Working Draft

X3T10.1/0991TR

Revision 1
May 23, 1996

Information Technology - PC Boot Considerations for Devices >8GB

Draft proposed American National Standard

This is a draft proposed Technical Report of Accredited Standards Committee X3. As such this is not a completed technical report. The X3T10 Technical Committee may modify this document as a result of comments received during public review and its approval as a technical report.

ABSTRACT
This technical report defines the problem related to address translation for disk drives greater than eight gigabytes in size and the operating system to boot considerations. A recommended solution is included.

Project Leader:
Duncan Penman
IIX Consulting
1045 Rubis Dr.
Sunnyvale, CA 94087

Voice: (408) 730-2565
FAX: (408) 730-5527
EMail penman@netcom.com

Technical Editor:
Lawrence Lamers
Adaptec
MS 293
691 South Milpitas Blvd
Milpitas CA, 95035
Voice: 408-957-7817
FAX: 408-957-7193
EMail ljlamers@aol.com

Reference number
ISO/IEC ***** : 199x
ANSI X3.*** - 199x
Printed 05/24/96

Other Points of Contact:

X3T10 Chairman	X3T10 Vice-Chair
John Lohmeyer	Lawrence Lamers
Symbios Logic Inc	Adaptec
4420 ArrowsWest Drive	MS 293
Colorado Springs, CO 80907-3444	691 South Milpitas Blvd
	Milpitas CA, 95035
Voice: 719-533-7560	408-957-7817
Fax: 719-533-7036	408-957-7193
Email: john.lohmeyer@symbios.com	ljlamers@aol.com

X3 Secretariat
Lynn Barra
Administrator Standards Processing
X3 Secretariat
1250 Eye Street, NW Suite 200
Washington, DC 20005

Voice: 202-626-5738
FAX: 202-638-4922
Email: lbarra@itic.nw.dc.us

Reflector

Internet address for subscription to the X3T10.1 reflector:	majordomo@symbios.com
Note should contain a line stating...	subscribe scsi your email address
Internet address for distribution via X3T10.1 reflector:	scsi@symbios.com
X3T10 Bulletin Board	719-533-7950
FTP Site:	ftp.symbios.com /pub/standards/io/x3t10
Web sites:	http://www.x3.org/x3t10 or http://www.symbios.com/x3t10 http://www.ssaia.org

Document Distribution
Global Engineering
15 Inverness Way East
Englewood, CO 80112-5704

Voice: 303-792-2181
or: 800-854-7179
FAX: 303-792-2192

PATENT STATEMENT

CAUTION: The developers of this technical report have requested that holders of patents that may be required for the implementation of the technical report, disclose such patents to the publisher. However, neither the developers nor the publisher have undertaken a patent search in order to identify which, if any, patents may apply to this technical report.

As of the date of publication of this technical report, following calls for the identification of patents that may be required for the implementation of the technical report, notice of one or more claims has been received.

By publication of this technical report, no position is taken with respect to the validity of this claim or of any rights in connection therewith. The known patent holder has, however, filed a statement of willingness to grant a license under these rights on reasonable and nondiscriminatory terms and conditions to applicants desiring to obtain such a license. Details may be obtained from the publisher.

No further patent search is conducted by the developer or the publisher in respect to any technical report it processes. No representation is made or implied that licenses are not required to avoid infringement in the use of this technical report.

ANSI®
X3.nnn-199n

American National Standard
for Information Systems -

PC Boot Considerations for Devices >8GB

Secretariat
Information Technology Industry Council

Approved Month dd, yy

American National Standards Institute, Inc.

Abstract

This technical report defines the problem related to address translation for disk drives greater than eight gigabytes in size and the operating system to boot considerations. A recommended solution is included.

X3's Technical Report Series

This Technical Report is one in a series produced by the American National Standards Committee, X3, Information Technology. The secretariat for X3 is held by the Computer and Business Equipment Manufacturers Association (CBEMA), 1250 Eye Street, NW Suite 200, Washington DC 20005.

As a by-product of the standards development process and the resources of knowledge devoted to it, X3 from time to time produces Technical Reports. Such Technical Reports are not standards, nor are they intended t o be used as such.

