



159159



G00002

SYSTEM AND METHOD FOR USING A MEMORY MAPPING FUNCTION TO MAP MEMORY DEFECTS

TECHNICAL FIELD

5 The present disclosure relates generally to computer systems and information handling systems, and, more particularly, to a system and method for creating a map of memory which governs its use by the computer system or information handling system.

BACKGROUND

10 As the value and use of information continues to increase, individuals and businesses seek additional ways to process and store information. One option available to these users is an information handling system. An information handling system generally processes, compiles, stores, and/or communicates information or data for business, personal, or other purposes thereby allowing users to take advantage of the value of the information. Because
15 technology and information handling needs and requirements vary between different users or applications, information handling systems may vary with respect to the type of information handled; the methods for handling the information; the methods for processing, storing or communicating the information; the amount of information processed, stored, or communicated; and the speed and efficiency with which the information is processed, stored, or communicated.
20 The variations in information handling systems allow for information handling systems to be general or configured for a specific user or specific use such as financial transaction processing, airline reservations, enterprise data storage, or global communications. In addition, information handling systems may include or comprise a variety of hardware and software components that may be configured to process, store, and communicate information and may include one or more
25 computer systems, data storage systems, and networking systems.

 An information handling system will typically include some type of temporary information storage medium, such as random access memory. In recent computers, the amount

of memory comprised by the information handling system may be on the order of gigabytes. As memory size increases, the likelihood that part of the memory will either be manufactured defective or become defective over time increases dramatically. If left unmanaged, the presence of defective memory cells, regardless of their size, can cause the information handling system to fail. Such failure can initiate an abrupt end to the current operation of the information handling system, resulting in the loss of critical data, or can prevent the information handling system from starting up altogether.

As systems continue to evolve and computer technology advances, the operational relationship between the CPU and memory becomes more profound. Many attributes of modern systems (specifically the introduction of multi-core processors and virtualization in general) are forcing an ever-growing memory footprint. Consequently, not only is system memory growing to be a much more substantial percentage of the overall solution cost, but the impact of erroneous behavior in the memory can have a much more adverse effect on the life cycle expense associated with service. Traditionally, embedded error correction schemes have successfully been used to avoid "crashes" on a single bit error and to detect multi-bit corruptions. However, as memory geometries become smaller and memory size grows larger, it is necessary to add another level of service protection that is quickly available and resident on the memory module itself.

SUMMARY

In accordance with the present disclosure, a system and method are herein disclosed for managing memory defects in an information handling system. More particularly, the present disclosure comprises a system and method for generating a usable memory map which excludes memory locations containing defect memory elements. In an information handling system, a local memory defect map, which comprises information about the location of defective memory elements within a memory device, is coupled to the memory device. BIOS uses the local memory defect map to generate a system memory defect map. As a map of memory usable by the system is created, usable memory regions containing defective memory elements are excluded from the usable memory map. The usable memory map is passed to the operating system piece by piece, which uses only those regions of memory designated as usable and non-defective.

The system and method disclosed herein are technically advantageous because a mechanism is provided for reducing the functional and financial impact of defective memory elements. In particular, the present disclosure provides a way for an information handling system to operate normally, despite the presence of one or more defective memory elements, with minimal impact on system performance. By effectively hiding defective memory blocks from the operating system, crashes of the information handling system are better avoided. In addition, by sequestering defective memory elements, the present disclosure extends the life of memory. Overall system costs for the information handling system are correspondingly reduced as memory need not be replaced even after multiple memory elements are found to be defective.

BRIEF DESCRIPTION OF THE DRAWINGS

A more complete understanding of the present embodiments and advantages thereof may be acquired by referring to the following description taken in conjunction with the accompanying drawings, in which like reference numbers indicate like features, and wherein:

5 Figure 1 is a block diagram of a typical information handling system;

 Figure 2A shows the addressable memory space for a typical information handling system;

 Figure 2B shows the memory space included in a usable memory map which excludes defect blocks;

10 Figure 3 is a flow diagram depicting an implementation of a memory mapping function;

 Figure 4 shows the code flow for an implementation of a memory mapping call.

