PATENT SPECIFICATION (11) 1588805

1588805

(21) Application No. 3058/80 (22) Filed 17 June 1977 (19)
(62) Divided out of No. 1588803
(31) Convention Application No. 721043 (32) Filed 7 Sept. 1976 in
(33) United States of America (US)
(44) Complete Specification published 29 April 1981
(51) INT. CL.³ G06F 11/20
(52) Index at acceptance
G4A 12T ES

(54) DISTRIBUTED POWER SYSTEM FOR A MULTIPROCESSOR SYSTEM

(71) We, TANDEM COMPUTERS INCORPORATED, a Corporation of the State of California, United States of America, of 20605 Valley Green Drive, Cupertino, California 95014, United States of America, do hereby declare the invention, for which we pray that a patent may be granted to us, and the method by which it is to be performed, to be particularly described in and by the following statement:—

The present invention relates to a distributed power system for a multiprocessor system of the kind in which separate processor modules are interconnected for parallel processing and relates particularly to a distributed power system for a multiprocessor system which can support high transaction rates to large on-line data bases and in which no single component failure can stop or contaminate the operation of the system.

There are many applications which require on-line processing of large volumes of data at high transaction rates. For example, such processing is required in retail applications for automated point of sale, inventory and credit transactions and in financial institutions for automated funds transfer and credit transactions.

In computing applications of this kind it is important, and often critical, that the data processing not be interrupted. A failure of an on-line computer system can shut down a portion of the related business and can cause considerable loss of data and money.

Thus, an on-line system of this kind must provide not only sufficient computing power to permit multiple computations to be done simultaneously, but it must also provide a mode of operation which permits data processing to be continued without interruption in the event some component of the system fails.

The system should operate either in a fail-safe mode (in which no loss of throughput occurs as a result of failure) or in a fail-soft mode (in which some slowdown occurs but full processing capabilities are maintained) in the event of a failure.

Furthermore, the system should also operated in a way such that a failure of a single component cannot contaminate the operation of the system. The system should provide fault-tolerant computing. For fault-tolerant computing all errors and failures in the system should either be corrected automatically, or if the failure or error cannot be corrected automatically, it should be detected, or if it cannot be detected, it should be contained and should not be permitted to contaminate the rest of the system.

Since a single processor module can fail, it is obvious that a system which will operate without interruption in an on-line application must have more than one processor module.

Systems which have more than one processor module can therefore meet one of the necessary conditions for non-interruptible operation. However, the use of more than one processor module in a system does not by itself provide all the sufficient conditions for maintaining the required processing capabilities in the event of component failure, as will become more apparent from the description to follow.

Computing systems for on-line, high volume, transaction oriented, computing applications which must operate without interrupton therefore require multiprocessors as a starting point. But the use of multiprocessors does not guarantee that all of the sufficient conditions will be met, and fulfilling the additional sufficient conditions for on-line systems of this kind has presented a number of problems in the prior art.

The prior art approach to uninterrupted data processing has proceeded generally along two lines — either adapting two or more large, monolithic, general purpose computers for joint operation or interconnecting a plurality of minicomputers to provide multiprocessing capabilities.

In the first case, adapting two large monolithic general purpose computers for joint operation, one conventional prior art approach has been to have the two computers share a common memory. Now in this type of multiprocessing system a failure in the shared memory can stop the entire system. Shared memory also presents a number of

SPECIFICATION AMENDED - SEE ATTACHED SLIP

other problems including sequencing accesses to the common memory. This system, while meeting some of the necessary conditions for uninterrupted processing, does not meet all of the sufficient conditions.

Furthermore, multiprocessing systems using large general purpose computers are quite expensive because each computer is constructed as a monolithic unit in which all components (including the packaging, the cooling system, etc.) must be duplicated each time another processor is added to the system even though many of the duplicated components are not required.

The other prior art approach of using a plurality of minicomputers has (in common with the approach of using large general purpose computers) suffered from the drawback of having to adapt a communications link between computers that were never originally constructed to provide such a link. The required links were, as a result, usually made through the input/output channel. Connections through the input/output channel are necessarily slower than internal transfers within the processor itself, and such interprocessor links have therefore provided relatively slow interprocessor communication.

Furthermore, the interprocessor connections required special adapter cards that added substantially to the cost of the overall system and that introduced the possibility of single component failures which could stop the system. Adding dual interprocessor links and adapter cards to avoid problems of critical single components failures increased the overall system cost even more substantially.

Providing dual links and adapter cards between all processors generally became very cumbersome and quite complex from the standpoint of operation.

Another problem of the prior art arose out of the way in which connections were made to peripheral devices.

If a number of peripheral devices are connected to a single input/output bus of one processor in a multiprocessor system and that processor fails, then the peripheral devices will be unavailable to the system even though the failed processor is linked through an interprocessor connection to another processor or processors in the system.

To avoid this problem, the prior art has provided an input/output bus switch for interconnecting input/output busses for continued access to peripheral devices when a processor associated with the peripheral devices on a particular input/output bus fails. The bus switches have been expensive and also have presented the possibility of single component failure which could down a substantial part of the overall system.

Providing software for the prior art multiprocessor systems has also been a major problem.

Operating systems software for such multiprocessing systems has tended to be nonexistent. Where software had been developed for such multiprocessor systems, it quite often was restricted to a small number of processors and was not adapted for the inclusion of additional processors. In many cases it was necessary either to modify the operating system or to put some of the operating system functions into the user's own program — an expensive, time-consuming operation.

The prior art lacked a satisfactory standard operating system for linking processors. It also did not provide an operating system for automatically accommodating additional processors in a multiprocessing system constructed to accommodate the modular addition of processors as increased computering power was required.

