Intel® Chipset 4 GB System Memory Support

White Paper

February 2005

Revision 1.0

Contents

Figures

Tables

Revision History

Revision Number	Description	Revision Date
0.5	• Initial Release	September 2004
1.0	• Redesignated as public document	February 2005

1 Preface

This document addresses potential customer questions that may arise when a system is populated with the maximum memory configuration (e.g., 4 GB for an Intel® 915G Express chipset), and the operating system (OS) reports less available memory than that populated (OS reports approximately 3.2 GB of available memory). It is intended for training of customer support and field personnel at hardware system manufacturers and for software developers of applications, operating system, and tools.

This document outlines the division and potential overlap of the addressable physical and virtual memory ranges by system BIOS for required system and PC architectural resources and operating system usage.

It offers background on ways that system developers may mitigate the potential overlaps and better describe the memory usage to the end customer. The first method used to decrease potential overlaps is memory remapping, used today by some workstation and server chipsets to maximize the amount of populated system memory that can be allocated to operating systems. Possible solutions that the system BIOS may incorporate to clearly delineate via setup menus include: populated, system resource, and operating system available memory.

This document also provides experimental results using a current desktop platform as a base line and a mid-range server platform that incorporates memory remapping to show the potential gains with multiple operating systems and varying amounts of system memory. Experimental data is not inclusive of all available or potentially available chipset solutions, memory sizes platform configurations, or operating system variations.

1.1 Key Takeaways

- Due to PC architectural requirements such as motherboard resources and OS limitations, desktop platforms using larger memory may be unable to take full advantage of all memory populated on the system.
- When systems with chipsets that support 4 GB or less of system memory are populated with the maximum system memory of 4 GB, the operating system (OS) may report a lower amount of available memory
- Standard PC Architecture System Resources require addressing which overlaps physical memory below 4 GB:
  - System BIOS
  - Motherboard Resources (I/OxAPIC)
  - Memory Mapped I/O
  - PCI Express* Configuration Space
  - Additional PCI Device Memory (Graphics Aperture)
  - VGA Memory
  - Others as included, etc.
- These requirements may reduce the addressable memory space available to and reported by the operating system

- These memory ranges, while unavailable to the OS, are still being utilized by subsystems such as I/O, PCI Express and Integrated Graphics and are critical to the proper functioning of the PC

§

2 Platform Memory Subsystem Usage

2.1 Definition of Address Ranges

Current PC memory usage model and memory mapping architecture are based on 32-bit CPUs, and can support up to 4 GB of physical memory. Addressable memory is defined as memory that the processor addresses on the front side bus. Current Intel CPUs support a 36-bit addressable front side bus giving a maximum potential physical address space of 64 GB. Physical memory is the actual memory hardware installed in the system. Each byte of physical memory is assigned a unique address, called a physical address. The addressable physical memory that chipsets support varies depending upon market segment requirements, but will normally be designed to be less than, or equal to, the addressable range of the processor bus. Any task or program running on an IA-32 processor can address a linear address space of up to 4 GB and a physical address space of up to 64 GB.

2.2 Memory Initialization and Assignment of Resources

When the system is powered on, system BIOS initializes all of the populated memory on the system, enables onboard and add-in devices, assigns Memory and IO resources (internal graphics, IOxAPICs, PCI Space, or other memory mapped IO ranges), and dynamically builds industry standard tables (Int 15h E820h, and ACPI Tables) that describe the memory range usage to the operating system.

In uni-processor based systems for mobile, desktop, workstation, and entry level servers, chipsets may be limited to 4 GB of maximum memory. In today's dual processor Intel server chipsets and workstations, maximum system memory size can be upwards of 16 GB. BIOS Table locations, memory ranges, and IO space by definition must be less than 4 GB. In platforms populated with physical memory sizes approaching 4 GB and greater, onboard system resource requirements will likely not allow the operating system to take advantage of all physical memory populated due to PCI specification requirements and other memory mapped IO resources. Portions of physical memory may overlap with the memory space dedicated to other subsystems and become unavailable to the operating system. While some memory ranges required for addressing IO devices (PCI add in cards, system BIOS flash, IOxAPIC ranges, ACPI ranges) are unused, other ranges (integrated graphics shared memory and system management memory ranges) which the operating system does not report as available, are utilized by the platform.

