DVJ Perspective on: Timing and synchronization for time-sensitive applications in bridges local area networks

Draft 0.201

Contributors:
See page xx.

Abstract: This working paper provides background and introduces possible higher level concepts for the development of Audio/Video bridges (AVB).
Keywords: audio, visual, bridge, Ethernet, time-sensitive

Editors' Foreword

Comments on this draft are encouraged. **PLEASE NOTE: All issues related to IEEE standards presentation style, formatting, spelling, etc. should be addressed, as their presence can often obfuscate relevant technical details.**

By fixing these errors in early drafts, readers wll be able to devote their valuable time and energy to comments that materially affect either the technical content of the document or the clarity of that technical content. Comments should not simply state what is wrong, but also what might be done to fix the problem.

Information on 802.1 activities, working papers, and email distribution lists etc. can be found on the 802.1 Website:

http://ieee802.org/1/

Use of the email distribution list is not presently restricted to 802.1 members, and the working group has had a policy of considering ballot comments from all who are interested and willing to contribute to the development of the draft. Individuals not attending meetings have helped to identify sources of misunderstanding and ambiguity in past projects. Non-members are advised that the email lists exist primarily to allow the members of the working group to develop standards, and are not a general forum.

Comments on this document may be sent to the 802.1 email reflector, to the editors, or to the Chairs of the 802.1 Working Group and Interworking Task Group.

This draft was prepared by:

David V James
JGG
3180 South Court
Palo Alto, CA 94306
+1.650.494.0926 (Tel)
+1.650.954.6906 (Mobile)
Email: dvj@alum.mit.edu

Chairs of the 802.1 Working Group and Audio/Video Bridging Task Group:.

Michael Johas Teener
Chair, 802.1 Audio/Video Bridging Task
Broadcom Corporation
3151 Zanker Road
San Jose, CA
95134-1933
USA
+1 408 922 7542 (Tel)
+1 831 247 9666 (Mobile)
Email:mikejt@broadcom.com

Tony Jeffree
Group Chair, 802.1 Working Group
11A Poplar Grove
Sale
Cheshire
M33 3AX
UK
+44 161 973 4278 (Tel)
+44 161 973 6534 (Fax)
Email: tony@jeffree.co.uk

Introduction to IEEE Std 802.1AS™

(This introduction is not part of P802.1AS, IEEE Standard for Local and metropolitan area networks—Timing and synchronization for time-sensitive applications in bridged local area networks.)

This standard specifies the protocol and procedures used to ensure that the synchronization requirements are met for time sensitive applications, such as audio and video, across bridged and virtual bridged local area networks consisting of LAN media where the transmission delays are fixed and symmetrical; for example,IEEE 802.3 full duplex links. This includes the maintenance of synchronized time during normal operation and following addition, removal, or failure of network components and network reconfiguration. The design is based on concepts developed within the IEEE Std 1588, and is applicable in the context of IEEE Stds 802.1D and 802.1Q.

Synchronization to an externally provided timing signal (e.g., a recognized timing standard such as UTC or TAI) is not part of this standard but is not precluded.

Contributors

This working paper is based on contributions or review comments from the people listed below. Their listing doesn't necessarily imply they agree with the entire content or the author's interpretation of their input.

Jim Battaglia	Pioneer
Alexei Beliaev	Gibson
Dirceu Cavendish	NEC Labs America
George Claseman	Micrel
Feifei (Felix) Feng	Samsung Electronics
John Nels Fuller	Independent
Geoffrey M. Garner	Samsung Electronics
Kevin Gross	Cirrus Logic
Jim Haagen-Smit	HP
David V James	JGG
Dennis Lou	Pioneer
Michael D. Johas Teener	Broadcom
Fred Tuck	EchoStar

