IEC 62842

Edition 1.0 2015-09

INTERNATIONAL STANDARD

NORME INTERNATIONALE

Multimedia home server systems – File allocation system with minimized reallocation

Systèmes de serveur domestique multimédias – Système d'allocation de fichiers avec réallocation réduite le plus possible

IEC 62842:2015-09(en-fr)

IEC Central Office
3, rue de Varembé
CH-1211 Geneva 20
Switzerland

Tel.: +41 22 919 02 11
info@iec.ch
www.iec.ch

About the IEC
The International Electrotechnical Commission (IEC) is the leading global organization that prepares and publishes International Standards for all electrical, electronic and related technologies.

About IEC publications
The technical content of IEC publications is kept under constant review by the IEC. Please make sure that you have the latest edition, a corrigendum or an amendment might have been published.

IEC publications search - webstore.iec.ch/advsearchform
The advanced search enables to find IEC publications by a variety of criteria (reference number, text, technical committee, …). It also gives information on projects, replaced and withdrawn publications.

IEC Just Published - webstore.iec.ch/justpublished
Stay up to date on all new IEC publications. Just Published details all new publications released. Available online and once a month by email.

IEC Customer Service Centre - webstore.iec.ch/csc
If you wish to give us your feedback on this publication or need further assistance, please contact the Customer Service Centre: sales@iec.ch.

IEC online collection - oc.iec.ch
Discover our powerful search engine and read freely all the publications previews. With a subscription you will always have access to up to date content tailored to your needs.

Electropedia - www.electropedia.org
The world's leading online dictionary on electrotechnology, containing more than 22 000 terminological entries in English and French, with equivalent terms in 18 additional languages. Also known as the International Electrotechnical Vocabulary (IEV) online.

A propos de l'IEC
La Commission Electrotechnique Internationale (IEC) est la première organisation mondiale qui élabore et publie des Normes internationales pour tout ce qui a trait à l'électricité, à l'électronique et aux technologies apparentées.

A propos des publications IEC
Le contenu technique des publications IEC est constamment revu. Veuillez vous assurer que vous possédez l'édition la plus récente, un corrigendum ou amendement peut avoir été publié.

Recherche de publications IEC - webstore.iec.ch/advsearchform
La recherche avancée permet de trouver des publications IEC en utilisant différents critères (numéro de référence, texte, comité d'études, …). Elle donne aussi des informations sur les projets et les publications remplacées ou retirées.

IEC Just Published - webstore.iec.ch/justpublished
Restez informé sur les nouvelles publications IEC. Just Published détaille les nouvelles publications parues. Disponible en ligne et une fois par mois par email.

Service Clients - webstore.iec.ch/csc
Si vous désirez nous donner des commentaires sur cette publication ou si vous avez des questions contactez-nous: sales@iec.ch.

IEC online collection - oc.iec.ch
Découvrez notre puissant moteur de recherche et consultez gratuitement tous les aperçus des publications. Avec un abonnement, vous aurez toujours accès à un contenu à jour adapté à vos besoins.

Electropedia - www.electropedia.org
Le premier dictionnaire d'électrotechnologie en ligne au monde, avec plus de 22 000 articles terminologiques en anglais et en français, ainsi que les termes équivalents dans 16 langues additionnelles. Egalement appelé Vocabulaire Electrotechnique International (IEV) en ligne.

IEC 62842

Edition 1.0 2015-09

INTERNATIONAL STANDARD

NORME INTERNATIONALE

Multimedia home server systems – File allocation system with minimized reallocation

Systèmes de serveur domestique multimédias – Système d'allocation de fichiers avec réallocation réduite le plus possible

INTERNATIONAL
ELECTROTECHNICAL
COMMISSION

COMMISSION
ELECTROTECHNIQUE
INTERNATIONALE

ICS 33.160.60 ISBN 978-2-8322-1029-9

Warning! Make sure that you obtained this publication from an authorized distributor.
Attention! Veuillez vous assurer que vous avez obtenu cette publication via un distributeur agréé.