Future generations of Intel chipset products and commercially available OS will allow more flexibility for system designers to use memory resources more efficiently.

Below is an example of memory use in a system with 4 Gbytes and 3 Gbytes of physical memory installed.

System Resource	Size	Physical Memory Remaining (4 GB Total System Memory)	Physical Memory Remaining (3 GB Total System Memory)
Firmware Hub flash memory (the BIOS)	*1 MB*	3.99	3.00
Local APIC	*4 KB*		
Area available to the chipset	*2 MB*		
IO APIC (4kBytes)	*4 KB*		
PCI Enumeration Area 1	*256 MB*	3.76	3.00
PCI Express* area (256 MB)	*256 MB*	3.51	3.00
PCI Enumeration Area 2 (if needed) (aligned on 256-MB boundary)	*512 MB*	3.01	3.00
VGA memory	*16 MB*	2.85	2.85
TSEG	*1 MB*	2.84	2.84
Memory available to BIOS, OS, and applications		**2.84**	**2.84**

In the above example, the size of the firmware hub (FWH) is set to 1 MB (or 8 Mbit). Below the FWH is the Local APIC area, the IO APIC, and other areas used by the chipset.

The chipset requires the PCI Express area to be aligned on a 256-MB boundary, and it is 256 MB in size. Therefore, the area was placed at 3.5 GB. For some BIOS' this may be a hard coded address.

The PCI enumeration algorithm may scan the buses to determine how much memory is needed for the different PCI devices in the system. If the amount of memory requested is less than 0.23 GB, the BIOS will assign that memory to PCI Enumeration Area 1. However, if the amount of memory required for PCI is larger than this size, the BIOS may allocate memory for PCI devices below the PCI Express area in PCI enumeration Area 2. This area is aligned on a 256-MB boundary and rounded up to the nearest multiple of 256 MB.

Note that the enumeration will ask for a contiguous memory area for all the devices below the host bridge. Therefore, if PCI needs greater than 256 MB of memory, the BIOS will put it at 3.0 GB – 3.5 GB, since 300 MB is rounded up to 512 MB. (For example, some video adapters will require 256 MB, or more, memory for graphics aperture)

If the system graphics device requires less than 0.23 GB, all the PCI devices should fit within PCI Enumeration Area 1. In that case the top amount of usable DRAM would move up to 3.5 GB increasing the amount of available memory by 512 MB.

While some memory ranges required for addressing IO devices (PCI add in cards, system BIOS flash, IOxAPIC ranges, ACPI ranges) are unused, other ranges (integrated graphics shared memory and system management memory ranges) which the operating system does not report as available, are utilized by the platform.

Note, that in the above example, the system BIOS reports the same amount of physical memory to the operating system in both scenarios even though one is populated with 4 GB of memory.

2.3 Microsoft Operating System Memory Support

The various Microsoft operating system releases and different versions of those releases support a varied maximum amount of physical memory. Below is a chart listing this maximum memory support:

Microsoft Operating System	Maximum Physical Memory Supported
Windows* 2000 Professional	4 GB
Windows 2000 Server	4 GB
Windows 2000 Advanced Server	8 GB
Windows 2000 Datacenter Server	32 GB
Windows XP Professional	4 GB
Windows XP 64 bit Edition	32 GB
Windows Server 2003 Web Edition	2 GB
Windows Server 2003 Standard Edition	4 GB
Windows Server 2003 Enterprise Edition	32 GB
Windows Server 2003 Datacenter Edition	64 GB
Windows Server 2003 Enterprise 64-bit Edition	64 GB
Windows Server 2003 Datacenter 64-bit Edition	512 GB

The memory support information, above, was gathered from a number of public Microsoft web pages. They are listed here:

http://www.microsoft.com/whdc/system/platform/server/PAE/PAEmem.mspx

http://www.microsoft.com/windowsserver2003/evaluation/features/compareeditions.mspx#2ram

http://www.microsoft.com/WindowsXP/64bit/evaluation/overviews/extended.mspx

§

3 Recommended System BIOS Updates and Documentation Changes

3.1 System BIOS Display of Available vs. Populated Memory

Many Intel chipset product specifications state support of 4 GB of physical memory. Since the chipsets comprehend 4 GB of memory in the system, then BIOS should indicate on the information screens and/or setup menu option the breakdown of memory sizes:

