. Part 1 and 2 have their own access paths to common devices located in part 3.



- 1 -

**ARCHITECTURE FOR MODULAR COMPUTER SYSTEM IN WHICH ONE OF
THE MODULES IS DEDICATED TO USER-INTERFACE TASK**

10 Large amounts of data are handled in the area of the user interface of computers. They take up a considerable quantity of the computer's output, because on the one hand the resolution of the imaging systems being used is rising constantly and new applications such as animation and desktop video require high data rates for structured or unstructured graphics, and on the other hand, with the constantly growing complexity of the operating structures, shorter and shorter response times are being demanded from inter-active systems.

20 It is an established fact that not only the main processor in the computer but also, if necessary, additional passive or active components are used in order to generate the presentations made accessible to a human being by the imaging system (see: Foley, van Dam, Feiner, Hughes, "Computer Graphics, Principles, and Practice", 2nd edition, Addison-Wesley Publishing Company, Reading, Massachusetts, ISBN 0-201-12110-7, and in the following especially EP 0350911 A2 for a service processor). It is known from Document D1 (EP-A-0 528396, Page 3, Lines 2 to 9) that the tasks of administering the user interface in a network of an industrial control system can be assigned to a special program (a "user interface server program") which can carry out the inputs and outputs at each node (client terminal). A computer architecture is presented in Document D2 (US-A_4 570 220) with a number of bus systems and special equipment for exchanging data and messages.

30

Present-day systems do not take account in their structures of the extreme demands that have to be

met in order to achieve the operational functionality of the user interface, nor the huge quantities of data that have to be imported from or exported to a point outside the graphics sub-system
5 and the main processor in a very short time, e.g. from a mass storage device or an input/output interface. For instance, the image memory of a system that strives to attain a resolution of 1600 x 1200 pixels with a colour depth of 32 bits per
10 pixel will take up about 7.5 MB. For an animation

15

20

25

30

35

40

- 2 -

sequence running at 25 images per second, some 187 MB/second of new data will have to be imported into this memory to recreate the image displayed. If an image repeat rate of 100 Hz is the aim, however, 750 MB/s have to be read from this memory and displayed simultaneously.

This leads to severe overloading of the part of the computer installation that is supposed to be working on the actual application, meaning calculating and retaining the data and suchlike. The performance of this part of the computer is severely affected both by the heavy load placed on its data paths for transporting the data to and from the graphics sub-system, which can cause a blockage of the main processor or other subordinate units (see also EP 0350911 A2 as an example of a multi-processor system in which the shared main memory of this close-coupled multi-processor system is described as the principle 'bottleneck'), and by the frequent interruptions suffered by the working procedure for the purpose of reacting instantly to an input from the user or for maintaining the actuality of the presentation. Moreover, with these kinds of architecture it is not possible to forecast their performance without limiting the availability of the system for user queries (see EP 0350911 A2, if the system support processor is occupied with handling interrupts, semaphores, etc.).

The present invention is thus based on the task of reducing the load on the parts of the computer handling the actual application caused by implementing the user interface for presentation and function, in order thereby to achieve more efficient computers with which, even if their familiar components are retained and even despite ever more stringent demands on the quality of the user interface, it will be possible greatly to increase

- 3 -

the capacity available for the application and to achieve a good level of predictability.

According to the invention there is provided a computer system with user interface containing a number of processing units using separate internal bus systems, and equipment for transferring data and messages, and software for the system wide administration of the graphical user interface, characterised by:

- 10 a) the hardware and software components of the computer system being divided into three parts in such a way that one of them, Part 1, takes over the function of providing the user interface in presentation and function; another, Part 2, mainly operates the application programs and thus does not contain any of the operating programs and computer components needed for the user interface; and the third, Part 3, takes charge of the remaining components of the computer system and of a connecting structure suitable for providing access to these components; Parts 1, 2 and 3 being implemented as a single unit of equipment (apart from the external peripheral components) which could itself be modular in design,
- b) Parts 1, 2 and 3 possessing their own data paths for their internal tasks, separate from those of the other parts,
- 20 c) a software and hardware interface existing in the unit of equipment between Part 1 and Part 2, which supports the exchange of messages and data without creating a direct, permanent connection which affects the independent running of the processes in Parts 1 and 2,
- d) Part 1 and Part 2 each having their own, direct connection to jointly used additional components in Part 3.