Version history

Version	Date	Author	Comments
0.082	2005Apr28	DVJ	Updates based on 2005Apr27 meeting discussions
0.085	2005May11	DVJ	– Updated front-page list of contributors – Updated book for continuous pages (Clause 1 discontinuity fixed) – Miscellaneous editing fixes
0.088	2005Jun03	DVJ	– Application latency scenarios clarified.
0.090	2005Jun06	DVJ	– Misc editorials in bursting and bunching annex.
0.092	2005Jun10	DVJ	– Extensive cleanup of Clause 5 subscription protocols, based on 2005Jun08 teleconference review comments.
0.121	2005Jun24	DVJ	– Extensive cleanup of clock-synchronization protocols, base on 2005Jun22 teleconference review comments.
0.127	2005Jul04	DVJ	– Pacing descriptions greatly enhanced.
0.200	2007Jan23	DVJ	Removal of non time-sync related information. Update based on recent teleconference suggestion (layering), as well as input available from others' drafts.
—	TBD	—	—

Formats

In many cases, readers may elect to provide contributions in the form of exact text replacements and/or additions. To simplify document maintenance, contributors are requested to use the standard formats and provide checklist reviews before submission. Relevant URLs are listed below:

General:	http://grouper.ieee.org/groups/msc/WordProcessors.html
Templates:	http://grouper.ieee.org/groups/msc/TemplateTools/FrameMaker/
Checklist:	http://grouper.ieee.org/groups/msc/TemplateTools/Checks2004Oct18.pdf

Topics for discussion

Readers are encouraged to provide feedback in all areas, although only the following areas have been identified as specific areas of concern.

a) Layering. Should be reviewed.

TBDs

Further definitions are needed in the following areas:

a) Details of the client time-sync services should be defined.

b) Details of the Ethernet per-port time-sync services should be defined.

c) How are leap-seconds handled?

d) How are rate differences distributed? Avoid whiplash?

e) When the grand-master changes, should the new clock transitioin to it free-run rate instantaneously or migrate there slowly over time?

Contents

List of figures 8

List of tables 9

1. Overview 11

1.1 Scope 11
1.2 Purpose 11
1.3 Introduction 11

2. References 13

3. Terms, definitions, and notation 14

3.1 Conformance levels 14
3.2 Terms and definitions 14
3.3 Service definition method and notation 15
3.4 State machines 16
3.5 Arithmetic and logical operators 19
3.6 Numerical representation 19
3.7 Field notations 20
3.8 Bit numbering and ordering 21
3.9 Byte sequential formats 22
3.10 Ordering of multibyte fields 22
3.11 MAC address formats 23
3.12 Informative notes 24
3.13 Conventions for C code used in state machines 24

4. Abbreviations and acronyms 25

5. Architecture overview 27

5.1 Application scenarios 27
5.2 Design methodology 28
5.3 Time-synchronization facilities 29
5.4 Client processing details 32
5.5 Key distinctions from IEEE Std 1588 33