According to the present invention there is provided a distributed power system for multiprocessor system of the kind in which separate processor modules each have a central processing unit are interconnected for parallel processing, said distributed power system comprising: a plurality of separate processor modules; a plurality of device controllers for controlling the transfer of data between a plurality of processor modules and a plurality of peripheral devices; multiple ports in each device controller and multiple input/output buses for connecting each device controller for access by different processor modules; and, distributed power supply means for insuring non-stop operation of the remainder of the multiprocessor system in the event of a failure of a single power supply to a part of the multiprocessor system, said distributed supply means including a separate power supply for each processor module and connecting means for connecting a plurality of the separate power supplies to each device controller and effective to supply power to each device controller from the remaining separate power supplies in the event of the failure of one of the power supplies connected to any device controller.

Preferably in a multiprocessor system incorporating the present invention every major component is modularized so that any major component can be removed or replaced without stopping the system. In addition, the system can be expanded in place (either horizontally by the addition of standard processor modules or in most cases vertically by the addition of peripheral devices) without system interruption or modification to hardware or software.

Preferably the distributed power system includes power-on circuit means for establishing that power levels are within certain established limits and which are effective to

function in four states—power-off, power going from off to on, power-on and power going from on to off.

In one preferred embodiment of the present invention all the processor modules may be connected by an interprocessor bus and preferably a power-on circuit means is located in each central processing unit and in each device controller.

The power-on circuit means may control interface circuits which drive all the interprocessor and input-output buses in the multiprocessor system and may present a signal in the event of a variation of logic power which signal establishes the level of power applied to the related processor module or device controller. The signal output may be used to directly disable appropriate interface signals of the processor module or device controller in which the power-on circuit is located, if the power-on is not within the certain established limits which will ensure correct logic operation.

A second interprocessor bus may interconnect all of the processor modules and the power-on circuit means may, for example, coact with a multi-path system provided, in a preferred embodiment, by the interprocessor buses and the multiple port device controllers and associated input/output buses to give a mode of operation in which a failing power supply for one processor or a multiple port device controller does not affect any other processor module or dual port controller on any of those paths.

Desirably the connecting means includes a diode switch arrangement which supplies power to the device controller from a plurality of power supplies when those associated power supplies are operative and which supplies power from the remaining power supplies in the event of a failure of one of the associated power supplies and in a changeover which is without interruption or pulsation so that operation of the device controller is never interrupted.

Preferably the separate power supplies are interconnected by a power bus having a plurality of lines with selected ones of the lines connected to selected ones of the power supplies and including taps on the power bus adjacent each device controller for permitting the device controller to be connected to any selected plurality of the power supplies by easy connection to related taps in the power.

In a preferred embodiment of the present invention any processor module or device controller can be powered down so that on-line maintenance can be performed in a powered off condition while the rest of the multiprocessor system incorporating the present invention is on-line and functional.

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

FIGURE 1 is an isometric, block diagram view of a multiprocessor system incorporating one embodiment of the present invention. Figure 1 shows several processor modules connected by two interprocessor buses (an X and a Y bus) with each bus controlled by a bus controller. Figure 1 also shows several dual-port device controllers with each device controller connected to the input/output (I/O) buses of two processor modules.

FIGURE 2 is a schematic view showing the details of the power on circuit (PON).

FIGURE 3 is a block diagram of a power distribution system. Figure 3 shows how a plurality of independent and separate power supplies are distributed and associated with the dual port device controllers for ensuring that each device controller has both a primary and an alternate power supply.

FIGURE 4 is an enlarged, detailed view of the switching arrangement for switching between a primary power supply and an alternative supply for a device controller. The switching structure shown in Figure 4 permits both automatic switching in the event of a failure of the primary power supply and manual switching in three different modes—off, auto and alternate.

FIGURE 5 is a block diagram showing details of one of the separate and independent power supplies 303 illustrated in Figure 3.

FIGURE 6 is a block diagram view showing details of the vertical buses and the horizontal buses for supplying power from the separate power supplies shown in Figure 3 to the individual device controllers. The particular bus arrangement shown in Figure 6 permits easy selection of any two of the individual power supplies as the primary and the alternate power supply for a particular device controller.

Fig. 1 is an isometric diagrammatic view of a part of a multiprocessor system incorporating one embodiment of the present invention. In Figure 1 the multiprocessor system is indicated generally by the reference numeral 31.

The multiprocessor system 31 includes individual processor modules 33. Each processor module 33 comprises a central processing unit 105, a memory 107, an input/output channel 109 and an interprocessor control 55.

The individual processor modules are interconnected by interprocessor buses 35 for interprocessor communications.

In a specific embodiment of the multiprocessor system 31, up to sixteen processor modules 33 are interconnected by two interprocessor buses 35 (indicated as the X bus and the Y bus in Figure 1).

Each interprocessor bus has a bus controller 37 associated with that bus.

The microprocessor 113 and microprogram 115 of the central processing unit 105 and an input/output control table 140 in the main memory 107 of each processor module 33 are operatively associated with the I/O channel 109.

The multiprocessor system includes a power distribution system 301 which distributes power from separate power supplies to the processor modules 33 and to the device controllers 41 in a way that permits on-line maintenance and also provides redundancy of power on each device controller.

As illustrated in Figure 3, the power distribution system of the present invention includes separate and independent power supplies 303.

A separate power supply 303 is provided for each processor module 33, and a bus 305 supplies the power from the power supply 303 to the central processing unit 105 and memory 105 of a related processor module 33.