X3 Technical Reports are produced in some cases to disseminate the technical and logical concepts reflected in standards already published or under development. In other cases, they derive from studies in areas where it is found premature to develop a standard due to a still changing technology, or inappropriate to develop a rigorous standard due to the existence of a number of viable options, the choice of which depends on the user's particular requirements. These Technical Reports, thus, provide guidelines, the use of which can result in greater consistency and coherence of information processing systems.

When the draft Technical Report is completed, the Technical Committee approval process is the same as for a draft standard. Processing by X3 is also similar to that for a draft standard.

Published by
American National Standards Institute
11 W. 42nd Street, New York, New York 10036

Contents

Figures

Tables

Foreword (This foreword is not part of X3 Technical Report X3.nnn-199n.)

The PC has some historical firmware and software limitations on the size of disk that can be addressed. The major ones are:

528MB Affects IDE drives only. Caused by interaction of 16 heads maximum (limited by disk register size) and 1024 cylinders maximum (a limit at the OS-to-BIOS interface).

2GB Independent of drive type. A file system limit for any OS using a FAT based (DOS/Windows type) file system.

8GB Affects any drive accessed through the BIOS INT13H interface. This limit is a function of the number of bits available at the BIOS interface to address the drive.

In the past 2 years the PC community has painfully overcome the 528MB barrier. Two parallel solutions have emerged:

A technical report method for address translation within the BIOS which supports addressing up to 8GB on a disk.

Direct disk support built into the OS, such that it does not use the BIOS interface to access disk storage. Using this approach, the 8GB barrier does not exist.

The latter solution has been implemented in all recent PC operating systems and appears to be the technical report approach for the future.

Unfortunately, the OS must still be booted from a disk through the BIOS interface. Until the OS is in memory and initialized, there is no support for bypassing the BIOS. It is very desirable that the BIOS be able to address the full capacity of the boot disk, even if that extends beyond 8GB.

The rest of this document provides more background on this situation and recommends that the INT 13H extensions as documented in the Enhanced Disk Drive (EDD) 1.1 specification be adopted as the technical report for the BIOS interface used to access drives with >8GB capacity.

This technical report was developed by Task Group X3T10.1 of Accredited Standards Committee X3 during 1993-96. The technical reports approval process started in 1995. The informative annexes contained in this document are not part of the technical report.

Requests for interpretation, suggestions for improvement and addenda, or defect reports are welcome. They should be sent to the X3 Secretariat, Information Technology Industry Council, 1250 Eye Street, NW, Suite 200, Washington, DC 20005-3922.

This technical report was processed and approved for submittal to ANSI by Accredited Standards Committee on Information Processing Systems, X3. Committee approval of the technical report does not necessarily imply that all committee members voted for approval. At the time it approved this technical report, the X3 Committee had the following members:

James D. Converse, Chair

Donald C. Loughry, Vice-Chair

Joanne M. Flanagan, Secretary

Organization Represented	Name of Representative
American Nuclear Society	Geraldine C. Main
	Sally Hartzell (Alt.)
AMP, Inc	Edward Kelly
	Charles Brill (Alt.)
Apple Computer	Karen Higginbottom
Association of the Institute for Certification of Professionals (AICCP)	Kenneth Zemrowski
AT&T/NCR	Thomas W. Kern
	Thomas F. Frost (Alt.)
Boeing Company	Catherine Howells
	Andrea Vanosdoll (Alt.)
Bull HN Information Systems, Inc.	William George
Compaq Computer Corporation	James Barnes
Digital Equipment Corporation	Delbert Shoemaker
	Kevin Lewis (Alt.)
Eastman Kodak	James D. Converse
	Michael Nier (Alt.)
GUIDE International	Frank Kirshenbaum
	Harold Kuneke (Alt.)
Hewlett-Packard	Donald C. Loughry
Hitachi America, Ltd.	John Neumann
	Kei Yamashita (Alt.)
Hughes Aircraft Company	Harold L. Zebrack
IBM Corporation	Joel Urman
	Mary Anne Lawler (Alt.)
National Communication Systems	Dennis Bodson
National Institute of Standards and Technology	Robert E. Roundtree
	Michael Hogan (Alt.)
Northern Telecom, Inc	Mel Woinsky
	Subhash Patel (Alt.)
Neville & Associates	Carlton Neville
Recognition Technology Users Association	Herbert P. Schantz
	G. Edwin Hale (Alt.)
SHARE, Inc	Gary Ainsworth
	David Thewlis (Alt.)
Sony Corporation	Michael Deese
Storage Technology Corporation	Joseph S. Zajaczkowski
	Samuel D. Cheatham (Alt.)
Sun Microsystems	Scott Jameson
	Gary Robinson (Alt.)
3M Company	Eddie T. Morioka
	Paul D. Jahnke (Alt.)
Unisys Corporation	John L. Hill
	Stephen P. Oksala (Alt.)
U.S. Department of Defense	William C. Rinehuls
	C. J. Pasquariello (Alt.)
U.S. Department of Energy	Alton Cox
	Lawrence A. Wasson (Alt.)
U.S. General Services Administration	Douglas Arai
	Larry L. Jackson (Alt.)
Wintergreen Information Services	Joun Wheeler
Xerox Corporation	Dwight McBain
	Roy Peirce (Alt.)

Technical Committee X3T10 on I/O Interfaces, which reviewed this technical report, had the following members:

John B. Lohmeyer, Chair

Lawrence J. Lamers, Vice-Chair

Ralph Weber, Secretary

Mr. I. Dal Allan
Mr. Tak Asami
Mr. Geoffrey Barton
Mr. Paul Boulay
Mr. John Cannon
Mr. Joe Chen
Ms. Nancy Cheng
Mr. Roger Cummings
Mr. William Dallas
Mr. Jan V. Dedek
Mr. Mark Evans
Mr. Edward Fong
Mr. Raymond Gilson
Mr. Dave Guss
Mr. Kenneth J. Hallam
Mr. Norm Harris
Mr. Mike Hetzel
Mr. Peter Johansson
Mr. Skip Jones
Mr. Thomas J. Kulesza
Mr. Edward Lappin
Mr. Robert Liu
Mr. John Masiewicz
Mr. Pete McLean
Dr. Akira James Miura
Mr. Ian Morrell
Mr. Jay Neer
Mr. Vit Novak
Mr. Duncan Penman
Mr. Kevin R. Pokorney
Mr. Ron Roberts
Mr. Steve D. Schueler
Mr. Robert N. Snively
Mr. Arlan P. Stone
Mr. Matthew Thomas
Mr. Dennis Van Dalsen
Mr. Roger Wang
Mr. Bob Whiteman
Mr. Devon Worrell
Mr. Ruben Yomtoubian

Mr. Paul D. Aloisi
Mr. Akram Atallah
Mr. Robert Bellino
Mr. Charles Brill
Mr. Ting Li Chan
Mr. Andy Chen
Mr. Mike Chennery
Mr. Chris D'Iorio
Mr. Joe Dambach
Mr. Dhiru N. Desai
Mr. Timothy Feldman
Mr. Edward A. Gardner
Mr. Chuck Grant
Mr. Peter Haas
Dr. William Ham
Mr. Edward Haske
Mr. Stephen Holmstead
Mr. Gerry Johnsen
Mr. Richard Kalish
Mr. Lawrence J. Lamers
Mr. Pat LaVarre
Mr. John Lohmeyer
Mr. David McFadden
Mr. Patrick Mercer
Mr. E.J. Mondor
Mr. John Moy
Mr. Marc A. Noblitt
Mr. Erich Oetting
Mr. George Penokie
Mr. Gary Porter
Mr. Frank Samela
Mr. J. R. Sims
Mr. Jeff Stai
Mr. Clifford E. Strang Jr.
Mr. Pete Tobias
Mr. Joseph Wach
Mr. Gary M. Watson
Mr. Jeffrey L. Williams
Mr. Anthony Yang