6. Medium access control (MAC) service and reference model 35

6.1 Overview 35
6.2 Overview of MAC services 35
6.3 MAC services to the client layer 35

7. Ethernet duplex-cable time synchronization 38

7.1 Design methodology 38
7.2 Time-synchronization operation 39
7.3 timeSync frame format 42

Annex A (informative) Bibliography 47

Annex B (informative) Bridging to IEEE Std 1394 48

B.1 Hybrid network topologies 48

Annex C (informative) Review of possible alternatives 50

C.1 Clock-synchronization alternatives 50

Annex D (informative) Time-of-day format considerations 52

D.1 Possible time-of-day formats 52
D.2 Time format comparisons 54

Index 55

Annex E (informative) C-code illustrations 59

List of figures

Figure 1.1—Topology and connectivity 12
Figure 3.1—Service definitions 15
Figure 3.2—Bit numbering and ordering 21
Figure 3.3—Byte sequential field format illustrations 22
Figure 3.4—Multibyte field illustrations 22
Figure 3.5—Illustration of fairness-frame structure 23
Figure 3.6—MAC address format 23
Figure 3.7—48-bit MAC address format 24
Figure 5.1—Garage jam session 27
Figure 5.2—Possible looping topology 28
Figure 5.3—Timing information flows 29
Figure 5.4—Grand-master precedence flows 30
Figure 5.5—Grand-master precedence 30
Figure 5.6—Hierarchical flows 31
Figure 5.7 —AVB service interface model 32
Figure 6.1 —AVB service and reference model relationship to the ISO/IEC OSI reference model 35
Figure 7.1—Timer snapshot locations 39
Figure 7.2—Timer snapshot locations 40
Figure 7.3—Rate-adjustment effects 41
Figure 7.4—timeSync frame format 42
Figure 7.5—*typeCount* format 43
Figure 7.6—*gmSelection* format 44
Figure 7.7—*uniqueID* format 44
Figure 7.8—Complete seconds timer format 45
Figure 7.9—Complete seconds timer format 45
Figure B.1—IEEE 1394 leaf domains 48
Figure B.2—IEEE 802.3 leaf domains 48
Figure B.3—Time-of-day format conversions 49
Figure B.4—Grand-master precedence mapping 49
Figure 4.1—Complete seconds timer format 52
Figure D.2—IEEE 1394 timer format 52
Figure D.3—IEEE 1588 timer format 53
Figure D.4—EPON timer format 53
Figure D.5—Compact seconds timer format 53
Figure D.6—Nanosecond timer format 53

List of tables

Table 3.1—State table notation example 17
Table 3.2—Called state table notation example 18
Table 3.3—Special symbols and operators 19
Table 3.4—Names of fields and sub-fields 20
Table 3.5—*wrap* field values 21
Table 7.1—Clock-synchronization intervals 42
Table C.1Protocol comparison 50
Table D.1—Time format comparison 54

DVJ Perspective on: Timing and synchronization for time-sensitive applications in bridges local area networks

1. Overview

1.1 Scope

This draft specifies the protocol and procedures used to ensure that the synchronization requirements are met for time sensitive applications, such as audio and video, across bridged and virtual bridged local area neworks consisting of LAN media where the transmission delays are fixed and symmetrical; for example, IEEE 802.3 full duplex links. This includes the maintenance of synchronized time during normal operation and following addition, removal, or failure of network components and network reconfiguration. It specifies the use of IEEE 1588 specifications where applicable in the context of IEEE Stds 802.1D and 802.1Q. Synchronization to an externally provided timing signal (e.g., a recognized timing standard such as UTC or TAI) is not part of this standard but is not precluded.

1.2 Purpose

This draft enables stations attached to bridged LANs to meet the respective jitter, wander, and time synchronization requirements for time-sensitive applications. This includes applications that involve multiple streams delivered to multiple endpoints. To facilitate the widespread use of bridged LANs for these applications, synchronization information is one of the components needed at each network element where time-sensitive application data are mapped or demapped or a time sensitive function is performed. This standard leverages the work of the IEEE 1588 WG by developing the additional specifications needed to address these requirements.

1.3 Introduction

1.3.1 Background

Ethernet has successfully propagated from the data center to the home, becoming the wired home computer interconnect of choice. However, insufficient support of real-time services has limited Ethernet's success as a consumer audio-video interconnects, where IEEE Std 1394 Serial Bus and Universal Serial Bus (USB) have dominated the marketplace. Success in this arena requires solutions to multiple topics:

a) Discovery. A controller discovers the proper devices and related streamID/bandwidth parameters to allow the listener to subscribe to the desired talker-sourced stream.

b) Subscription. The controller commands the listener to establish a path from the talker.
Subscription may pass or fail, based on availability of routing-table and link-bandwidth resources.

c) Synchronization. The distributed clocks in talkers and listeners are accurately synchronized.
Synchronized clocks avoid cycle slips and playback-phase distortions.

d) Pacing. The transmitted classA traffic is paced to avoid other classA traffic disruptions.

This draft covers the “Synchronization” component, assuming solutions for the other topics will be developed within other drafts or forums.

1.3.2 Interoperability

AVB time synchronization interoperates with existing Ethernet, but the scope of time-synchronization is limited to the AVB cloud, as illustrated in Figure 1.1; less-precise time-synchronization services are available everywhere else. The scope of the AVB cloud is limited by a non-AVB capable bridge or a half-duplex link, neither of which can support AVB services.