The following provides a non-restrictive outline of certain possibly preferable features of the invention which will be more fully described hereinafter.

- 4 -

The advantages attainable with this invention consist mainly of the effects described in the following points:

The separation of Part 1, which handles the user interface, from Part 2, which handles the actual application (for instance: making calculations, administering data, or making logical combinations), leads to the clear uncoupling of these areas and thus to a far more efficient computer. For instance, a change of context by the processor, as is customary with present-day architecture, can in many cases be avoided completely.

10 Parts 1 and 2 can be optimized much more effectively for their actual tasks than is possible with any system in which both tasks are handled by one and the same part, for instance a single processing unit consisting of a processor, a memory, and an input/output.

20 Part 1's and Part 2's own access paths to the shared components in Part 3 increase the potential performance of the system, because for instance while Part 1 is making use of its access to a component in Part 3 (e.g. in order to load animation data from a mass storage device or a network), Part 2 can continue to work without interruption on its local data paths. In the systems customary at present, this kind of access either has to be handled by the central processor itself, for instance by reading data from the mass storage device and passing it on to the graphics card, or at least a bus system in the central processor has to be used which then more or less blocks it.

If the invention is implemented with suitable synchronisation equipment for access by Parts 1 and 2 to the jointly used components in Part 3, it can be guaranteed that even if maximum throughput is being fully utilized the part that has the highest priority at any given moment will have access with only a minimum of delay. Any necessary synchronisation can, for instance, be provided via the software and hardware interface between Parts 1 and 2 or

- 5 -

by means of a connection between the components in Parts 1 and 2 responsible for access.

10 The structural unification of the part of the computer responsible for the presentation and that responsible for the operational functionality of the user interface permits extensive optimization such as would not be possible if these components were separated. In the architectures customary today, although the presentation is admittedly accelerated by it's own processors, handling the response to the user (meaning the functionality of the user interface) still always has to be the responsibility of the central processor, which therefore has to interrupt whatever calculations are being carried out thus has to stop them for the necessary length of time. For instance, the object oriented structure of modern user interfaces can be better represented with the new architecture on the computer's architecture because any object of the user interface will generally consist of presentation and function, which are handled here by one and the same unit.

20 Data retention becomes easier because it is possible to do, at precisely that place where the graphic presentation is being made, the correlation of a co-ordination point on the screen surface (meaning, for instance, the position of a pointer) to the individual objects or structures of the user interface. For instance, the calculation of the mutual covers of such objects now only has to be done once.

The structure that is defined by the architecture in the computer is similar to the client / server software structure of modern operating system / user interface combinations. A division with message exchange and data transfer is provided here as well. This will facilitate a particularly simple adaptation of such systems to the new architecture. Older systems can likewise be transformed by means of appropriate measures, for instance by the provision of a multiple-access or dual ported memory.

- 6 -

Unlike genuine client / server structures, as are to be found for instance in so-called X-terminals, all the advantages are retained in this architecture of decentralized hardware with its own computing capacity for applications. Moreover, and in contrast to the possibilities described in Patent EP 401803 and elsewhere (the combination of several separate computer units), it is also possible to make use of techniques which are only feasible when the connections are kept short, meaning when they are all inside the one computer.

10 A distinguishing feature of the architecture is a clear modularisation, even if in actual practice it may not be visible in a physical or spatial form, and this makes it simple to exchange parts in order to modify or increase the output or to use various different processors or components in the three parts.

It is easy to see that the advantages of modularisation listed here and the similarity of this architecture to modern operating systems makes it possible to use this architecture independently of the design of the part-components, and that this makes it suitable for use with all known systems with a number of different processors and different operating system / user interface combinations.