Mr. David Andreatta
Mr. Wayne Baldwin
Mr. Rick Bohn
Mr. Michael Bryan
Mr. Shufan Chan
Mr. Jack Chen
Mr. Dan Colegrove
Mr. Zane Daggett
Mr. Brian N. Davis
Mr. Mike Eneboe
Mr. Stephen G. Finch
Mr. John Geldman
Mr. Louis Grantham
Mr. Douglas Hagerman
Mr. Tom Hanan
Mr. Stephen F. Heil
Mr. Gerald Houlder
Mr. Brian Johnson
Mr. Greg Kapraun
Mr. Dennis Lang
Mr. Jaff Lin
Mr. Bill Mable
Mr. James McGrath
Mr. Gene Milligan
Mr. Charles Monia
Mr. S. Nadershahi
Mr. Tim Norman
Mr. Dennis Pak
Mr. Doug Piper
Mr. Robert Reisch
Mr. John P. Scheible
Mr. Michael Smith
Mr. Gary R. Stephens
Mr. Nicos Syrimis
Mr. Tokuyuki Totani
Mr. Dean Wallace
Mr. Ralph O. Weber
Mr. Michael Wingard
Mr. Danny Yeung

Introduction

This technical report is divided into the following clauses and annexes.

Clause 1 defines the scope.

Clause 2 specifies the normative references.

Clause 3 defines the definitions, symbols and abbreviations.

Clause 4 contains the conventions.

Clause 5 defines the boot considerations for IDE devices.

Clause 6 defines the boot considerations for SCSI devices.

Clause 7 defines the 8GB barrier.

Clause 8 defines the extensions.

Clause 9 defines the recommendation.

Clause 10 defines the effect on existing standards.

X3 Technical Report for Information Technology -

PC Boot Considerations for Devices >8GB

1. Scope

This technical report defines the problem related to address translation for disk drives greater than eight gigabytes in size and the operating system to boot considerations. A recommended solution is included.

2. Normative references

None.

3. Definitions, symbols and abbreviations

3.1 Definitions

3.1.1. x: definition.

3.2 Symbols and Abbreviations

&	logical AND
=	assignment or comparison for EQUAL
≠	comparison for NOT EQUAL
<	comparison for LESS THAN
≤	comparison for LESS THAN OR EQUAL TO
>	comparison for GREATER THAN
+	ADD
-	SUBTRACT
*	MULTIPLY
±	PLUS OR MINUS
≈	APPROXIMATELY
»	MUCH GREATER THAN
GB	gigabyte
MB	megabyte

4. Conventions

Certain words and terms used in this technical report have a specific meaning beyond the normal English meaning. These words and terms are defined either in the glossary or in the text where they first appear. Lower case is used for words having the normal English meaning.

Fields containing only one bit are usually referred to as the "named" flag instead of the "named" field. When a bit is set its value is 1. When a bit is cleared, its value is 0.

Numbers that are not immediately followed by lower-case "b" or "h" are decimal values.

Numbers immediately followed by lower-case "b" (xxb) are binary values.

Numbers immediately followed by lower-case "h" (xxh) are hexadecimal values.

Decimal numbers are indicated with a comma(e.g., two and one half is represented as "2,5"). Decimal numbers having a value exceeding 999 are represented with a space (e.g., 24 255).

The bit ordering within a byte used in SSA-TL1 is illustrated in Table 1.

Table 1 - Bit ordering in a byte

Bit 7	6	5	4	3	2	1	Bit 0
msb							lsb

The bits in an encoded character are designated a b c d e i f g h j . Bit a is transmitted on the line first and the other bits follow in the order shown.

Reserved bits, fields, bytes, and code values are set aside for future standardization. Their use and interpretation may be specified by future extensions to this technical report. A reserved bit, field, or byte shall be set to zero, or in accordance with a future extension of this technical report.

Ignored bits, fields, or bytes shall be ignored by the device. Any value shall be considered valid.

Field names are shown as small capital letters, such as the FRAME TYPE field. Field values are shown as all capital letters, such as the FRAME TYPE field APPLICATION FRAME value. Variable names are in italics, such as N.