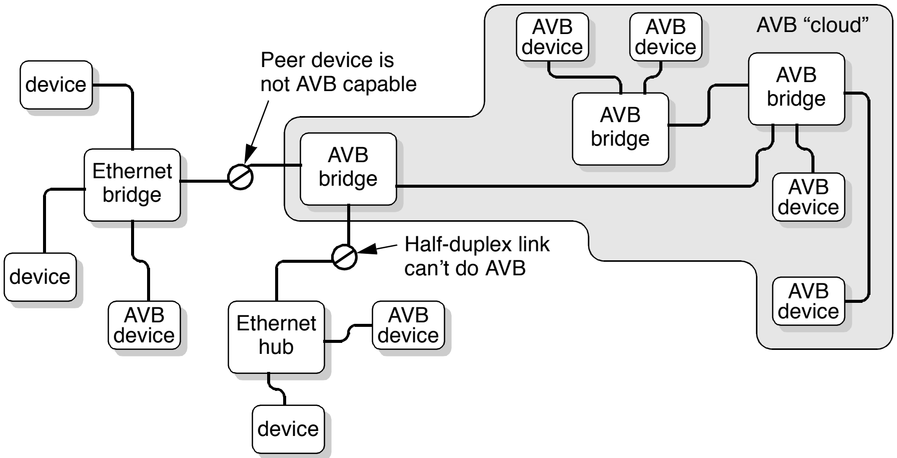


Figure 1.1—Topology and connectivity

Separation of AVB devices is driven by the requirements of AVB bridges to support subscription (bandwidth allocation) and pacing of time-sensitive transmissions, as well as time-of-day clock-synchronization.

1.3.3 Document structure

The clauses and annexes of this working paper are listed below.

- Clause 1: Overview
- Clause 2: References
- Clause 3: Terms, definitions, and notation
- Clause 4: Abbreviations and acronyms
- Clause 5: Architecture overview
- Clause xx: FrameFormats
- Clause xx: xx
- Annex A: Bibliography
- Annex B: Bridging to IEEE Std 1394
- Annex C: Review of possible alternatives
- Annex D: Time-of-day format considerations
- Annex E: C-code illustrations

2. References

The following documents contain provisions that, through reference in this working paper, constitute provisions of this working paper. All the standards listed are normative references. Informative references are given in Annex A. At the time of publication, the editions indicated were valid. All standards are subject to revision, and parties to agreements based on this working paper are encouraged to investigate the possibility of applying the most recent editions of the standards indicated below.

ANSI/ISO 9899-1990, Programming Language-C.[1,2]

IEEE Std 802.1D-2004, IEEE Standard for Local and Metropolitan Area Networks: Media Access Control (MAC) Bridges.

[1] Replaces ANSI X3.159-1989

[2] ISO documents are available from ISO Central Secretariat, 1 Rue de Varembe, Case Postale 56, CH-1211, Geneve 20, Switzerland/Suisse; and from the Sales Department, American National Standards Institute, 11 West 42 Street, 13th Floor, New York, NY 10036-8002, USA

3. Terms, definitions, and notation

3.1 Conformance levels

Several key words are used to differentiate between different levels of requirements and options, as described in this subclause.

3.1.1 may: Indicates a course of action permissible within the limits of the standard with no implied preference ("may" means "is permitted to").

3.1.2 shall: Indicates mandatory requirements to be strictly followed in order to conform to the standard and from which no deviation is permitted ("shall" means "is required to").

3.1.3 should: An indication that among several possibilities, one is recommended as particularly suitable, without mentioning or excluding others; or that a certain course of action is preferred but not necessarily required; or that (in the negative form) a certain course of action is deprecated but not prohibited ("should" means "is recommended to").

3.2 Terms and definitions

For the purposes of this working paper, the following terms and definitions apply. The Authoritative Dictionary of IEEE Standards Terms [B2] should be referenced for terms not defined in the clause.

3.2.1 bridge: A functional unit interconnecting two or more networks at the data link layer of the OSI reference model.