DETAILED DESCRIPTION

For purposes of this disclosure, an information handling system may include any instrumentality or aggregate of instrumentalities operable to compute, classify, process, transmit, receive, retrieve, originate, switch, store, display, manifest, detect, record, reproduce, handle, or
5 utilize any form of information, intelligence, or data for business, scientific, control, or other purposes. For example, an information handling system may be a personal computer, a network storage device, or any other suitable device and may vary in size, shape, performance, functionality, and price. The information handling system may include random access memory (RAM), one or more processing resources such as a central processing unit (CPU) or
10 hardware or software control logic, ROM, and/or other types of nonvolatile memory. Additional components of the information handling system may include one or more disk drives, one or more network ports for communication with external devices as well as various input and output (I/O) devices, such as a keyboard, a mouse, and a video display. The information handling system may also include one or more buses operable to transmit communications
15 between the various hardware components.

Shown in Figure 1 is a block diagram of a typical information handling system. A processor or CPU 10 of the typical information handling system 5 is communicatively coupled to a memory controller hub or northbridge 30. Memory controller hub 30 is coupled to memory 20 and a graphics processing unit 40. Memory controller hub 30 is also coupled to an I/O controller
20 hub or southbridge 60. I/O hub 60 is coupled to storage elements of the computer system, including a storage element 50 such as flash ROM for the BIOS of the computer system and the hard drive 70 of the computer system. I/O hub 60 may also be coupled to a Super I/O chip 80, which is itself coupled to many of the I/O ports of the computer system, including keyboard 90, mouse 100, and parallel ports.

25 During startup of the information handling system 10, the components of the system are initialized during BIOS setup. The initialization process includes mapping the system memory that will be available to the operating system (OS) once the initialization is complete.

The system memory map is created by BIOS and communicated to the OS through the memory mapping call, such as INT15-E820. During BIOS initialization, memory defect information stored in local memory defect maps is translated into a system memory defect map. The operating system uses iterative calls to the BIOS memory mapping functions to generate the complete usable memory map. The complete map details usable, non-usable, and ACPI reclaimable address ranges.

Figure 2A is a diagram of the memory space that is reviewed by a memory mapping call, such as an E820 call, within the information handling system. The information handling system 10 may be comprised of numerous distinct components capable of being addressed as memory, including dual inline memory modules (DIMMs) 120 and PCI devices 130, such as discrete graphics cards. Devices that include addressable memory are detected during the BIOS initialization period, and the memory within these devices is scanned and identified as system-accessible memory during the memory scanning operation. The memory scanning operation results in the iterative interrogation of possible memory locations, beginning with memory having the lowest possible address 110 within the system. The memory scanning operation reviews all addressable memory locations in succession, including memory blocks 124, 125 and 126, even though memory block 125 contains a memory defect. As system-accessible memory is located, including DIMMs 120, memory regions are identified based on the type of memory in that region. These type-based memory region designations may be stored in numerous locations, including CMOS or RAM.

Consistent with this disclosure, the BIOS also generates a system memory defect map during POST based on known memory defects. Information regarding previously detected memory defects may be stored in a number of locations, including the serial presence detect (SPD) EEPROM of the DIMM module that includes the associated defective memory cell. During BIOS initialization, the stored local memory defect map is read to determine the location of memory defect blocks in the corresponding DIMM. This information is then translated into a system memory defect map which stores information for identifying defective memory blocks at

the system level. In the system memory defect map, for instance, the memory defect blocks may be represented by a starting address in the system memory address space, the length of the memory block, and a flag indicating whether the memory block contains defective memory or not. By contrast, the local memory defect map for each memory device may simply be a series of flags representing the presence or absence of defective memory in sequentially identified blocks of the corresponding memory device. More sophisticated local memory defect maps are possible, but regardless of complexity, the information in the local memory defect map should be translated into the system memory address space used by the OS. This function is best performed by the BIOS, though the OS could do the same.

The OS ultimately identifies the usable memory available to it by calling a memory mapping function such as an INT15-E820 call. Consistent with this disclosure, the memory mapping function reads the memory types of the memory regions identified by the memory scanning operation, and also reads the system memory defect map. The memory mapping function may check for overlaps to determine whether the memory defect blocks indicated by the system memory defect map coincide with the boundaries of the memory type regions identified by the memory scanning operation. The memory mapping function then returns information to the OS indicating the next region of usable memory. Because the system memory defect map is consulted during the E820 call, defective memory blocks will not be identified as usable memory. As a result, the operating system has no knowledge of or access to the defective memory, but instead sees only the usable and non-defective memory. Defective memory is thus effectively quarantined without the use of software and without any decrease in hardware performance.

Figure 2B is a diagram of a usable memory map in which the unusable or defective locations in the memory of the computer system have been logically mapped out of the computer system. For example, memory blocks 124 and 126 are included in the usable memory map, but memory block 125, which contains a defective memory element, is not.