The byte ordering convention is Big Endian (i.e. the most significant byte of a number is sent first).

5. Booting From an IDE Drive

When power is applied to a PC, it begins executing self test and initialization code, usually called POST (Power On Self Test), from nonvolatile memory that is resident on the motherboard. Once the motherboard resources, including add-on I/O adapters, have been initialized, POST makes an INT 19H call to the BIOS boot routine.

Current BIOSes allow booting from the first floppy (the A drive) or the first hard disk (the C drive). The user is generally allowed to configure which drive, A or C, is tried first as a boot device. There is also a mechanism implemented in some BIOSes for booting from a CD-ROM.

The BIOS boot routine reads the Boot Sector from the hard disk. This contains a small loader program that brings in the first part of the operating system. For details on this portion of the process, see Hale Landis' writeups on disk partitions and OS2 and Windows Boot Sector formats. These can be retrieved by ftp access from fission.dt.wdc.com under /pub/otherdocs/.... When finished, the boot routine passes control to the loader program.

The loader uses the INT 13H interface to read a portion of the operating system into memory. That portion of the OS then takes control and brings in the rest of the load image, typically still using the INT 13H interface in this phase. Figure 1 illustrates this. It is in these two steps that the addressability of the boot disk through the BIOS is a concern, even though the operating system may bypass the BIOS later and access the disk directly. The reason for the concern is that it is undesirable to have to place restrictions on where the files referenced at boot time are placed on the hard disk.

Throughout this process, the drive is unaware that this is a boot operation. It simply sees a series of read commands.

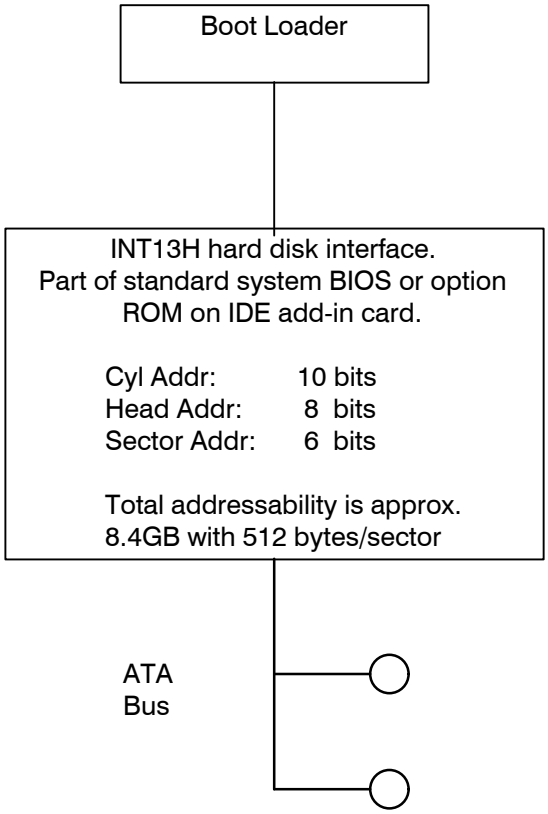

Figure 1 - Basic IDE Boot Interface

6. Booting From a SCSI Drive

If the C drive has a SCSI interface, rather than IDE, the process is identical except for one of the steps in POST.

During POST, a reserved range of memory addresses in the first megabyte of storage is scanned to see if any option ROMs are present. When an option ROM is found, control is temporarily passed to its code in order to allow for product specific initialization. In order for a SCSI drive to be the boot device, its host adapter must have an associated option ROM. One of its initialization functions is to replace the INT 13H entry address in the Interrupt Vector Table (IVT) with its own entry address. This is commonly called 'hooking INT 13'. Figure 2 illustrates this.

As a result of this initialization, any calls to the INT 13H interface are first directed to the SCSI code in the option ROM (shadowed into RAM storage). If the disk being addressed is a SCSI drive, the INT 13H request parameters are translated into a SCSI CDB and issued to the target drive. If the request is for an IDE drive, then the SCSI option code simply filters the request and passes control to the address it previously overlaid at the INT 13H location of the IVT. Using this mechanism, the boot process and any runtime software that uses the INT 13H interface can be indifferent to whether a drive is IDE or SCSI. It treats it as if it were IDE in either case.