20 The following pages will now describe three possible examples of the implementation of the invention in practice, as shown in the drawings:

Fig. 1 Architecture for a computer, schematic examples.

Fig. 2 Possible designs for Parts 1 and 2.

Fig. 3 Possible design for Part 1.

The abbreviations used in the drawings are explained in the text.

- 7 -

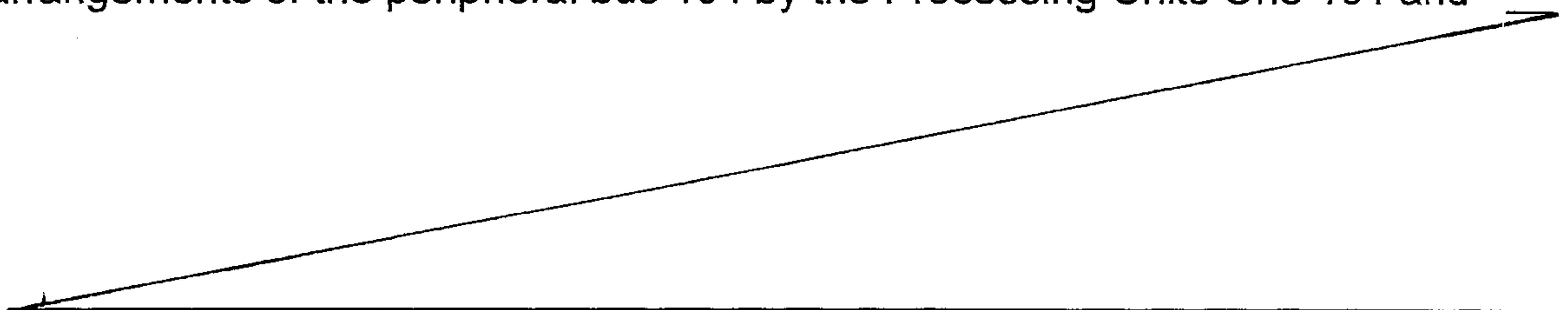
There now follows an explanation of the invention on the basis of the drawings of three implementation examples in terms of the construction and, where appropriate, of the way in which this invention works.

The details of the possible implementations such as special processors, bus systems, etc., are intended to serve only to communicate a better understanding of the present invention. It will be clear to the expert reader that the invention can be put into practice even without these special details. Schematic presentations are used in the drawings in order not to burden the invention with unnecessary details.

10 Fig. 1 shows a schematic example of the implementation of the architecture. A Processing Unit One 101 representing Part 1, which for instance could consist of a microprocessor with memory and peripheral interfaces, has an interface 103 for messages and data connecting it to another part, Processing Unit Two 102 likewise representing Part 2. The separated internal bus systems of the two processing units are not illustrated.

Part 3 is shown in this example as the additional units 107 (mass storage), 108 (network or remote data transmission) and 109 (input / output control, I/O), the periphery bus connecting these components 104 and the external peripherals 110 (printer, scanner, etc.) connected to I/O 109.

20 Processing Unit One 101 is connected via its own access path 105 to the peripheral bus 104 in Part 3. Accordingly, Processing Unit Two 102 is connected with its own access path 106 to the peripheral bus 104 and thus to the jointly used components 107, 108, 109, and 110 in Part 3. The mutual terminal arrangements of the peripheral bus 104 by the Processing Units One 101 and



- 8 -

Two 102 can be made, for instance, by a collision-detection process, similar to that used in the Ethernet.

10 The type of access to the individual components 107, 108, 109, and 110 using the peripheral bus 104, and the selection of the components, are shown only by way of an example and do not represent any inherent characteristics of the present invention. For instance, it would be equally possible to use a crossbar switch, which would permit the simultaneous access of the processing units representing Parts 1 and 2 to the components of Part 3 so long as the component groups it controlled were disjunctive.

The design of Processing Units One 101 and Two 102, which represent Parts 1 and 2, is likewise only an example. Another, more advanced, design is shown in Fig. 2. The processing units in this drawing are shown as 201 (Processing Unit One, PE1), and 202 (Processing Unit Two, PE2), as they merely represent a further possible design.