CONTENTS

INTERNATIONAL ELECTROTECHNICAL COMMISSION

MULTIMEDIA HOME SERVER SYSTEMS – FILE ALLOCATION SYSTEM WITH MINIMIZED REALLOCATION

FOREWORD

1) The International Electrotechnical Commission (IEC) is a worldwide organization for standardization comprising all national electrotechnical committees (IEC National Committees). The object of IEC is to promote international co-operation on all questions concerning standardization in the electrical and electronic fields. To this end and in addition to other activities, IEC publishes International Standards, Technical Specifications, Technical Reports, Publicly Available Specifications (PAS) and Guides (hereafter referred to as "IEC Publication(s)"). Their preparation is entrusted to technical committees; any IEC National Committee interested in the subject dealt with may participate in this preparatory work. International, governmental and non-governmental organizations liaising with the IEC also participate in this preparation. IEC collaborates closely with the International Organization for Standardization (ISO) in accordance with conditions determined by agreement between the two organizations.

2) The formal decisions or agreements of IEC on technical matters express, as nearly as possible, an international consensus of opinion on the relevant subjects since each technical committee has representation from all interested IEC National Committees.

3) IEC Publications have the form of recommendations for international use and are accepted by IEC National Committees in that sense. While all reasonable efforts are made to ensure that the technical content of IEC Publications is accurate, IEC cannot be held responsible for the way in which they are used or for any misinterpretation by any end user.

4) In order to promote international uniformity, IEC National Committees undertake to apply IEC Publications transparently to the maximum extent possible in their national and regional publications. Any divergence between any IEC Publication and the corresponding national or regional publication shall be clearly indicated in the latter.

5) IEC itself does not provide any attestation of conformity. Independent certification bodies provide conformity assessment services and, in some areas, access to IEC marks of conformity. IEC is not responsible for any services carried out by independent certification bodies.

6) All users should ensure that they have the latest edition of this publication.

7) No liability shall attach to IEC or its directors, employees, servants or agents including individual experts and members of its technical committees and IEC National Committees for any personal injury, property damage or other damage of any nature whatsoever, whether direct or indirect, or for costs (including legal fees) and expenses arising out of the publication, use of, or reliance upon, this IEC Publication or any other IEC Publications.

8) Attention is drawn to the Normative references cited in this publication. Use of the referenced publications is indispensable for the correct application of this publication.

9) Attention is drawn to the possibility that some of the elements of this IEC Publication may be the subject of patent rights. IEC shall not be held responsible for identifying any or all such patent rights.

International Standard IEC 62842 has been prepared by technical area 8: Multimedia home systems and applications for end-user network of IEC technical committee 100: Audio, video and multimedia systems and equipment.

The text of this technical report is based on the following documents:

CDV	Report on voting
100/2367/CDV	100/2459/RVC

Full information on the voting for the approval of this standard can be found in the report on voting indicated in the above table.

This publication has been drafted in accordance with the ISO/IEC Directives, Part 2.

The committee has decided that the contents of this publication will remain unchanged until the stability date indicated on the IEC website under "http://webstore.iec.ch" in the data related to the specific publication. At this date, the publication will be

- reconfirmed,
- withdrawn,
- replaced by a revised edition, or
- amended.

INTRODUCTION

Recently, hard disk and Blu-ray Disc[1] recorders have become popular in the home to record television programmes. Normally a Hard Disk Recorder (HDR) is used for time shift and a Blu-ray Disc (BD) is used for library. When an HDR is used for time shift, television programmes are recorded and played, then many of them are deleted to reuse the spaces for other programmes to be recorded. These programmes are stored as files in a hard disk drive (HDD) using a file system. Continuous recording and deletion of programmes involves the continuous storing and deletion of files in the file system. Television programme streams include at least videos and an electronic programme guide (EPG). The HDR stores videos in a long, variable length file depending on the quality and recording hours. Compared with videos, EPG related information is stored in a shorter file or files but is often updated. This continuous creation, deletion and updating of files of different lengths finally causes the files to be stored in fragments, and the system performance becomes very low.

In a computer, defragmentation tools are provided to solve the problem of a fragmented file system. Normally defragmentation with reallocation of files in sequence takes a long time and the end user cannot but wait for the completion of the defragmentation, with no other activity. In the home server environment, a smarter solution to resolve this problem needs to be provided.

The recent newly developed HDD features will be reflected in the next version of the standard.

1 Blu-ray Disc™ is a trademark of the Blu-ray Disc Association. This information is given for the convenience of users of this document and does not constitute and endorsement by IEC of the product named.