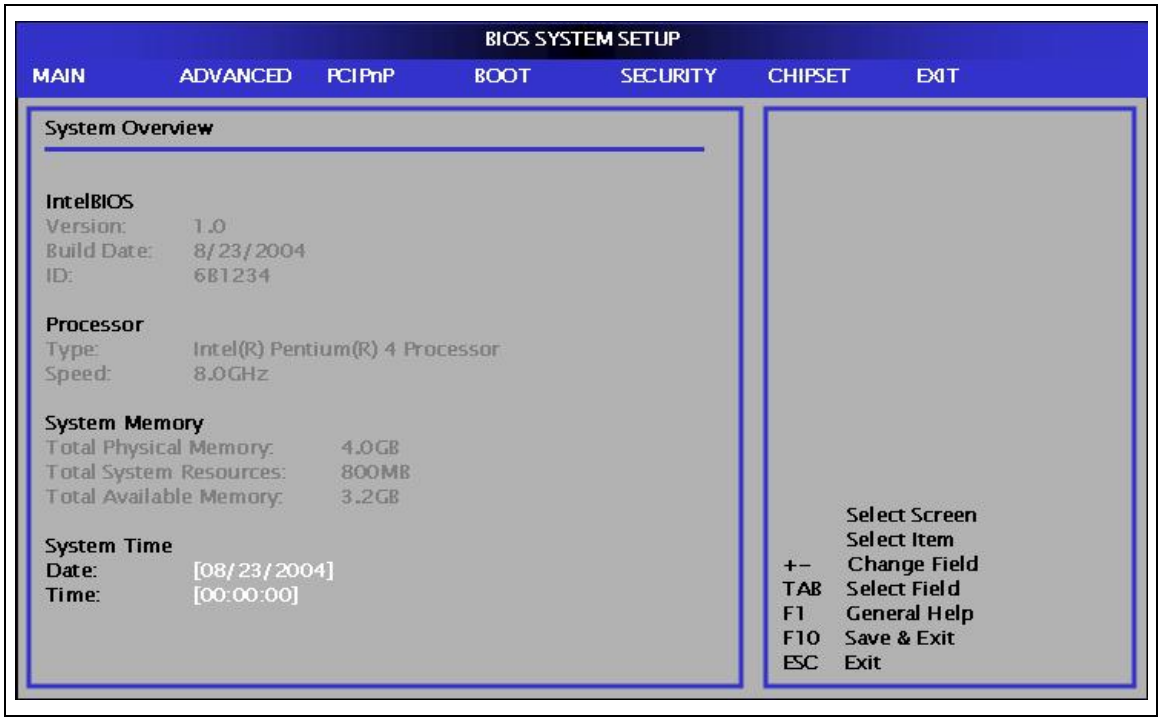
1. The amount of physical memory populated in the system;
2. The amount of populated memory dedicated to motherboard resources; and
3. The amount being reported as available to the operating system. This information should align to the INT15 E820h standard that BIOS uses to communicate memory usage to the operating system.

Where the sum of the amount of memory available to the operating system and the memory dedicated to system resources equals the total amount of physical memory on the system. By enumerating how populated memory is utilized and accounting for all physical memory populated, this BIOS feature will clarify the memory subsystem support and usage for the end-user.

3.2 Documenting Chipset Capabilities

Regarding 4 GB support with Intel chipsets that support up to (and not exceeding) 4 GB, Intel recommends a suitable memory space notification be included in all documentation.

Figure 1: Example BIOS Information Screen



§

4 Memory Remapping

Intel® enterprise chipsets have remapping capability to recover addressable memory space lost to MMIO space. Details on this remapping capability on Intel enterprise chipsets can be found in product-specific datasheets at www.intel.com .

In order to use remapping, the operating system must be able to address ranges higher than 4 GB of memory; enterprise operating systems have this capability. By comparison, today's mainstream desktop operating systems have limited address support. As mainstream client operating systems evolve to support greater than 4 GB of memory space, Intel will provide solutions as markets require in either the mobile, desktop, workstation, or low end server chipsets as well as offering similar remapping mechanisms to those included on Intel enterprise chipsets today. While this will improve performance on the system, and help ensure full physical memory usage, there will still be ranges of memory that will be dedicated to and actively used by other platform subsystems (such as integrated graphics), remaining unavailable to the operating system.