The bus controllers 37, interprocessor buses 35 and interprocessor controls 55 (Figure 1), together with associated microprocessors 113, microprograms 115 and bus receive tables provide an interprocessor bus system.

The multiprocessor system 31 has an input/output (I/O) system for transferring data between the processor modules 33 and peripheral devices, such as the discs 45, terminals 47, magnetic tape drives 49, card readers 51, and line printers 53 shown in Figure 1.

The I/O system includes one I/O bus 39 associated with each I/O channel 109 of a processor module and one or more multiport device controllers 41 may be connected to each I/O bus 39.

In the specific embodiment illustrated, each device controller 41 has two ports 43 for connection to two different processor modules 33 so that each device controller is connected for access by two processor modules.

The I/O system includes a microprocessor 119 and a microprogram 121 in the I/O channel 109 which are dedicated to input/output transfers.

As also illustrated in Figure 3, each device controller 41 is connected for supply of power from two separate power supplies 303 through an automatic switch 311. If one power supply 303 for a particular device controller 41 fails, that device controller is supplied with power from the other power supply 303; and the changeover is accomplished smoothly and without any interruption or pulsation in the power supplied to the device controller.

The power distribution system coacts with the dual port system of the device controller to provide nonstop operation and access to the peripheral devices in the event of a failure of either a single port 43 or a single power supply 303.

The multiprocessor system includes a power on (PON) circuit 182 (the details of which are shown in Figure 2) in several components of the system to establish that the power to that particular component is within certain acceptable limits.

For example, the PON circuit 182 is located in each CPU 105, in each device controller 41, and in each bus controller 37.

The purpose of the PON circuit is to present a signal establishing the level of power applied to that particular component; and if the power is not within certain predetermined acceptable limits, then the signal output is used to directly disable the appropriate bus signal of the component in which the PON is located.

The power-on circuit functions in four states—power off; power going from off to on; power on; and power going from on to off.

The power-on circuit initializes all of the logic states of the system as the power is brought up; and in the present invention, the power-on circuit provides an additional and very important function of providing for a fail-safe system with on line maintenance. To do this, the power-on circuit in the present invention is used in a unique way to control the interface circuits which drive all of the intercommunication buses in the system.

The construction and operation of the power distribution system according to the present invention are illustrated in Figs. 3—6 and are described in detail below under the subtitle Power Distribution System.

The multiprocessor system includes a memory system in which the physical memory is divided into four logical address areas—user data, system data, user code and system code.

The memory system includes a map and control logic for translating all logical addresses to physical addresses and for indicating pages absent from primary storage bit present in secondary storage as required to implement a virtual memory system in which the physical page addresses are invisible to users.

The memory system incorporates a dual port access to the memory by the central processing unit 105 and the I/O channel 109. The I/O channel 109 can therefore access the memory 107 directly (without having to go through the central processing unit 105) for data transfers to and from a device controller 41.

An error detection system is incorporated in the memory system for correcting all single bit and detecting all double bit errors when semiconductor memory is used in the memory system. This error detection system

utilizes a 16 bit data field and a 6 bit check field and includes a data bit complementer for correcting single bit errors.

Before going into the detailed description of the distributed power system, it should be noted that certain terminology will have the following meanings as used in this application.

The term "software" will refer to an operating system or a user program instructions; the term "firmware" will refer to a microprogram in read only memory; and the term "hardware" will refer to actual electronic logic and data storage.

The operating system is a master control program executing in each processor module which has primary control of the allocation of all system resources accessible to that processor module. The operating system provides a scheduling function and determines what process has use of that processor module. The operating system also allocates the use of primary memory (memory management), and it operates the file system for secondary memory management. The operating system also manages the message system. This provides a facility for information transfer over the interprocessor bus.

The operating system arrangement parallels the modular arrangement of the multiprocessor system components described above, in that there are no "global" components.

At the lowest level of the software system, two fundamental entities are implemented—processes and messages.

A process is the fundamental entity of control within a system.

Each process consists of a private data space and register values, and a possibly shared code set. A process may also access a common data space.

A number of processes coexist in a processor module 33.

The processes may be user written programs, or the processes may have dedicated functions, such as, for example, control of an I/O device or the creation and deletion of other processes.

A process may request services from another process, and this other process may be located in the same processor module 33 as a process making the request, or the other process may be located in some other processor module 33.

The processes work in an asynchronous manner, and the processes therefore need a method of communication that will allow a request for services to be queued without "races" (a condition in which the outcome depends upon the sequence of which process started first)—thus the need for "messages" (an orderly system of interprocessor module communication described in more detail below).

Also, all interprocessor module communication should appear the same to the processes, regardless of whether the processes are in the same or in different processor modules.

As will become more clear from the description to follow, the software structure parallels the hardware; and different processes can be considered equivalent to certain components of the hardware in arrangement and function.

For example, just as the I/O channel 109 communicates over the I/O bus 39 to the device controller 41, a user process can make a request (using the message system) to the process associated with that device controller 41; and then the device process returns status back similar to the way the device controller 41 returns information back to the I/O channel 109 over the I/O bus 39.

The other fundamental entity of the software system, the message, consists of a request for service as well as any required data. When the request is completed, any required values will be returned to the requesting process.

When a message is to be sent between processes in two different processor modules 33, the interprocessor buses 35 are used. However, as noted above, all communication between processes appears the same to the processes, regardless of whether they are in the same or in different processor modules 33.

This software organization provides a number of benefits.

This method of structuring the software also provides for significantly more reliable software. By being able to compartmentalize the software structure, smaller module sizes can be obtained, and the interfaces between modules are well defined.

The system is also more maintainable because of the compartmentalization of function.