3.2.2 clock master: A bridge or end station that provides the link clock reference.

3.2.3 clock slave: A bridge or end station that tracks the link clock reference provided by the clock master.

3.2.4 cyclic redundancy check (CRC): A specific type of frame check sequence computed using a generator polynomial.

3.2.5 grand clock master: The clock master selected to provide the network time reference.

3.2.6 link: A unidirectional channel connecting adjacent stations (half of a span).

3.2.7 listener: A sink of a stream, such as a television or acoustic speaker.

3.2.8 local area network (LAN): A communications network designed for a small geographic area, typically not exceeding a few kilometers in extent, and characterized by moderate to high data transmission rates, low delay, and low bit error rates.

3.2.9 MAC client: The layer entity that invokes the MAC service interface.

3.2.10 medium (plural: **media):** The material on which information signals are carried; e.g., optical fiber, coaxial cable, and twisted-wire pairs.

3.2.11 medium access control (MAC) sublayer: The portion of the data link layer that controls and mediates the access to the network medium. In this working paper, the MAC sublayer comprises the MAC datapath sublayer and the MAC control sublayer.

3.2.12 network: A set of communicating stations and the media and equipment providing connectivity among the stations.

3.2.13 plug-and-play: The requirement that a station perform classA transfers without operator intervention (except for any intervention needed for connection to the cable).

3.2.14 protocol implementation conformance statement (PICS): A statement of which capabilities and options have been implemented for a given Open Systems Interconnection (OSI) protocol.

3.2.15 span: A bidirectional channel connecting adjacent stations (two links).

3.2.16 station: A device attached to a network for the purpose of transmitting and receiving information on that network.

3.2.17 topology: The arrangement of links and stations forming a network, together with information on station attributes.

3.2.18 transmit (transmission): The action of a station placing a frame on the medium.

3.2.19 unicast: The act of sending a frame addressed to a single station.

3.3 Service definition method and notation

The service of a layer or sublayer is the set of capabilities that it offers to a user in the next higher (sub)layer. Abstract services are specified in this working paper by describing the service primitives and parameters that characterize each service. This definition of service is independent of any particular implementation (see Figure 3.1).

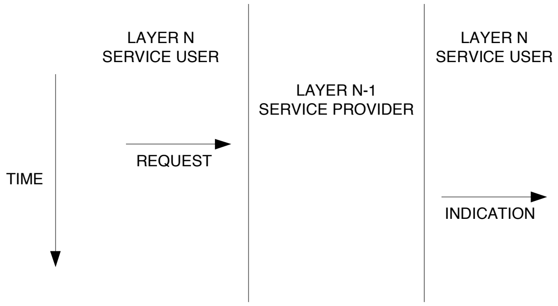


Figure 3.1—Service definitions

Specific implementations can also include provisions for interface interactions that have no direct end-to-end effects. Examples of such local interactions include interface flow control, status requests and indications, error notifications, and layer management. Specific implementation details are omitted from this service specification, because they differ from implementation to implementation and also because they do not impact the peer-to-peer protocols.

3.3.1 Classification of service primitives

Primitives are of two generic types.

a) REQUEST. The request primitive is passed from layer N to layer N-1 to request that a service be initiated.

b) INDICATION. The indication primitive is passed from layer N-1 to layer N to indicate an internal layer N-1 event that is significant to layer N. This event can be logically related to a remote service request, or can be caused by an event internal to layer N-1.

The service primitives are an abstraction of the functional specification and the user-layer interaction. The abstract definition does not contain local detail of the user/provider interaction. For instance, it does not indicate the local mechanism that allows a user to indicate that it is awaiting an incoming call. Each primitive has a set of zero or more parameters, representing data elements that are passed to qualify the functions invoked by the primitive. Parameters indicate information available in a user/provider interaction. In any particular interface, some parameters can be explicitly stated (even though not explicitly defined in the primitive) or implicitly associated with the service access point. Similarly, in any particular protocol specification, functions corresponding to a service primitive can be explicitly defined or implicitly available.