The following Memory Mapped IO devices and ranges are typically located just below 4 GB:

- High BIOS
- H-Seg
- XAPIC
- Local APIC
- FSB Interrupts
- EXPA0 through EXPC1 M, PM and BAR regions

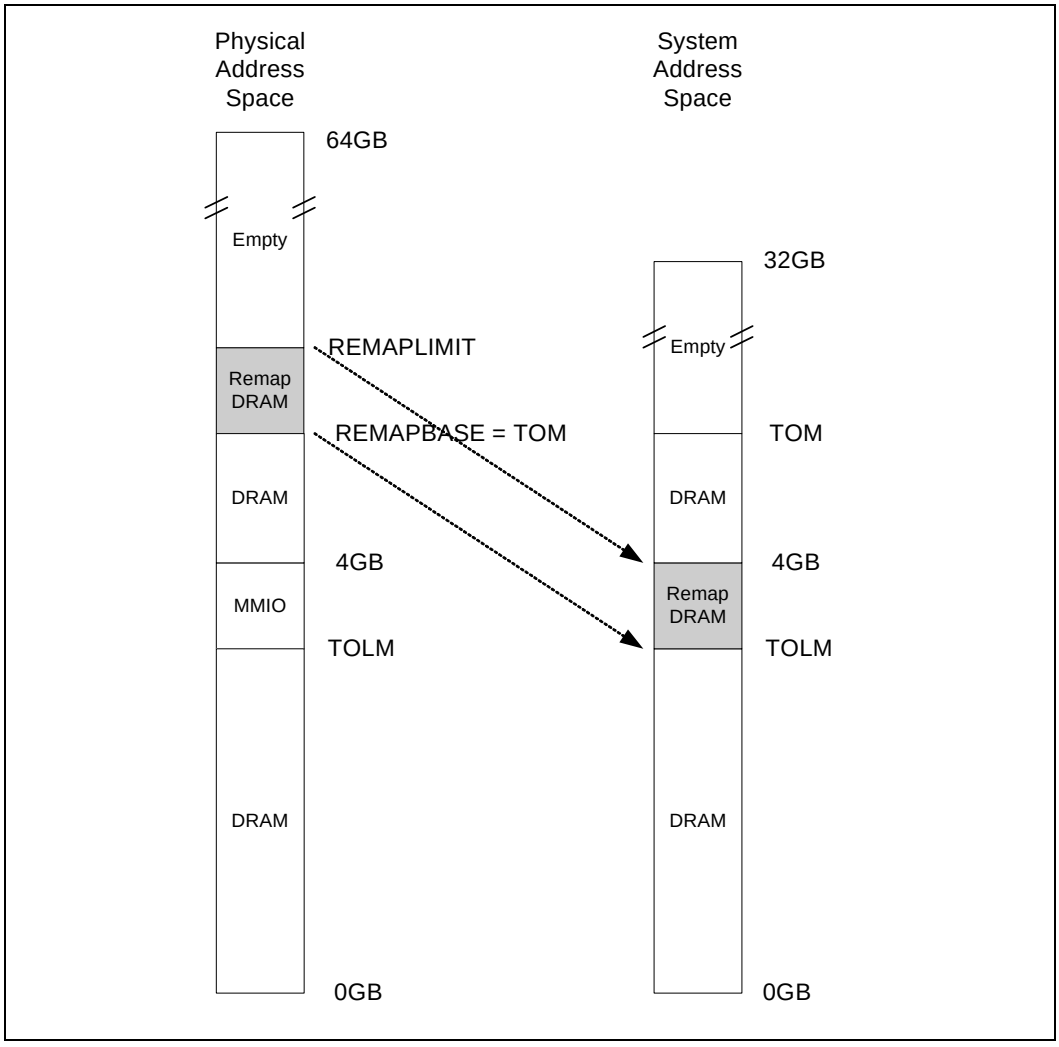
In uni-processor MCH architectures, the physical DRAM memory overlaps by the logical address space allocated to these Memory Mapped IO devices was unusable. In dual-processor MCH systems and potentially in upcoming uni-processor MCH based systems the memory allocated to memory mapped IO devices could easily exceed 1 GB. This creates the possibility of a large amount of physical memory populated in the system becoming usable.