20 PE1 201 is connected via an access path of its own 207, which in this case could for instance be implemented in the form of an SCSI or an IEEE bus, with a peripheral bus 206. PE2 202 could, but must not necessarily, be connected in the same fashion via 208 to 206. Let an SCSI or an IEEE bus be assumed as the connecting equipment 209 (periphery bus control unit One, PBK1) and 213 (periphery bus control unit Two, PBK2). The connection 203 for message interchange between PBK1 and PBK2 is optional, but could for instance be made through an additional connection between the host adapters or by special protocols on the peripheral bus 206 or via a roundabout route via the access paths 204 for data or 205 for messages.

- 9 -

PBK1 209 is furthermore connected to 210 (memory One, S1), which would be feasible via a direct memory access (DMA) to memory S1, and to 211 (extension unit 1, EE1). EE1 211 can, for instance, contain adapters for image output 217 (image output device BAG) or for the connection to input devices 218 (input device EG). Also, in addition to the connection with PBK1 209 already mentioned, it has one connection to S1 210 and another to processor 212 (processor One P1). The latter could be a microprocessor adapted for graphics. This has an additional internal connection inside PE1 201 with S1 210. This will make the internal, independent bus systems in PE1 201 representing Part 1 clear.

10

The internal construction of PE2 202 is similar in all relevant respects. PBK2 213 is connected to S2 214 (memory Two, S2) and EE2 215 (extension unit 2, EE2), which in this example could consist of a mathematical co-processor. S2 214 has connections to PBK2 213, EE2 215, and the processor P2 216, which in turn has contact with EE2 215. P2 216 could be any suitable microprocessor. It emerges at this point that the present invention can be readily adapted to fit in with existing architectures by letting the old architecture take the place of Part 2 (in this case PE2 202) which is augmented in a suitable manner, for instance using one of the methods described here, by a Part 1 (in this case PE1 201) and connected to it. In this way, it should be possible to use components from the existing architecture, e.g. a graphics card in PE1 201.

20

Connection 204, for handling data between S1 210 and S2 214, could take the form of DMA transfer that may be initiated by processors P1 212 and P2 216. Connection 205, for handling messages, could take the form of interrupt lines by means of which the processors would inform one another of events which could if necessary be accompanied by a transfer on 204. Alternatively, queues could be implemented in hardware (as FIFO memories, for instance).

If connections 203, 204, and 205 have been clearly defined, PE1 201 could also for instance be replaced if need be, event at a later date, by the variant PE1 301

- 10 -

shown in Fig. 3. The reverse also applies, of course, for any other fictitious implementation of PE2 202. This makes the possible modular concept visible.

PE1 301 as shown in Fig. 3 has the same basic structure as PE1 201. A correlation results, in respect of the basic function : P1 302 corresponds to P1 212, but in this case should instead be implemented as a universal processor because the graphics functions, as will be explained later, can be taken on by other components. It has a connection with S1 303, corresponding to S1 210, but can perhaps also have different key data. This in turn is connected to PBK1 304, which corresponds to PBK1 209. P1 302, S1 303, and PBK1 304 each has one or more than one connection to the remainder of PE1 301, which is the counterpart to EE1 211. Let the following components of EE1 211 be presented here in detail by way of examples: a digital signal processor (DSP) 308 connected to S1 303 and with interfaces to 312 (audio input device AEG) and 313 (audio output device AAG); an input processing unit EPE 307, which has an interface to one or more than one input devices EG 311; and a video processing unit (VPE) which on the one hand has an interface to 309 (image output device BAG) and another to 310 (image input device BEG) but on the other hand also has access to a scalable number of image processing units (BPE1, BPE2, BPE3, ...) 350, 351, 352 ... The BPE's could take the form of small modules composed of a simple graphics processor, possibly combined with a BITBLT function, and a part of the image memory. They generate in parallel the presentation which is read from VPE 306 and issued as a video signal, or alternatively arrives as a video signal at VPE 306, where it is converted and distributed to the BPE's. The BPE's are connected, individually or in groups, with compression/decompression units

10

20

30

- 11 -

(KDE1, KDE2, ...) 330, 331 ... which send data streams to and from PBK1 304 and the various different KDE's, converting them in whichever way is required.