The well defined modules and the well defined interfaces in the software system also provide advantages in being able to make it easily expandible—as in the case of adding additional processor modules 33 or device controllers 41 to the multiprocessor system.

Furthermore, there is a benefit to the user of the multiprocessor system and software system in that the user, writing his program, need not be aware of either the actual machine configuration or the physical location of other processes.

Just as the hardware provides multiple functionally equivalent modules with redundant interconnects, so does the software.

For example, messages going between processes in different processor modules 33 may use either interprocessor bus 35. Also, device controllers 41 may be operated by processes in either of the processor modules

33 connected to the device controller 41.

POWER DISTRIBUTION SYSTEM:

The multiprocessor system shown in Figure 1 and discussed above incorporates an embodiment of the power distribution system of the present invention.

In many prior art systems it was necessary to stop the processor system in order to perform required maintenance on a component of the system. Also, in many prior art systems, a failure in the power supply could stop the entire processor system.

The power distribution system embodying the present invention incorporates a plurality of separate and independent power supplies and distributes the power from the power supplies to the processor modules and to the device controllers in a way that permits on-line maintenance and also provides redundancy of power on each device controller.

In this regard "on-line" is used in the sense that when a part of the system is on-line, that part of the system is not only powered on, but it is also functioning with the system to perform useful work.

The term "on-line maintenance" therefore means maintaining a part of the system (including periodic preventative maintenance or repair work) while the remainder of the system is on-line as defined above.

In the present invention any processor module or device controller can be powered down so that on-line maintenance can be performed in a power off condition on that processor module or a device controller while the rest of the multiprocessor system is on-line and functional. The on-line maintenance can be performed while fully meeting Underwriters Laboratory safety requirements.

Also, in the power distribution system of the present invention each device controller is connected for supply of power from two separate power supplies and by a diode switching arrangement that permits the device controller to be supplied with power from both power supplies when both power supplies are operative and to be supplied with power from either one of the power supplies in the event the other power supply fails; and the changeover in the event of failure of one of the power supplies is accomplished smoothly and without any interruption or pulsation in the power supply so that an interrupt to a device controller is never required in the event of a failure of one of its associated power supplies.

A power distribution system for insuring both a primary supply and an alternate power supply for each individual dual port device controller 41 is illustrated in Fig. 30. The power distribution system is indicated generally by the reference numeral 301 in Figure 3.

The power distribution system 301 insures that each dual port device controller 41 has both a primary power supply and an alternate power supply. Because each device controller does have two separate and independent sources of power supply, a failure of the primary power supply for a particular device controller does not render that device controller (and all of the devices associated with that controller) inoperative. Instead a switching arrangement provides for an automatic switchover to the alternate power supply so that the device controller can continue in operation. The power distribution system thus coacts with the dual port system of the device controller to provide non-stop operation and access to the devices in the event of a failure of either a single port or a single power supply.

The power distribution system 301 shown in Figure 3 provides the further advantage that each processor module 33 and associated CPU 105 and memory 107 has a separate and independent power supply which is dedicated to that processor module. With this arrangement, a failure of any one power supply or a manual disconnection of any one power supply for repair or servicing of the power supply or associated processor module is therefore limited in effect to only one particular processor module and cannot affect the operation of any of the other processor modules in the multiprocessor system.

The power distribution system 301 shown in Figure 3 thus works in combination with the individual processor modules and the dual port device controllers to insure that a failure or disconnection of any one power supply does not shut down the overall system or make any of the devices ineffective.

The power distribution system 301 includes a plurality of separate and independent power supplies 303, and each power supply 303 has a line 305 (actually a multiline bus 305 as shown in Figure 6) which is dedicated to supplying power to the CPU and memory of a particular, related processor module.

Each device controller 41 is associated with two of the power supplies 303 through a primary line 307 and an alternate line 309 and an automatic switch 311.

A manually operated switch 313 is also associated with each device controller 41 between the device controller and the primary line 307 and the alternate line 309.

The switches 311 and 313 are shown in more detail in Figure 4.

Figure 5 shows details of the component construction of a power supply 303.

As shown in Figure 5, each power supply 303 has an input connector 315 for taking power from the mains. The input 315 is connected to an AC to DC converter 317, and the output of the AC to DC converter

provides, on a line 319, a five volt interruptable power supply (IPS). This five volt interruptable power supply is supplied to the CPU 105, the memory 107 and the device controller 41. See also Figure 6.

The AC to DC converter 317 also provides on a second output line 321 a sixty volt DC output which is supplied to a DC to DC converter 323. See Figure 5.

The DC to DC converter in turn provides a five volt output on a line 325 and a twelve volt output on a line 327.

The outputs from the lines 325 and 327 are, in the system of the present invention, uninterruptable power supply (UPS) outputs in that these power supply outputs are connected to the CPU and memory when semiconductor memory is used. The power supply to a semiconductor memory must not be interrupted because a loss of power to a semiconductor memory will cause loss of all data stored in the memory.

The five volt interruptable power supply on line 319 is considered an interruptable power supply because this power is supplied to parts of the multiprocessing system in which an interruption of power can be accepted. Thus, the five volts interruptable power is supplied to parts of the CPU other than semiconductor memory and to only those parts of the memory which are core memory (and for which a loss of power does not cause a loss of memory) and to the device controller (which as will be described in more detail below) is supplied with an alternate source of power in the event of a failure of the primary power supply.

Since the power supply on lines 325 and 327 must be an uninterruptable power supply, the present invention provides a battery back-up for the input to the DC to DC converter 323. This battery back-up includes a battery and charger module 329. The module 329 is connected to the DC to DC converter 323 by a line 331 and a diode 333.