4.1.1 Memory Re-mapping Algorithm

First we need to clarify some terminology. The address presented to the MCH is traditionally called a "physical address," because Intel architecture processors contain both segmentation and paging hardware, and all compatible software differentiates between logical addresses, virtual addresses, and physical addresses. The algorithm for re-mapping addresses presented to the MCH to reclaim DRAM address space must be implemented such that the mechanism is invisible to compatible software. This creates yet another type of address internal to the MCH, which is referred to in this document as a "system address." The system address applies to the internal MCH interface to physical DRAM memory, and is not directly visible to software, other than through certain internal logging registers used to store decoded DRAM address information for error isolation.

An incoming address (referred to as a physical address) is checked to see if it falls in the memory re-map window. The re-map window is defined by chipset registers which are programmed by the BIOS. An address that falls within this window is remapped to the system address by offset from the base of the remap window to the top of memory populated in the system.

4.1.2 Example: Greater than 4 GB of Physical Memory:



Example: 5 GB of system Memory, with 1 GB allocated to Memory Mapped IO

TOM = 5 GB TOLM = 3 GB

Amount of system Memory to be Remapped = 1 GB

REMAPBASE = 5 GB REMAPLIMIT = 6 GB-1

In this example, the amount of memory remapped is the range between **TOLM** and 4 GB. This system memory will be mapped to the physical address range defined between the **REMAPBASE** and **REMAPLIMIT**.

4.1.3 Interaction with Other Overlapping Address Space

The following Memory Mapped IO address spaces are all logically addressed below 4 GB where they do not overlap the physical address of the re-mapped memory region:

- H-Seg — At fixed address below 4 GB
- T-Seg — At (TOLM-TSEGSIZE) to TOLM
- High BIOS — Reset vector just under 4-GB boundary
- XAPIC — At fixed address below 4 GB
- Local APIC — At fixed address below 4 GB
- FSB Interrupts — At fixed address below 4 GB
- PCI memory spaces — At relocatable addresses below 4 GB and above TOLM
- PCI BARs — 32-bit BARs used by PCI and PCI Express devices – restricted below 4 GB

4.1.4 Implementation Notes

Remapping applies to transactions from all interfaces. All coherent inbound transactions are remapped by the host interface logic (necessary to ensure that the correct physical address is issued on the FSB for the snoop cycle). Non-coherent inbound transactions (hit non-coherent address window or non-snoop attribute set) get remapped within the inbound/outbound arbiter block.

Note: Accesses from PCI Express ports and from HI should be decoded to determine their type before they are remapped. For instance a PCI Express write to FEEx_xxxx is an interrupt transaction, but there is a physical PCI Express address (above 4 GB) that will be re-mapped to the system address of FEEx_xxxx. This is true of HighSMM and other ranges that lie between 4G and **TOLM**. In all cases the remapping of the address is done only after all other decodes have taken place, and a destination has been determined.

4.1.5 Unmapped Addresses between TOLM and 4 GB

Accesses that don't hit DRAM or PCI space should be treated the same as they are by other chipsets. This means that they are subtractively decoded to downstream PCI devices. Because the TOLM register is used to mark the upper boundary of DRAM space below the 4 GB boundary, no physical address between TOLM and 4 GB (inclusive of TOLM, but exclusive of 4 GB) ever decodes to main memory. Thus even if remapping is disabled, any address in this range has a non-memory destination.

All interfaces that generate or propagate traffic into the MCH must properly decode and route transactions based on their destination addresses.

§

5 Experimental Data

EXAMPLE DATA: For today's desktop platforms, such as the 915G/P and 925X Express chipset-based platforms, memory mapped IO (MMIO) and other system resources may utilize upwards of 512 MB of the 4 GB of linear addressable space. When the maximum physical memory supported is populated, MMIO will overlap with physical memory. Standard PCI Architecture limitations have prevented end users from taking advantage of all memory when 4 GB of memory populated.

The following E820h table was captured from a reference 915G Express chipset-based platform with Internal Graphics enabled. This table indicates to the operating system how much memory is available to the system.