One method of mapping out defective memory blocks involves creating the entire system memory defect map during BIOS power-on self-test (POST) and storing each memory map entry until the OS requests it. Because each standard E820 entry is 20 bytes, however, mapping any significant number of defects with this method requires a large quantity of storage.

5 For an 8 DIMM system supporting 8 defects per DIMM, for example, this method would create 64 reserved entries and 64 usable entries, as well as standard entries for base memory, HECBase, reclaim area and reserved area below 4GB. This totals 132 entries, or 2640 bytes from the BIOS runtime area. This quantity of storage is sufficiently large to require placement in a reserved memory area outside F000, slowing down access to the map information and tying up limited

10 memory resources.

Another method of mapping the memory could entail scanning the local memory defect map, for example from a DIMM's SPD, during E820 entry creation, interpreting the local memory defect map and returning holes in real time as the OS calls the memory mapping function. This method requires significant time to process the data, however, causing longer

15 boot-up times.

In a third method for building the system memory defect map with defect information, the map would be built during POST for all of the reserved areas described in the DIMM defect data. Using efficient addressing techniques, the size of each entry could be reduced to as little as 8 bytes. In the 8 DIMM system described previously, the entire map would

20 be as small as 512 bytes, a reduction in size by a factor of 5. This method requires a more intelligent memory mapping function than the current version of the E820 call, and could implement, for instance, an algorithm to retain pointer offset in the memory defect map (located, for instance, in the DIMM SPD) created during POST and scanning forward and backward in the structure to determine next entry types, overlaps, etc.

25 In one example of a more efficient memory mapping algorithm, a continuation value would temporarily store information about the progress of the iterative calls to the memory mapping function. The continuation value could be stored, for example, in the BX register of a

typical information handling system. Figure 3 depicts one possible sequence of events implementing an efficient memory mapping algorithm. The system is initially powered on 305 and begins the POST process 310. As part of the POST, the memory space is scanned 315 to determine which locations are reserved for other devices, which are reclaimable, and which are usable. Next, the local memory defect map is read 317, followed by the creation of the system memory defect map 318. After the POST process concludes 319, the OS initiates the first memory mapping call 320.

During the memory mapping call, the BIOS first examines the continuation value 325 to determine where the previous call, if any, left off. If no previous calls have been made, the continuation value will still be in its initialized state, pointing to the zero address of the memory defect map. The BIOS reads the memory defect map entry at the offset indicated by the continuation value 330 and compares that memory location against the reserved and reclaimable memory addresses 335 determined during its previous memory scan 315 and stored in CMOS or memory. After reconciling any overlaps, the BIOS generates the memory map entry requested by the memory mapping function call 340. Before passing the memory map entry to the OS 350, however, the BIOS updates the continuation value to reflect its progress through the memory defect map 345. Once the memory map entry is sent to the OS 350, the OS determines whether or not the continuation value indicates the end of the usable memory space 355. If so, the memory mapping function is complete at the iterative process terminates 360. If not, the memory mapping call is repeated, beginning with the BIOS examining to the continuation value to determine where it left off 325.

Those skilled in the art will recognize that the continuation value could be structured in numerous ways to implement the foregoing memory mapping algorithm. In one implementation, a single bit could be used to determine whether the next memory map entry falls within a defective memory block. Assuming a memory defect map size of 128 entries or less, seven bits would be sufficient to store an offset to either the current or previous location read from the memory defect map during the memory mapping iterations. Another bit could be used

to determine whether or not the current memory location is above the 4GB threshold or not, and two more bits could serve as a counter to indicate how many entries have been created since crossing the 4GB threshold.

To further reduce the large size of individual memory map entries, the excess
 5 addressability of the fields of each memory map entry can be trimmed. In traditional E820 maps, for instance, each memory map entry includes a base address, the size of the memory block being sent to the memory map, and the type of memory. When excluding defective memory elements from the memory map, the length of the memory map entry field corresponding to the size of the memory block at issue can be reduced to the megabyte range, since DIMMs with capacity greater
 10 than 1GB are still fairly rare. Likewise, since only three general types of memory are returned to the OS, the length of the memory map entry field denoting memory type could be reduced to as little as two bits.