In a particular embodiment of the present invention the battery 323 supplies power at 48 volts to the converter 323, which is within the input range of the converter 323.

The diode 333 insures that power from the battery is supplied to the converter 323 if the voltage on the line 321 drops below 48 volts. The diode 333 also stops the flow of current from the battery and the line 333 when the output of the AC to DC converter on line 321 exceeds 48 volts.

Each power supply 303 also includes a power warning circuitry 335 for detecting a condition in the AC power input on line 315 that would result in insufficient power out on the output lines 319, 325 and 327. The power warning circuit 335 transmits a power failure warning signal on a line 337 to the related CPU 105.

Because of the capacity storage in the power supply 303, there is enough time between the power warning signal and the loss of the five volts interruptable power on line 319 for the CPU to save its state before the power is lost.

However, the uninterruptable power supply on lines 325 and 327 must not be interrupted, even for an instant of time; and the battery back-up provided by the arrangement shown in Figure 5 insures that there is no interruption in the power supply on lines 325 and 327 in the event of a power failure in the input line 315.

One particular power supply 303 itself can fail for some reason with the other power supplies 303 still operating. In that event, the power distribution system 301 of the present invention limits the effect of the failure of the power supply 303 to the loss of one particular, associated CPU and memory; and the automatic switch 311 provides for an automatic switchover from the failed power supply to the alternate power supply to keep the associated device controller 41 in operation. The device controller 41 which had been connected to the failed power supply therefore continues in operative association with the other processor modules and components of the multiprocessor system, because the required power is automatically switched in from the alternate power supply.

As best illustrated in Figure 4, each automatic switch 311 includes two diodes—a diode 341 associated with the primary power line 307 and a diode 343 associated with the alternate power line 309.

The function of the diodes 341 and 343 is to permit power to be supplied to a device controller 41 from either the primary power line 307 and a related power supply 303 or the alternate power line and its related power supply 303 while keeping the supplies isolated. This prevents a failed power supply from causing its associated alternate or primary from failing.

In normal operation each diode permits a certain amount of current to flow through the diode so that the power to each device controller 41 is actually being supplied by both the primary and alternate power supplies for that device controller.

In the event that one of the power supplies fails, the full power is supplied by the other power supply, and this transition occurs without any loss of power at all.

Since there is a small voltage drop across the diodes 341 and 343, the voltage on the lines 307 and 309 must be enough higher than five volts to accomodate the voltage drop across the diodes 341 and 343 and still supply exactly five volts to the device controller 41. The lines 305 are in parallel with the lines 307 and 309, and the power actually received at the CPU in memory must also be five volts; so balancing diodes 339 are located

in the lines 305 to insure that the voltage after the diodes 339 as supplied to each CPU is exactly five volts.

The manual switch 313 permits a device controller 41 to be disconnected from both the primary and the alternate power sources when the device controller needs to be disconnected for removal and service.

Details of the construction of the switch 313 are shown in Figure 4. As shown in Figure 4, the switch 313 includes a manual switch 345, a transistor 347, a capacitor 348 and a resistor 350 and a resistor 352.

The manual switch 345 is closed to turn on the transistor 347 which then supplies power to the device controller 41.

It is important that both the turn on and the turn off of power to the device controller 41 be accomplished in a smooth way and without fluctuations which could trigger the PON circuit 182 more than once. The feedback capacitor 348 acts in conjunction with the resistor 352 to cause the required smooth ramp build-up of power when the switch 345 is closed to turn the transistor 347 on.

When the transistor 347 is turned off by opening the switch 345, the feedback capacitor 348 acts in conjunction with resistor 350 to provide a smooth fall off of power.

In a preferred embodiment of the invention all of diodes 341, 343 and 339 are Schottky diodes which have a very low forward voltage drop, and this reduces power dissipation.

Each device controller 41 has a power on circuit (PON) 182 for detecting when the five volt power is below specifications. The PON circuit 182 is shown in more detail in Figure 2 and resets the device controller 41 to lock everything off of the device controller and holds the device controller itself in a state that is known when the power is turned off by the switch 313. The PON circuit 182 also releases the device controller and returns it to operation after the power is turned on by switch 313 and five volt power supply at the proper specification is supplied to the device controller 41.

The power on circuit 182 acts in association with transceivers to control the behavior of the transceivers as the device controller 41 is powered up or powered down, in a way which prevents erroneous signals from being placed on the I/O bus while power is going up or down. This feature is particularly significant from the standpoint of on line maintenance.

Each transceiver comprises a receiver 198 and a transmitter 200.

The transmitter is enabled by an enable line 202.

There are several terms which are on the enable line 202. These include the select bit 173, a required input function on the T bus, and a signal from the PON circuit 182.

The signal from the PON circuit, in a particular embodiment of the present invention, is connected in a "wire or" connection to the output of the gate which combines the other terms so that the output of the PON circuit overrides the other terms by pulling down the enable line 202. This insures that the transmitter 200 (in one specific embodiment, an 8T26A or 7438) is placed in a high impedence state until the PON circuit detects that the power is at a sufficient level that the integrated circuits will operate correctly. The PON circuit output stage is designed to take advantage of a property of the specific transceiver integrated circuit used. On this particular type IC if the driver enable line 202 is held below two diode drops above ground potential, the transmitter output transistors are forced into the off state regardless of the level of power applied to the integrated circuit. This insures that the driver cannot drive the bus.

This particular conbination of features provides a mode of operation wherein the output of the integrated circuit is controlled as power comes up or goes down, whereas normally the output of an integrated circuit is undefined when power drops below a certain level.