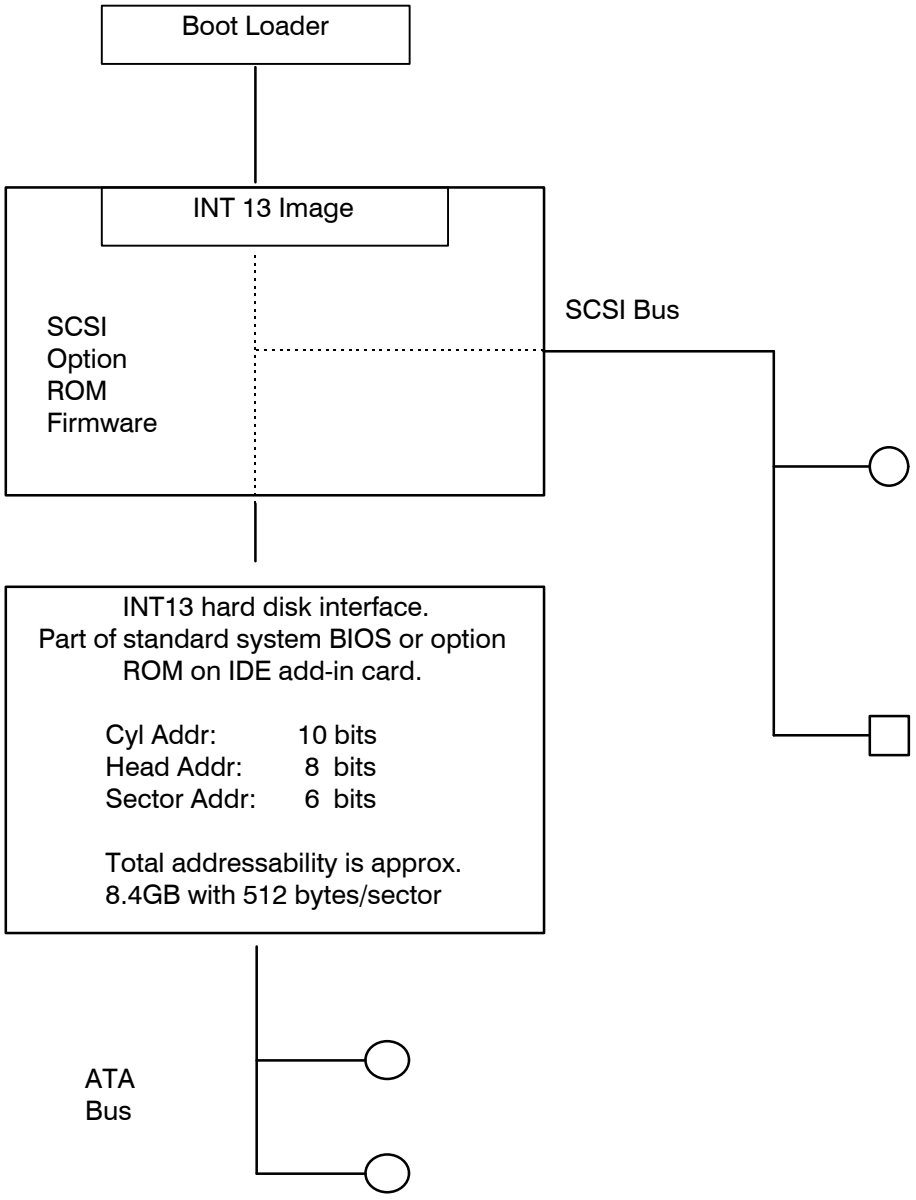

Figure 2 - SCSI Boot Interface

7. The 8GB Barrier

The INT 13H interface has only enough bits to address approximately 8.4GB of storage on a disk. This is referred to as the 8GB barrier. There are already SCSI disks with more capacity than that, and we can expect to see IDE drives exceeding that limit in the future.