MULTIMEDIA HOME SERVER SYSTEMS – FILE ALLOCATION SYSTEM WITH MINIMIZED REALLOCATION

1 Scope

This International Standard specifies the method for allocating requested file space with no fragmentation, to minimize the need for reallocation of fragmented files in the Universal Disc Format (UDF) file system applied to hard disk drives used in hard disk recorders.

2 Normative references

The following documents, in whole or in part, are normatively referenced in this document and are indispensable for its application. For dated references, only the edition cited applies. For undated references, the latest edition of the referenced document (including any amendments) applies.

ISO/IEC 13346 (all parts), *Information technology – Volume and file structure of write-once and rewritable media using non-sequential recording for information*

ISO/IEC 13346-1:1995, *Information technology – Volume and file structure of write-once and rewritable media using non-sequential recording for information interchange – Part 1: General*

ISO/IEC 13346-3:1999, *Information technology – Volume and file structure of write-once and rewritable media using non-sequential recording for information interchange – Part 3: Volume structure*

ISO/IEC 13346-4:1999, *Information technology – Volume and file structure of write-once and rewritable media using non-sequential recording for information interchange – Part 4: File structure*

OSTA UDF2.01:200, *Information technology – OSTA Universal Disk Format Specification, Revision 2.01*

Secure Universal Disk Format Specification Revision 1.00, *Optical Storage Technology Association (OSTA),* http://www.osta.org/

3 Terms, definitions, abbreviations and notation

3.1 Terms and definitions

For the purposes of this document, the following terms and definitions apply.

3.1.1
partition
region allocated to a file system by a disk volume space management system

3.1.2
virtual container partition
virtual partition containing a partition which has a minimum size of power-of-2 of allocation unit size of the disk

3.1.3
buddy
region allocation method where a given region having a power-of-2 unit size is recursively divided into two equal size regions (as 'buddies') until it reaches to one unit in size and, if a region of a given size is requested, the smallest free power-of-2 unit size region that can contain the requested size region is allocated

3.1.4
Concatenation of power of 2
CoPo2
basic allocation method of this standard

3.1.5
divide
process of obtaining divided-partitions through first identifying master-divided partitions, then applying the buddy method to them to get divided-partitions and finally allocating divided-partition numbers

3.1.6
master-dividing
process of identifying the master-divided-partition in the first process of dividing

Note 1 to entry: When the size of a partition is a power-of-2 unit size, the partition as a whole constitutes a master-divided partition.

Note 2 to entry: When the size of a partition is expressed as the sum of mutually different power-of-2 sizes with the power-of-2 size that constitutes the sum, constitute the partition as a concatenation of the areas in the sequence of those sizes as master-divided partitions in decreasing order of size.

3.1.7
master-divided-partition
divided partition identified by master-dividing

3.1.8
one unit in size
predetermined minimum unit of memory size that can be allocated

3.1.9
divided-partition
partition identified by dividing the master-divided partitions in a given partition recursively in half until it reaches one unit in size

3.1.10
divided-partition level
level value of a divided-partition expressed by the power-of-2 value of the divided partition compared to unit size

Note 1 to entry: This is often abbreviated as level n (where n specifies level number).

3.1.11
partition level
level n
abbreviation of divided-partition level, normally used in a simple form, 'level n'

3.1.12
divided-partition pair
pair of divided-partition formed by dividing a divided-partition into two halves as buddies

3.1.13
divided-partition number
number assigned to divided-partitions in the process of dividing a given partition from top level to the lowest level incrementally

Note 1 to entry: Numbers are assigned to each divided-partition from top to bottom and from left to right in the same level with the starting number 1, counting up by one.

3.1.14
master-divided-partition number
divided-partition number assigned to each master-divided-partition

3.1.15
master-divided partitions table
table for managing the master-divided partitions

3.1.16
master-divided-partition number management table
table for managing the master-divided-partition number in contrast with each divided-partition level

3.1.17
end position management table
table for managing the maximum divided-partition number assigned in each divided-partition level

3.1.18
highest divided-partition level management table
table for managing the highest divided-partition level for the partition

3.1.19
multilevel-divided partition management tables
three tables for managing the divided-partitions

Note 1 to entry: The three tables are the master divided-partition number management table, the end position management table and the highest divided-partition level management table.

3.1.20
multilevel-divided-partition allocation table
table for managing the multilevel-divided partitions state in each level with 2 bits

3.1.21
segment
region to be taken or allocated from a partition

Note 1 to entry: If the size of a region is power-of-2 unit size, the region consists of one divided partition.