PBK1 304 can of course, for instance, consist of a number of parallel access units in order to attain a greater band-width.

10 All these components are controlled by a configuration and message bus 305, through which the processor P1 302 co-ordinates everything. Interrupt signals and very small quantities of data, such as a single character of the keyboard, can also be transmitted via this bus.

The possible further access paths of their own which PE1 or PE2 might have to their own components in Part 3 have not been illustrated. These might for instance consist of a modem connected via an interface allocated exclusively to PE2, or something similar.

- 12 -

CLAIMS

1. Computer system with user interface containing a number of processing units (101; 102) using separate internal bus systems, and equipment for transferring data and messages (103), and software for the system wide administration of the graphical user interface

characterised by

10

a) the hardware and software components of the computer system being divided into three parts in such a way that one of them, Part 1 (101), takes over the function of providing the user interface in presentation and function; another, Part 2 (102), mainly operates the application programs and thus does not contain any of the operating programs and computer components needed for the user interface; and the third, Part 3, takes charge of the remaining components (107, 108, 109, 110) of the computer system and of a connecting structure (104) suitable for providing access to these components; Parts 1, 2 and 3 being implemented as a single unit of equipment (apart from the external peripheral components (110)) which could itself be modular in design,

20

b) Parts 1, 2, and 3 possessing their own data paths for their internal tasks, separate from those of the other parts,

c) a software and hardware interface (103) existing in the unit of equipment between Part 1 (101) and Part 2 (102), which supports the exchange of messages and data without creating a

30

- 13 -

direct, permanent connection which affects the independent running of the processes in Parts 1 and 2,

d) Part 1 and Part 2 each having their own, direct connection (105, 106) to jointly used additional components in Part 3.

10 2. The computer system according to claim 1 characterised by Part 1 (201) being implemented in the form of at least one processor of its own (212) with an accompanying memory (210) and access components (209) to the components in Part 3 used jointly with Part 2 (202).

3. The computer system according to claim 1 or 2 characterised by Part 2 (202) being implemented in the form of at least one processor of its own (216) with an accompanying memory (214) and access components (213) to the components in Part 3 used jointly with Part 1 (201).

4. The computer system according to any one of claims 1 to 3 characterised by the interface (103) between Part 1 and Part 2 supporting the efficient exchange of data (204) and messages (205) through the use of interrupt signals and queues.

20 5. The computer system according to any one of claims 1 to 4 characterised by the presence of additional access paths specifically for Part 1 and/or Part 2 to their own components in Part 3 allocated exclusively to them.

6. The computer system according to any one of claims 1 to 5 characterised by the presence of additional specific equipment (211) within Part 1 which supports the relevant functions, preferably those of accelerating the graphics assembly (350, 351, 352, ...) and/or supporting (306, 307, 308) various input/output devices (309...313) and/or local mass storage and/or the compression/decompression of data (330, 331...).

- 14 -

7. The computer system according to any one of claims 1 to 6 characterised by the presence of additional specific equipment (215) within Part 2 which supports the various functions, preferable those of accelerating mathematical functions and/or functions for retaining data, such as the phonetic search in complete texts, and/or local mass storage.

10 8. The computer system according to any one of claims 1 to 7 characterised by the presence of a connection (203) between access components (209, 213) in Parts 1 and 2 facilitating access to the jointly used components in Part 3 which preferably handles (meaning prevents or deliberately uses) a collision by priority control on the jointly used data paths.

9. The computer system according to claim 8 characterised by the connection (203) being implemented in the form of priority control which preferably uses dynamically changeable priorities, for instance enabling priorities to be changed in particular cases in favour of Part 1 or in favour of Part 2 with regard to the base configuration.