In order to access the full capacity of a disk larger than this through the BIOS, the INT 13H interface must be enhanced.

8. INT 13H Extensions

In 1993 IBM and Microsoft developed a set of INT 13H extended functions that, among other things, provided a way to address the full capacity of disks bigger than 8GB.

In 1994 Phoenix Technologies published the Enhanced Disk Drive (EDD) BIOS interface specification, which included a more complete description of this interface than the original specification from IBM and Microsoft.

The basic INT 13H interface uses registers to pass request parameters to the BIOS. These are in the form of cylinder, head, and sector values for the data location on the disk. In contrast, the extended interface places the request parameters in a data structure in memory. A pointer to this structure is passed in a pair of registers to the BIOS.

The data structure for an extended INT 13H request contains a 64 bit field for a logical block address (LBA). This is more than adequate for the next several years, as SCSI is currently limited to a maximum LBA size of 32 bits and IDE to 28 bits. See Figure 3.

9. Recommendation For Boot Support

The recommended approach for supporting boot devices larger than 8GB is to implement the INT 13H extensions as published in EDD 1.1. For operation after the boot processing is complete, the OS can continue to use the extended INT 13H interface or can use a device driver that bypasses the BIOS.

10. Effect On Existing Standards

No changes are required in SCSI to support drives larger than 8GB until we reach a point where we want to use an LBA larger than 32 bits. For a block size of 512 bytes, this means a drive with a capacity greater than 2 terabytes. There are no current plans to expand SCSI LBAs beyond 32 bits, so this document does not affect any SCSI-3 standards.

ATA-4 needs to describe IDENTIFY DEVICE data for drives larger than 8GB. This should probably be an annex similar to the one in ATA-2 and ATA-3 describing IDENTIFY DEVICE data for drives less than 8GB.

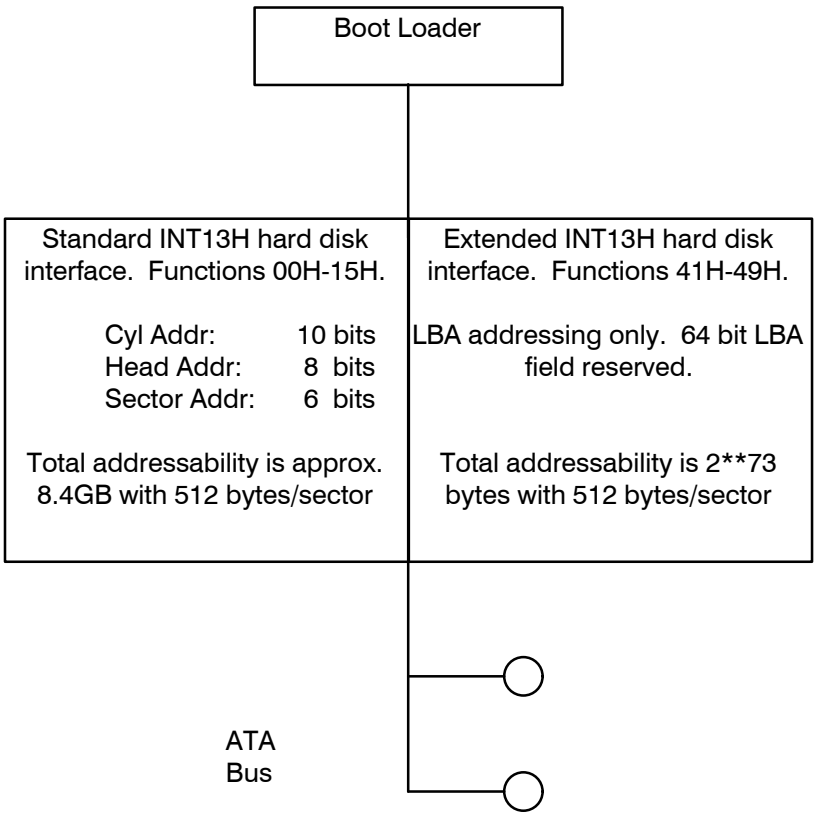

Figure 3 - INT13H Extensions Boot Interface