3.4 State machines

3.4.1 State machine behavior

The operation of a protocol can be described by subdividing the protocol into a number of interrelated functions. The operation of the functions can be described by state machines. Each state machine represents the domain of a function and consists of a group of connected, mutually exclusive states. Only one state of a function is active at any given time. A transition from one state to another is assumed to take place in zero time (i.e., no time period is associated with the execution of a state), based on some condition of the inputs to the state machine.

The state machines contain the authoritative statement of the functions they depict. When apparent conflicts between descriptive text and state machines arise, the order of precedence shall be formal state tables first, followed by the descriptive text, over any explanatory figures. This does not override, however, any explicit description in the text that has no parallel in the state tables.

The models presented by state machines are intended as the primary specifications of the functions to be provided. It is important to distinguish, however, between a model and a real implementation. The models are optimized for simplicity and clarity of presentation, while any realistic implementation might place heavier emphasis on efficiency and suitability to a particular implementation technology. It is the functional behavior of any unit that has to match the standard, not its internal structure. The internal details of the model are useful only to the extent that they specify the external behavior clearly and precisely.

3.4.2 State table notation

NOTE—The following state machine notation was used within 802.17, due to the exactness of C-code conditions and the simplicity of updating table entries (as opposed to 2-dimensional graphics).
Early state table descriptions can be converted (if necessary) into other formats before publication.

Each row of the table is preferably provided with a brief description of the condition and/or action for that row. The descriptions are placed after the table itself, and linked back to the rows of the table using numeric tags.

3.4.2.1 Parallel-execution state tables

State machines may be represented in tabular form. The table is organized into two columns: a left hand side representing all of the possible states of the state machine and all of the possible conditions that cause transitions out of each state, and the right hand side giving all of the permissible next states of the state machine as well as all of the actions to be performed in the various states, as illustrated in Table 3.1. The syntax of the expressions follows standard C notation (see 3.13). No time period is associated with the transition from one state to the next.

Table 3.1—State table notation example

Current		Row	Next	
state	**condition**		**action**	**state**
START	sizeOfMacControl > spaceInQueue	1	—	START
	passM == 0	2		
	—	3	TransmitFromControlQueue();	FINAL
FINAL	SelectedTransferCompletes()	4	—	START
	—	5	—	FINAL

Row 3.1-1: Do nothing if the size of the queued MAC control frame is larger than the PTQ space.
Row 3.1-2: Do nothing in the absence of MAC control transmission credits.
Row 3.1-3: Otherwise, transmit a MAC control frame.

Row 3.1-4: When the transmission completes, start over from the initial state (i.e., START).
Row 3.1-5: Until the transmission completes, remain in this state.

Each combination of current state, next state, and transition condition linking the two is assigned to a different row of the table. Each row of the table, read left to right, provides: the name of the current state; a condition causing a transition out of the current state; an action to perform (if the condition is satisfied); and, finally, the next state to which the state machine transitions, but only if the condition is satisfied. The symbol "—" signifies the default condition (i.e., operative when no other condition is active) when placed in the condition column, and signifies that no action is to be performed when placed in the action column. Conditions are evaluated in order, top to bottom, and the first condition that evaluates to a result of TRUE is used to determine the transition to the next state. If no condition evaluates to a result of TRUE, then the state machine remains in the current state. The starting or initialization state of a state machine is always labeled "START" in the table (though it need not be the first state in the table). Every state table has such a labeled state.

Each row of the table is preferably provided with a brief description of the condition and/or action for that row. The descriptions are placed after the table itself, and linked back to the rows of the table using numeric tags.

3.4.2.2 Called state tables

A RETURN state is the terminal state of a state machine that is intended to be invoked by another state machine, as illustrated in Table 3.2. Once the RETURN state is reached, the state machine terminates execution, effectively ceasing to exist until the next invocation by the caller, at which point it begins execution again from the START state. State machines that contain a RETURN state are considered to be only instantiated when they are invoked. They do not have any persistent (static) variables.