This same circuit is used on the X and Y buses of the interprocessor bus system to control the transceivers and control signals generated by the interprocessor control 55. As indicated in Figure 3, each central processor unit (CPU) 105 has a PON circuit 182 which is similar to the PON circuit 182 in the device controller. The PON circuits therefore control the transmitters for all of the device controllers 41 and all of the interprocessor controls 55.

Details of the power-on (PON) circuit are shown in Figure 2 where the circuit is indicated generally by the reference numeral 182.

The purpose of the PON circuit is to sense two different voltage levels of the five volt supply.

If power is failing, the circuit senses the point at which the power drops below a certain level which renders the logic in the device controller or CPU an indeterminate state or condition. At this point the circuit supplies signals to protect the system against the logic which subsequently goes into an undefinable state.

The second voltage level which the PON circuit will sense is a value that is perceived when power is coming up. This second level at which power is sensed will be greater than the first level by roughly 100 millivolts to provide hysteresis for the system to eliminate any conditions of oscillation.

The PON circuit stays in a stable condition after it senses one of the voltage conditions until it senses the other voltage condition, at

which point it changes state. The state at which the PON circuit is in at any particular time determines the voltage level at which the transition to the other state will be made.

The power on circuit 182 thus presents a signal establishing an indication that the power is within predetermined, acceptable operating limits for the device controller 41. If the power is not within those predetermined, acceptable operating limits, the signal output of the power-on circuit 182 is used to directly disable the appropriate bus signals of the device controller 41.

The output of the PON circuit 182 is a binary output. If the output is a one, the power is within satisfactory limits. If the output of the PON circuit is a zero, this is an indication that the power is below the acceptible limit.

The power-on circuit 182 shown in Figure 3 and to be described in detail below is used with the device controller 41 and has seven output driver stages which are used in the application of the power-on circuit 182 to the device controller 41. However, the same power-on circuit 182 is also used with the CPU 105 and the bus controller 37, but in those applications the power-on circuit will have a lesser number of output driver stages.

As illustrated in Figure 2, the PON circuit 182 comprises a current source 184 and a differential amplifier 186.

The differential amplifier 186 has, as one input, a temperature compensated reference voltage input on a line 188 and has a second input on a line 190 which is an indication of the voltage that is to be sensed by the power-on circuit.

The reference voltage on line 188 is established by a zener diode 192.

The differential amplifier 186 comprises a matched pair of transistors 194 and 196.

The voltage applied on the line 190 is determined by resistors 198, 200 and 202. The resistors 198, 200 and 202 are metal film resistors which provide a high degree of temperature stability in the PON circuit.

The outputs on lines 204 and 206 of the differential amplifier 186 are applied to a three transistor array (the transistors 208, 210 and 212), and this three transistor array in turn controls the main output control transistor 214.

The main output control transistor 214 drives all output drivers that are attached. For example, in the application of the PON circuit 182 for the device controller 41 (as illustrated in Figure 2), the main output transistor 214 drives output stages 216 through 228. The output stage 216 is used to clear the logic, the output stages 218, 220 and 222 are used in combination with the interface devices of one port 43 of the device controller 41, and the output stages 224, 226 and 228 are used in combination with the interface device of the other port 43 of the device controller 41.

Finally, the PON circuit 182 includes a hysteresis control 230. The hysteresis control 230 includes resistors 232, 234 and a transistor 236.

In operation, assuming that operation is started from a power off state to a power on condition, the power is applied through the current source 182 to the differential amplifier 186 and to the main output control transistor 214. At this time the voltage on the line 190 is less than the voltage on the line 188 so the differential amplifier 186 holds the output of the main output control transistor 214 in the off state. This, in turn, will force the output stages 216 through 228 on.

This asserts the output of the PON circuit 182 in the zero state, the state indicating that power is not within acceptable limits.

As voltage rises, the input voltage on line 190 will increase until it equals the reference voltage on line 188. At this point the differential amplifier 186 drives the main output control transistor 214, turning it on. This removes the base drive from the output stages 216 through 228, forcing these output stages off. The output of the PON circuit 182 is then a one, indicating that the power is within acceptable limits.

At this point the hysteresis control circuit 230 comes into play. While power was coming on, the transistor 236 of the hysteresis control circuit 230 was on. When the transistor 236 is on, the resistance value of the resistor 202 appears to be less than the resistance value of this resistor 202 is when the transistor 236 is off.

The point at which the main output control transistor 214 turns on is the point at which the hysteresis transistor 236 turns off. Turning off the hysteresis transistor 236 causes a slight voltage jump in the line 190 which further latches the differential amplifier 186 into the condition where the differential amplifier 186 sustains the main output transistor 214 in the on state.

The state of the PON circuit will remain stable in this condition with the main output control transistor 214 on and the output drivers 216 through 228 off until the plus five volts drops below a lower threshold point, as determined by the voltage applied on the line 190.

As the voltage on the line 190 decreases below the reference voltage on the line 188, (because the five volts supply is going down in a power failure condition), then the differential amplifier 186 turns off the main output control transistor 214. This, in turn, turns on the output driver stages 216 through 228.

Since the hysteresis transistor 236 was off as power dropped, the voltage applied to the input of the PON circuit 182 must drop

somewhat farther than the point at which the PON circuit 182 sensed that power was within the acceptable limits during the power-up phase of operation.

This differential or hysteresis is used to inhibit any noise on the five volt power supply from causing any oscillation in the circuit that would erroneously indicate that power is failing.

The PON circuit 182 shown in Figure 2 provides very accurate sensing of the two voltages used by the PON circuit to determine its state (whether a one or a zero output of the PON circuit).