Table 1. E820h Table from Reference Intel® 915G Express Chipset-Based Platform with Internal Graphics

Start	Length (bytes)	Size (bytes)	Status
0000 0000h	**9FC00h**	**654336**	**Available**
0009 FC00h	0400h	1024	Reserved
000E 0000h	020000h	131072	Reserved
0010 0000h	**CF530000h**	**3478323200**	**Available**
CF63 0000h	010000h	65536	Reserved
CF64 0000h	0C0000h	786432	Reserved
CF70 0000h	100000h	1048576	Reserved
CF80 0000h	800000h	8588608	Reserved
FEC0 0000h	1000h	4096	Reserved
FEE0 0000h	1000h	4096	Reserved

The operating system uses the available memory regions above to determine the available memory on the system. For example:

654336 Bytes + 347832300 Bytes = 3,478,977,536 Bytes = **3,397,439 KB** = **3.24 GB**

Note that in the examples below, the value of 3.24 GB is reported by the operating system.

The following data was collected on a reference 915G Express chipset-based platform with Microsoft Windows* XP with Service Pack 2. The configurations shown below are using 2 GB and 4 GB of total physical memory. In order to show differences in system resources, three different video solutions are shown.

Table 2. Intel® 915G Express Chipset-Based Platform with Microsoft Windows* XP and Service Pack 2

	2 GB (Int Gfx)	2 GB (PCI Gfx)	2 GB (PCIEx Gfx)	4 GB (Int Gfx)	4 GB (PCI Gfx)	4 GB (PCIEx Gfx)
MTRR: Total WB Memory	2 GB	2 GB	2 GB	4 GB	4 GB	4 GB
SysINFO: Total Physical Mem	2048.00 MB	2048.00 MB	2048.00 MB	4096.00 MB	4096.00 MB	4096.00 MB
SysINFO: Available Physical Mem	1.73 GB	1,75 GB	1.73 GB	2.99 GB	2.97 GB	3.01 GB
TaskMan: Total Physical Mem	2086636 KB	2086636 KB	2094828 KB	**3397356 KB[1]**	**3397356 KB[1]**	**3405548 KB[1]**
TaskMan: Available Mem	1832568 KB	1840935 KB	1826628 KB	3148444 KB	3122688 KB	3158026 KB
SysProp: Total RAM	1.99 GB	1.99 GB	2.00 GB	**3.24 GB[1]**	**3.24 GB[1]**	**3.25 GB[1]**

NOTES:
1. Corresponds to BIOS E820h table of available memory.

The following data was collected on a reference 915G chipset-based platform with Microsoft Windows* XP 64-Bit Edition. The configurations shown below are using 2 GB and 4 GB of total physical memory. In order to show differences in system resources, three different video solutions are shown.

Table 3. Intel® 915G Express Chipset-Based Platform with Microsoft Windows* XP 64-Bit Edition

	2 GB (Int Gfx)	2 GB (PCI Gfx)	2 GB (PCIEx Gfx)	4 GB (Int Gfx)	4 GB (PCI Gfx)	4 GB (PCIEx Gfx)
MTRR: Total WB Memory	2 GB	2 GB	2 GB	4 GB	4 GB	4 GB
SysINFO: Total Physical Mem	2048.00 MB	2048.00 MB	2048.00 MB	4096.00 MB	4096.00 MB	4096.00 GB
SysINFO: Available Physical Mem	1.77 GB	1.78 GB	1.77 GB	3.02 GB	3.02 GB	3.03 GB
TaskMan: Total Physical Mem	2086636 KB	2086636 KB	2094828 KB	**3397356 KB[1]**	**3397356 KB[1]**	**3405548KB[1]**
TaskMan: Available Mem	1874576 KB	1872544 KB	1875440 KB	3169716 KB	3179044 KB	3186524 KB
SysProp: Total RAM	1.99 GB	1.99 GB	2.00 GB	**3.24 GB[1]**	**3.24 GB[1]**	**3.25 GB[1]**

NOTES:
1. Corresponds to BIOS E820h table of available memory.

The following data was collected on a reference 915G Express chipset-based platform with RedHat* Enterprise Linux. The configurations shown below are using 2 GB and 4 GB of total physical memory.

Table 4. Intel® 915G Express Chipset-Based Platform with RedHat* Enterprise Linux

	2 GB (Int Gfx)	4 GB (Int Gfx)
FREE: Total Physical Mem	2051772 KB	**3343188 KB**[1]
FREE: Used Physical Mem	86176	101872
FREE: Free Physical Memory	1965596	3241316

NOTES:
1. Corresponds to BIOS E820h table of available memory.