20 10. The computer system according to any one of claims 1 to 9 characterised by Parts 1, 2, and 3 being implemented as a modular design in which hardware and software interfaces (103, 105, 106) are defined in such a way that various different configurations and/or a retro-fitting facility can be supported to meet changing requirements.

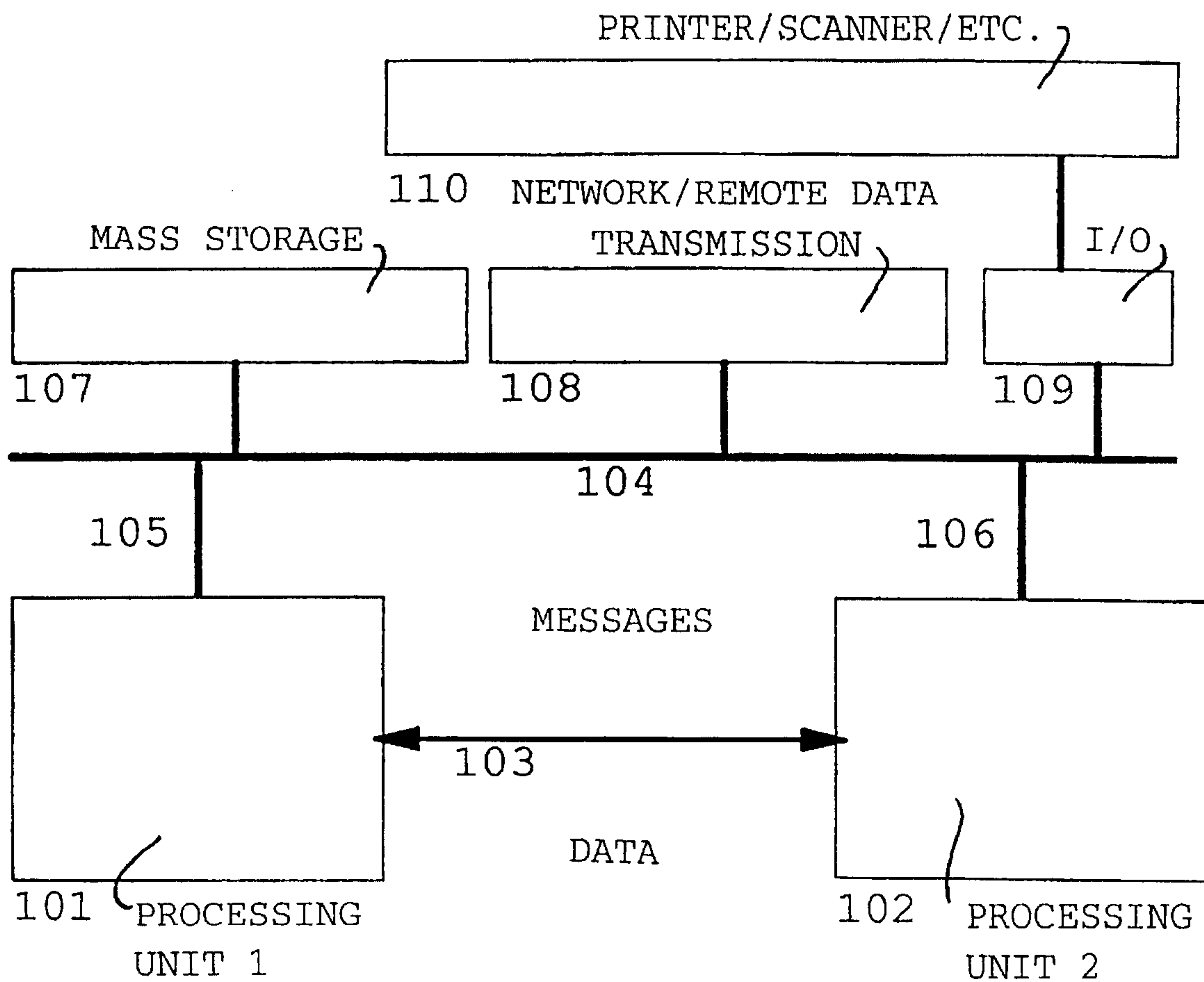


FIG. 1

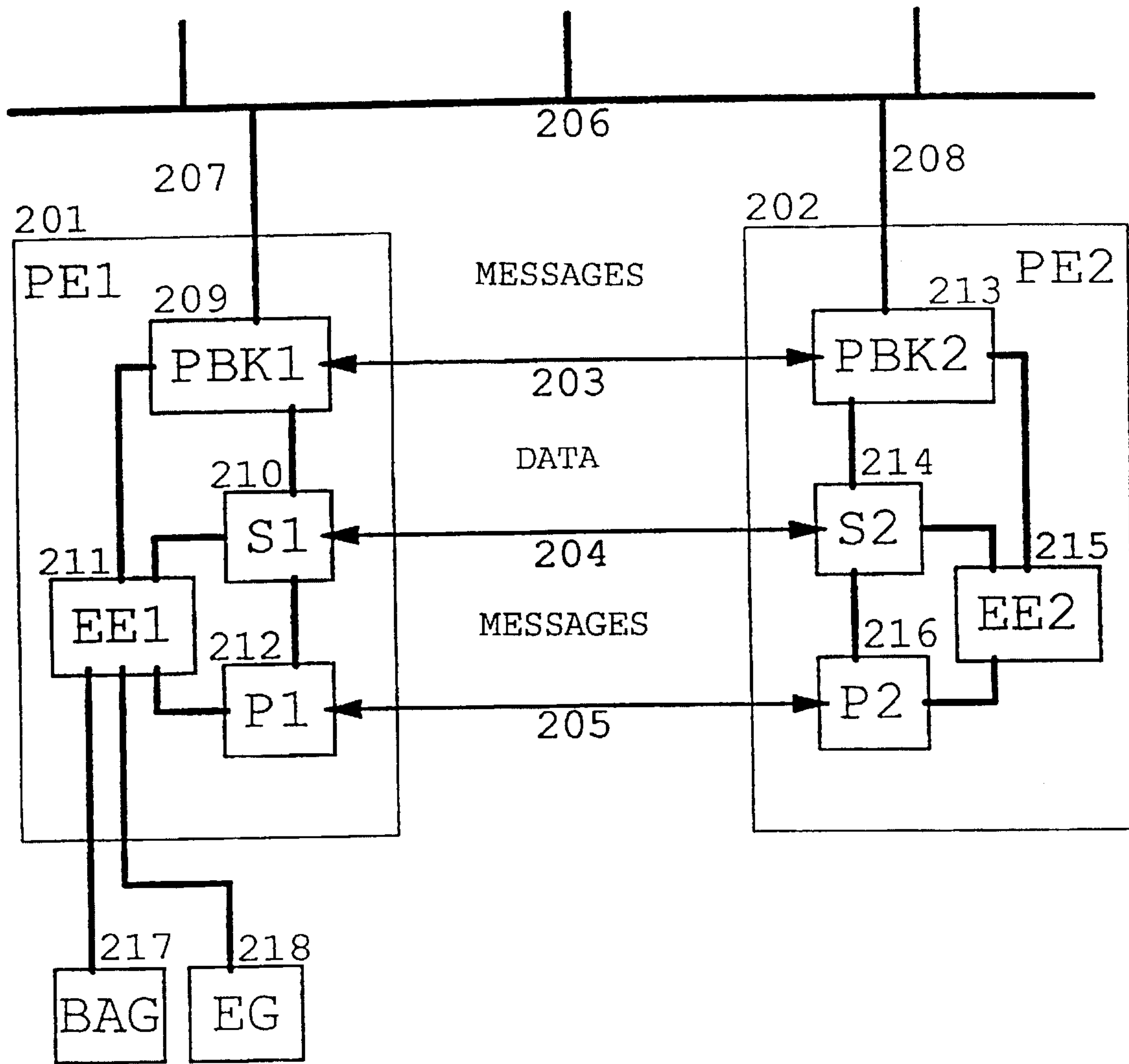


FIG. 2

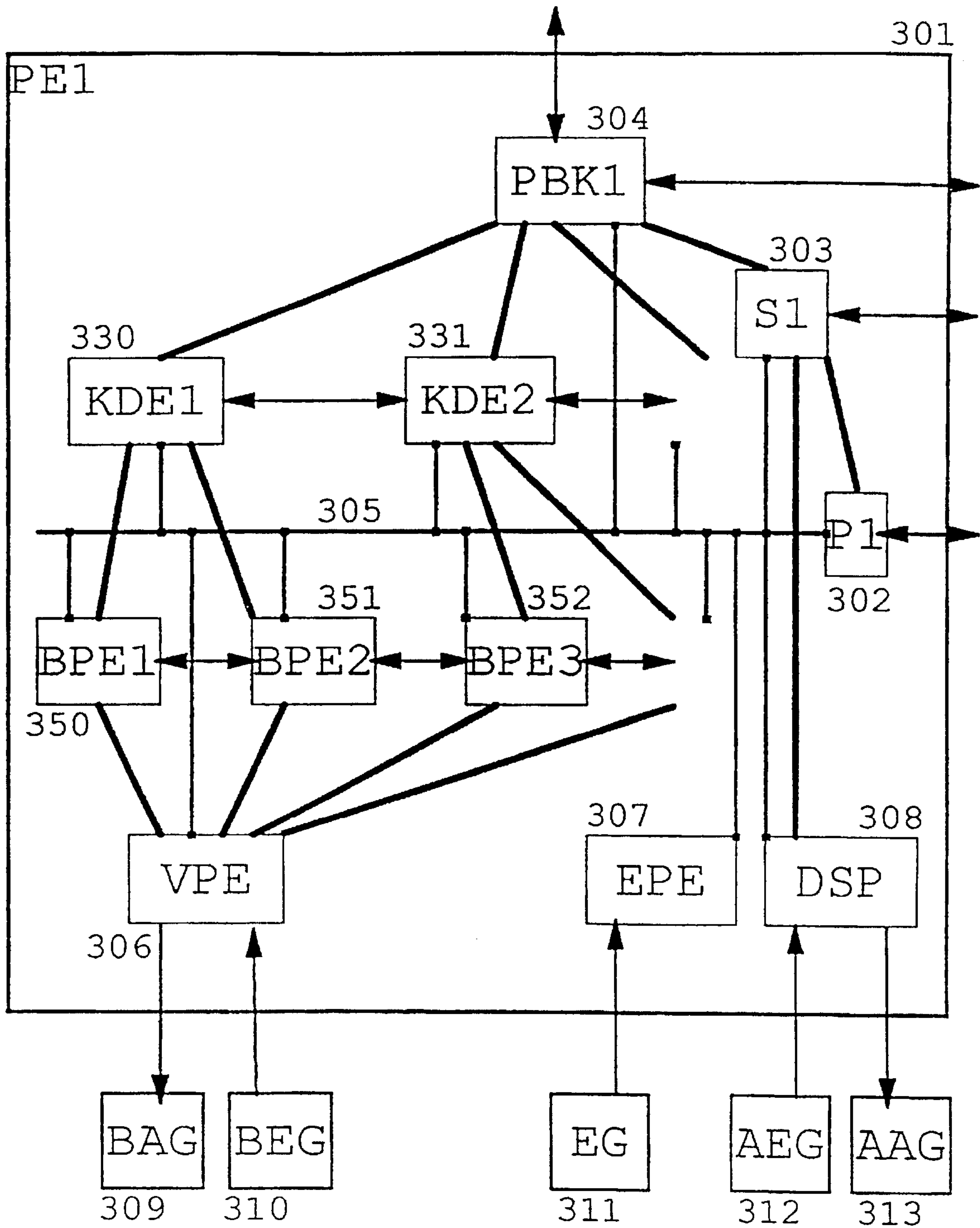


FIG. 3