In order to sense these two voltages very accurately the PON circuit must have the capability of compensating for initial tolerances of the different components and also the capability to compensate for changes in temperature during operation. In the PON circuit 182, the zener diode 192 is the only critical part that must be compensated for because of its initial tolerance, and this compensation is provided by selecting the resistor 198.

Temperature compensation is achieved because the zener diode 192 is an active zener diode and is not a passive zener diode. Effective temperature compensation is also achieved because the two transistors in the differential amplifier 186 are a matched pair of transistors and the resistors 198, 200 and 202 are metal film resistors.

With reference to Figure 6, the power from each power supply 303 is transmitted to a related CPU by the vertical bus 305, and each vertical bus 305 is a laminated bus bar which has five layers of electrical conductors.

As indicated by the legends in Figure 6, each vertical bus 305 has two different conductors connected to ground.

One conductor provides the ground for both the five volt interruptable power supply (IPS) and the five volt uninterruptable power supply (UPS).

A separate conductor provides a ground for the memory voltage. This separate ground for the memory voltage insures that the relatively large fluctuations in current to the memory will not have any effect on either the five volt IPS or the five volt UPS supplied to the CPU.

The horizontal bus 305, 307 includes the primary and alternate power supply lines 307 and 309 (as indicated by the reference numerals in Fig. 30). In a particular embodiment of the present invention the bus 305, 307 is actually a nine layer laminated bus which has a single ground and eight voltage layers (V1 through V8 as indicated by the legends and notations in Figure 6).

Each voltage layer is connected to the five volt interruptable output of a different power supply 303. Thus, the layer V1 is connected at 351 to the five volt IPS power for the power supply 303 and related processor module farthest to the left as viewed in Figure 6, and the layer V2 is connected at 353 to the five volt IPS power supply 303 for the processor module at the center as viewed in Figure 6, and so on.

Since there are eight layers (V1 through V8) and a common ground available to each device controller in the horizontal bus, upstanding vertical taps 355 to these eight layers at spaced intervals along the horizontal bus permit each device controller 41 to be associated with any two of the power supplies 303 merely by connecting the primary line 307 and the alternate line 309 to a particular set of taps. By way of example, the device controller 41 on the lefthand side of Fig. 33 is shown connected to the taps V1 and V2 and the device controller 41 on the righthand side of Figure 6 is shown connected to the taps V2 and V3.

Thus, any device controller 41 can be connected to any two of the power supplies 303 with any one of the power supplies serving as the primary power supply and any one of the other power supplies serving as the alternate power supply.

The power distribution system of the present invention thus provides a number of important benefits.

The power distribution system permits on line maintenance to be performed because one processor module or device controller can be powered down while the rest of the multiprocessor system is on line and functional.

The power distribution system fully meets all Underwriter Laboratory safety requirements for doing on line maintenance of a powered down component while the rest of the multiprocessor system is on line and in operation.

Each device controller is associated with two separate power supplies so that a failure in one of the power supplies does not cause the device controller to stop operation. Instead, the electronic switch arrangement of the present invention provides such a smooth transition of power from the two power supplies to only one of the power supplies that the device controller is maintained in continuous operation without an interrupt.

Reference is directed to our co-pending Application No. 25490/77 (Serial No. 1,588,803) from which the present Application is divided and to Applications Nos. 3057/80 (Serial No. 1,588,804), 3072/80 (Serial No. 1,588,806) and 3073/80 (Serial No. 1,588,807).

WHAT WE CLAIM IS:—

1. A distributed power system for multiprocessor system of the kind in which separate processor modules each having a central processing unit are interconnected for paral-

somewhat farther than the point at which the PON circuit 182 sensed that power was within the acceptable limits during the power-up phase of operation.

This differential or hysteresis is used to inhibit any noise on the five volt power supply from causing any oscillation in the circuit that would erroneously indicate that power is failing.

The PON circuit 182 shown in Figure 2 provides very accurate sensing of the two voltages used by the PON circuit to determine its state (whether a one or a zero output of the PON circuit).

In order to sense these two voltages very accurately the PON circuit must have the capability of compensating for initial tolerances of the different components and also the capability to compensate for changes in temperature during operation. In the PON circuit 182, the zener diode 192 is the only critical part that must be compensated for because of its initial tolerance, and this compensation is provided by selecting the resistor 198.

Temperature compensation is achieved because the zener diode 192 is an active zener diode and is not a passive zener diode. Effective temperature compensation is also achieved because the two transistors in the differential amplifier 186 are a matched pair of transistors and the resistors 198, 200 and 202 are metal film resistors.

With reference to Figure 6, the power from each power supply 303 is transmitted to a related CPU by the vertical bus 305, and each vertical bus 305 is a laminated bus bar which has five layers of electrical conductors.

As indicated by the legends in Figure 6, each vertical bus 305 has two different conductors connected to ground.

One conductor provides the ground for both the five volt interruptable power supply (IPS) and the five volt uninterruptable power supply (UPS).

A separate conductor provides a ground for the memory voltage. This separate ground for the memory voltage insures that the relatively large fluctuations in current to the memory will not have any effect on either the five volt IPS or the five volt UPS supplied to the CPU.

The horizontal bus 305, 307 includes the primary and alternate power supply lines 307 and 309 (as indicated by the reference numerals in Fig. 30). In a particular embodiment of the present invention the bus 305, 307 is actually a nine layer laminated bus which has a single ground and eight voltage layers (V1 through V8 as indicated by the legends and notations in Figure 6).