Figure 4 depicts one possible code sequence to implement the third method above, where the system memory defect map is built during POST for the entries found in the local
 15 memory defect map. First, an entry handler 400 uses a continuation value to determine special handling routines for calling specific routines as well as basic fault checks. Next, the first INT15 E820 call 410 is made. During the first INT15 call 410, offsets are initialized into tables and returned entries and the standard base memory area is returned. The second INT15 call 430 comprises five steps. First, the information handling system checks for valid memory defect
 20 entries and overlaps for a start address. Next, the structure data is checked for invalid ranges and sizes. Third, the information handling system checks for the next entry type and overlaps with reserved areas. Fourth, flags and offsets are set for the next entry type and routine calls are made. Finally, the needed entry type is return with its associated address and range.

Subsequent INT15 calls differ for memory ranges above 4 gigabytes and below 4
 25 gigabytes. If the memory range is blow 4 gigabytes, the INT15 call 460 first checks for offsets and flags to determine the entry log type required. Next, the structure data is checked for invalid ranges and sizes. Third, the information handling system checks for overlaps and required

ranges. Fourth, the information handling system scans through structures to determine the next entry, setting flags and offsets. Fifth, the needed entry type is returned with its associated address and range. Finally, the system checks if the current entry is the last entry below 4GB. For INT15 calls above 4GB 480, the information handling system first checks for valid structure entries and overlaps. Next, the information handling system checks the structure data for invalid ranges and sizes. Third, flags and offsets are set for the next entry. Finally the needed entry type is returned with the calculated address and range.

Although this disclosure has been described with respect to the creation of a usable memory map, such as an E820 map, in an information handling system, it should be recognized that the memory mapping system and method described herein may be implemented with any physical storage device with potential defects. Although the present disclosure has been described in detail, it should be understood that various changes, substitutions, and alterations can be made hereto without departing from the spirit and the scope of the invention as defined by the appended claims.

WHAT IS CLAIMED IS:

1. A system for managing memory defects, comprising:
 - an information handling system;
 - 5 an operating system;
 - memory;
 - a memory defect map, wherein the memory defect map comprises information regarding the location of regions of said memory which contain defective memory elements;
 - a memory mapping function, wherein the memory mapping function is operable to
 - 10 create a map of usable memory and send the map of usable memory to the operating system, and wherein the memory mapping function is further operable to reference the memory defect map and exclude memory regions containing defective memory elements from the usable memory map.
- 15 2. The system of claim 1, wherein the device storing the memory defect map is physically coupled to the memory.
3. The system of claim 1, wherein the memory mapping function is a modification of the INT15-E820 memory description call.
- 20 4. The system of claim 1, wherein the memory mapping function is called before the operating system is initialized.
5. The system of claim 1, wherein the entire usable memory map is built during the
- 25 power-on self-test period of the information handling system.

6. The system of claim 1, wherein the usable memory map is generated by scanning the memory defect map during the creation of each entry in the usable memory map.

7. The system of claim 1, wherein the usable memory map is built during power-on self-test for each of the reserved memory regions indicated in the memory defect map.

8. The system of claim 1, wherein the size of memory map entries is reduced by referencing memory through offsets.

9. The system of claim 1, wherein the size of memory map entries is reduced by limiting the length of the field corresponding to the size of a memory block.

10. The system of claim 1, wherein the size of memory map entries is reduced by limiting the length of the field corresponding to the type of memory being mapped during the call.

11. The system of claim 1, wherein a continuation value register stores information about the progress of the memory mapping function.

12. The system of claim 11, wherein the continuation value register comprises an offset value pointing to a location in the memory defect map.

13. The system of claim 11, wherein the continuation value register comprises information about the type of memory identified by the following usable memory map entry.

14. A method of managing memory defects in an information handling system, comprising the steps of:

storing a memory defect map comprising information about the location of defective memory elements;

5 creating a map of available memory for use by the information handling system;
 excluding from the map of available memory regions of memory containing defect memory elements;

 making the map of available memory accessible to the operating system.

10 15. The method of claim 14, wherein the entire map of available memory is built during the power-on self-test period of the information handling system.

 16. The method of claim 14, wherein the map of available memory is generated by scanning the memory defect map during the creation of each entry in the map of available
15 memory.

 17. The method of claim 14, wherein the map of available memory is built during power-on self-test for each of the reserved memory regions indicated in the memory defect map.

18. A method for building a usable memory map, comprising the steps of:
calling a memory mapping function;
reading a continuation value;
reading information from a memory defect map based on the continuation value;
5 returning information regarding the location of usable and non-defective memory.

19. The method of claim 18, further comprising the step of determining whether an overlap exists between a memory range read from the memory defect map and other information regarding the type of memory in the memory range.

10

20. The method of claim 18, wherein the continuation value comprises information regarding the progress of the memory mapping function as it iterates through the memory defect map.