Note 2 to entry: If the size of a region is a sum of polynomials of power-of-2 of a unit size, the region consists of a concatenation of multilevel divided-partitions, each corresponding to one polynomial component, and it is taken from the container divided-partition for a segment.

3.1.22
multilevel-dividing
procedure for obtaining a region as a segment, consisting of the concatenation of multilevel divided-partition decreasing order in size, of which each corresponds to one power-of-2 polynomial component

3.1.23
container divided partition for a segment
simple power-of-2 region determined by rounding up the size of a polynomial-based requested segment to the size of the nearest simple power-of-2 region and that simple region will consist of the segment to be allocated and adjacent-multilevel-segment

3.1.24
adjacent-multilevel-segment
segment adjacent to the allocated multilevel-segment of which each component divided-partition is allocated in increasing order of size, and which constitutes the rest of the container divided-partition for a segment

3.1.25
first-pass-allocation
first step to allocate a segment, to obtain a minimum divided partition including the requested segment

Note 1 to entry: If a segment is a single level divided-partition, the allocation is completed, but if a segment is multilevel, the process continues to a second-pass allocation.

3.1.26
second-pass-allocation
second step to allocate a multilevel segment, in a segment allocation, where the multilevel-dividing is applied to the divided-partition obtained from the first-pass-allocation

3.1.27
provisional-allocation
procedure to get a segment, where, if a divided-partition including the target segment is not found, a search for the available upper level of a divided-partition is requested and the divided-partition found is allocated as provisional allocation

Note 1 to entry: After getting a provisional divided-partition, first-pass-allocation continues and in the case of allocating a multilevel-segment, second-pass-allocation continues.

3.1.28
allocation state
divided-partition state in view of allocation, classified into available, occupied and reserved

3.1.29
available
allocation state showing the divided-partition is free to use, meaning available

Note 1 to entry: Divided-partitions are classified into first-pass-available and second-pass-available. When searching for a free divided-partition, start by searching for first-pass-available and then search for second-pass-available.

3.1.30
first-pass-available
initial available state assigned to free divided-partitions

Note 1 to entry: When an occupied divided-partition becomes free, its allocation state is set as first-pass-available.

3.1.31
second-pass-available
available allocation state assigned to each component divided-partition of adjacent-multilevel-segments, after allocation of the multilevel-segment

3.1.32
occupied
allocation state showing the divided-partition is in use

3.1.33
reserved
divided-partition allocation state that is not appropriate to classify into available or occupied

Note 1 to entry: Divided-partitions other than available or occupied are set as reserved. When getting a segment through provisional-allocation, the allocated top divided-partition and the divided-partitions in lower levels shall be set as reserved.

3.1.34
master boot record
conventional partitioning record in logical block address 0 of HDD

3.1.35
logical block address
address scheme used for HDD volume space

3.1.36
logical sector number
address scheme used for UDF volume in HDD volume space

Note 1 to entry: UDF volume might be restricted in the first partition of HDD. LSN = LBA − 63.

3.1.37
logical block number
address scheme used for UDF partition in UDF volume space

Note 1 to entry: LBN = LSN − 257.

3.2 Abbreviations

HDD	Hard Disk Drive
MBR	Master Boot Record
LBA	Logical Block Address
LSN	Logical Sector Number
LBN	Logical Block Number
OS	Operating System
UDF	Universal Disk Format

3.3 Notation

Universal disk format (UDF) is an application of the ISO/IEC 13346 series and the specification of the data structure always has a reference to the corresponding part of ISO/IEC 13346 in the form of “ISO/IEC-x.a.b.c” where x is a part number and a.b.c is the paragraph number. In Clause 9, data structure is described in C-like form. The fields in the form that need explanation are written in bold, and the explanation is provided just after the structure.

These standard-specific compound words are expressed by words concatenation with hyphen.

4 Precondition and the policy

4.1 Preconditions

When applying UDF on HDD, the preconditions considered are as follows:

a) a disk volume is configured with partitions managed by the partition manager for the HDD system. Each partition is a contiguous region;

b) a partition is a contiguous region which is an integer multiple of a cluster, which is an integer multiple of a sector.