Each voltage layer is connected to the five volt interruptable output of a different power supply 303. Thus, the layer V1 is connected at 351 to the five volt IPS power for the power supply 303 and related processor module farthest to the left as viewed in Figure 6, and the layer V2 is connected at 353 to the five volt IPS power supply 303 for the processor module at the center as viewed in Figure 6, and so on.

Since there are eight layers (V1 through V8) and a common ground available to each device controller in the horizontal bus, upstanding vertical taps 355 to these eight layers at spaced intervals along the horizontal bus permit each device controller 41 to be associated with any two of the power supplies 303 merely by connecting the primary line 307 and the alternate line 309 to a particular set of taps. By way of example, the device controller 41 on the lefthand side of Fig. 33 is shown connected to the taps V1 and V2 and the device controller 41 on the righthand side of Figure 6 is shown connected to the taps V2 and V3.

Thus, any device controller 41 can be connected to any two of the power supplies 303 with any one of the power supplies serving as the primary power supply and any one of the other power supplies serving as the alternate power supply.

The power distribution system of the present invention thus provides a number of important benefits.

The power distribution system permits on line maintenance to be performed because one processor module or device controller can be powered down while the rest of the multiprocessor system is on line and functional.

The power distribution system fully meets all Underwriter Laboratory safety requirements for doing on line maintenance of a powered down component while the rest of the multiprocessor system is on line and in operation.

Each device controller is associated with two separate power supplies so that a failure in one of the power supplies does not cause the device controller to stop operation. Instead, the electronic switch arrangement of the present invention provides such a smooth transition of power from the two power supplies to only one of the power supplies that the device controller is maintained in continuous operation without an interrupt.

Reference is directed to our co-pending Application No. 25490/77 (Serial No. 1,588,803) from which the present Application is divided and to Applications Nos. 3057/80 (Serial No. 1,588,804), 3072/80 (Serial No. 1,588,806) and 3073/80 (Serial No. 1,588,807).

WHAT WE CLAIM IS:—

1. A distributed power system for multiprocessor system of the kind in which separate processor modules each having a central processing unit are interconnected for paral-

* C L M S *

lel processing, said distributed power system comprising: a plurality of separate processor modules; a plurality of device controllers for controlling the transfer of data between a plurality of processor modules and a plurality of peripheral devices; multiple ports in each device controller and multiple input/output buses for connecting each device controller for access by different processor modules; and distributed power supply means for insuring non-stop operation of the remainder of the multiprocessor system in the event of a failure of a single power supply to a part of the multiprocessor system, said distributed supply means including a separate power supply for each processor module and connecting means for connecting a plurality of the separate power supplies to each device controller and effective to supply power to each device controller from the remaining separate power supplies in the event of the failure of one of the power supplies connected to any device controller.

2. A system according to claim 1, including power-on circuit means for establishing that power levels are within certain established limits and effective to function in four states—power-off, power going from off to on, power-on and power going from on to off.

3. A system according to claim 2, wherein an interprocessor bus interconnects all of the processor modules and wherein a power-on circuit means is located in each central processing unit, and in each device controller.

4. A system according to claim 3 wherein the power-on circuit means controls interface circuits which drive all of the interprocessor and input/output buses in the multiprocessor syste.

5. A system according to claim 2, 3 or 4 wherein the power-on circuit means presents a signal in the event of a variation of logic power which signal establishes the level of power applied to the related processor module or the device controller and wherein the signal output is used to directly disable appropriate interface signals of the processor module or device controller in which the power-on circuit is located if the power-on is not within the certain established limits which will ensure correct logic operation.

6. A system according to claim 3, 4 or 5, including a second interprocessor bus interconnecting all of the processor modules and wherein the power-on circuit means coacts with a multi-path system provided by the interprocessor buses and the multiple port device controllers and associated input/output buses to give a mode of operation in which a failing power supply for one processor or a multiple port device controller does not affect any other processor module or dual port controller on any of those paths.

7. A system according to any preceding claim, where the connecting means includes a diode switch arrangement which supplies power to the device controller from a plurality of power supplies when those associated power supplies are operative and when supplies power from the remaining power supplies in the event of a failure of one of the associated power supplies and in a changeover which is without interruption or pulsation so that operation of the device controller is never interrupted.

8. A system according to any preceding claim, wherein the separate power supplies are interconnected by a power bus having a plurality of lines with selected ones of the lines connected to selected ones of the power supplies and including taps on the power bus adjacent each device controller for permitting the device controller to be connected to any selected plurality of the power supplies by easy connection to related taps in the power bus.

9. A system according to any preceding claim, wherein each device controller is associated with two separate power supplies.

10. A distributed power system according to claim 1 substantially as hereinbefore described with reference to, and as shown in the accompanying drawings.

FORRESTER, KETLEY & CO.,
Chartered Patent Agents,
Forrester House,
52 Bounds Green Road,
London N11 2EY,
also at
Rutland House,
148 Edmund St.,
Birmingham B3 2LD,
and
Scottish Provident Building,
29 St. Vincent Place,
Glasgow G1 2DT.
Agents for the Applicants.

Printed for Her Majesty's Stationery Office by Burgess & Son (Abingdon) Ltd.—1981. Published at The Patent Office, 25 Southampton Buildings, London, WC2A 1AY, from which copies may be obtained.

* C L M S *

(9)

SPECIFICATION NO 1588805

By a direction given under Section 17 (1) of the Patents Act 1949 this application proceeded in the name of TANDEM COMPUTERS INCORPORATED, a Corporation organised and existing under the laws of the State of Delaware, of 19333 Vallco Parkway, Cupertino, California 95015, United States of America.

Bas 83005/5

THE PATENT OFFICE