4.2 Policy

This standard applies the allocation method which tries to allocate the larger free space as far as possible. Using this method of allocation, the possibilities of fragmentation and reallocation are minimized.

The buddy system can allocate the minimum contiguous power-of-2 cluster region which contains the requested size. The region of the clusters left over the requested size, which are waste, are called slack. The slack size becomes larger and larger, when the requested size become larger. This is the reason why the buddy system only applies to memory management systems in the small system and does not apply to the file system.

The CoPo2 (concatenation of power-of-2) resolves the problem of buddy, which helps to use the slack, and can be applied to the file system which minimizes the possibilities of fragmentation and reallocation.

5 Method to be applied-CoPo2

In buddy, a region expressed in binary 1000 in cluster size, a possible divided partition that could be allocated can be expressed as a tree of divided partitions from size 1000 (level 3) to 0001 (level 0):

Divided partitions	Divided-partition level
(#0,#1)	level 3 (size=1000)
(#0,#2) (#1,#3)	level 2 (size=0100)
(#0,#4) (#1,#5) (#2,#6) (#3,#7)	level 1 (size=0010)
(#0,#8) (#1,#9) (#2,#10) (#3,#11) (#4,#12) (#5,#13) (#6,#14) (#7,#15)	level 0 (size=0001)

In the above tree, each divided partition is expressed as “(#level-internal-divided-partition-number, #divided-partition-number-in-tree)”. In other words, in two parenthesis, the hash sign expresses a sequential number, followed by the divided partition number followed by a comma to separate the following expression followed by a hash sign to express a sequential number followed by the divided-partition number in the total tree.

To obtain a 0101 size segment in the tree, a solution is getting divided partition number in tree #2 and #12. (in the following, #n expresses divided-partition-number-in-tree). In buddy this solution is not available, instead the minimum size of divided-partition that contains #2 and #12 is obtained, which is #1. In reality #2 and #12 are used in #1, and the #13 and #7 are left and become slack and could not be used.

In buddy, a cluster that is the multiple of a sector is used as the basic allocation unit. To manage the cluster allocation state in buddy, one bit is used to express 'occupied' or 'available'. With this method, the slack cannot be managed. To manage slack, each divided-partition state requires two bits for state management.

The key states of a divided-partition can be 'occupied' or 'available'. Initially, only #1 is available where a space can be allocated. The other divided-partitions belong to the #1 and cannot be allocated. But when a size 0101 segment is allocated, #1 gives control for space management to the lower layer divided-partitions and #2 and #12 become occupied and concatenated together to form a segment of size 0101. As a result, the states of the upper and lower layer divided-partitions change. In this process #13 and #7, that is a slack in buddy, form an adjacent-multilevel-segment and become available space in the adjacent area of the allocated segment.

Considering the practicality of optimum usage of 2 bits, divided-partitions are classified as follows:

a) available(0x), where space can be allocated, is classified in two states as available1(00) and available2(01) that is normally a slack area in buddy but available in this system.
b) occupied(1x), where space cannot be allocated, is classified in two states as in-use(11) and reserved(10). In-use is used for allocated divided-partitions. Reserved is used for the divided-partitions which depend on the upper or lower layer of a divided-partition in allocations and are reserved for allocation decision.

In summary, the state states shrink into four, as follows:

– available(0x)
    - available1(00)-available in normal;
    - avaiable2(01)-become available in adjacent-multilevel-segment;
– occupied(1x)
    - reserved(10);
    - in-use(11).

Using those states explains the state change using a simple partition:

The initial state of a partition in tree form is as follows:

(#0,#1;00)								level 3 (size=1000)
(#0,#2;10)				(#1,#3;10)				level 2 (size=0100)
(#0,#4;10)		(#1,#5;10)		(#2,#6;10)		(#3,#7;10)		level 1 (size=0010)
(#0,#8;10)	(#1,#9;10)	(#2,#10;10)	(#3,#11;10)	(#4,#12;10)	(#5,#13;10)	(#6,#14;10)	(#7,#15;10)	level 0 (size=0001)

only #1 is available1 (00) and the others are all reserved(10). If a size 0101 segment is allocated, the states change as follows:

– #1 changes to reserved(10);
– #2 changes to in-use(11) and no change in the lower level divided-partitions in the group;
– #3 has no change, but
– #12 changes to in-use(11);
– #7 and #13 change to available2;
– #14 and #15 have no change.