Table 3.2—Called state table notation example

Current		**Row**	**Next**	
state	**condition**		**action**	**state**
START	sizeOfMacControl > spaceInQueue	1	—	FINAL
	passM == 0	2		
	—	3	TransmitFromControlQueue();	RETURN
FINAL	MacTransmitError();	4	errorDefect = TRUE	RETURN
	—	5	—	

Row 3.2-1: The size of the queued MAC control frame is less than the PTQ space.
Row 3.2-2: In the absence of MAC control transmission credits, no action is taken.
Row 3.2-3: MAC control transmissions have precedence over client transmissions.

Row 3.2-4: If the transmission completes with an error, set an error defect indication.
Row 3.2-5: Otherwise, no error defect is indicated.

3.5 Arithmetic and logical operators

In addition to commonly accepted notation for mathematical operators, Table 3.3 summarizes the symbols used to represent arithmetic and logical (boolean) operations. Note that the syntax of operators follows standard C notation (see 3.13).

Table 3.3—Special symbols and operators

Printed character	Meaning
&&	Boolean AND
||	Boolean OR
!	Boolean NOT (negation)
&	Bitwise AND
|	Bitwise OR
^	Bitwise XOR
<=	Less than or equal to
>=	Greater than or equal to
==	Equal to
!=	Not equal to
=	Assignment operator
//	Comment delimiter

3.6 Numerical representation

NOTE—The following notation was taken from 802.17, where it was found to have benefits:
– The subscript notation is consistent with common mathematical/logic equations.
– The subscript notation can be used consistently for all possible radix values.

Decimal, hexadecimal, and binary numbers are used within this working paper. For clarity, decimal numbers are generally used to represent counts, hexadecimal numbers are used to represent addresses, and binary numbers are used to describe bit patterns within binary fields.

Decimal numbers are represented in their usual 0, 1, 2, … format. Hexadecimal numbers are represented by a string of one or more hexadecimal (0-9,A-F) digits followed by the subscript 16, except in C-code contexts, where they are written as `0x123EF2` etc. Binary numbers are represented by a string of one or more binary (0,1) digits, followed by the subscript 2. Thus the decimal number “26” may also be represented as “$1A_{16}$” or “11010_2”.

MAC addresses and OUI/EUI values are represented as strings of 8-bit hexadecimal numbers separated by hyphens and without a subscript, as for example “01-80-C2-00-00-15” or “AA-55-11”.

3.7 Field notations

3.7.1 Use of italics

All field names or variable names (such as *level* or *myMacAddress*), and sub-fields within variables (such as *thisState.level*) are italicized within text, figures and tables, to avoid confusion between such names and similarly spelled words without special meanings. A variable or field name that is used in a subclause heading or a figure or table caption is also italicized. Variable or field names are not italicized within C code, however, since their special meaning is implied by their context. Names used as nouns (e.g., subclassA0) are also not italicized.

3.7.2 Field conventions

This working paper describes values that are packetized or MAC-resident, such as those illustrated in Table 3.2.

Table 3.4—Names of fields and sub-fields

Name	Description
newCRC	Field within a register or frame
thisState.level	Sub-field within field *thisState*
thatState.rateC[*n*].*c*	Sub-field within array element *rateC*[*n*]

Run-together names (e.g., *thisState*) are used for fields because of their compactness when compared to equivalent underscore-separated names (e.g., *this_state*). The use of multiword names with spaces (e.g., “This State”) is avoided, to avoid confusion between commonly used capitalized key words and the capitalized word used at the start of each sentence.

A sub-field of a field is referenced by suffixing the field name with the sub-field name, separated by a period. For example, *thisState.level* refers to the sub-field *level* of the field *thisState*. This notation can be continued in order to represent sub-fields of sub-fields (e.g., *thisState.level.next* is interpreted to mean the sub-field *next* of the sub-field *level* of the field *thisState*).

Two special field names are defined for use throughout this working paper. The name *frame* is used to denote the data structure comprising the complete MAC sublayer PDU. Any valid element of the MAC sublayer PDU, can be referenced using the notation *frame.xx* (where *xx* denotes the specific element); thus, for instance, *frame.serviceDataUnit* is used to indicate the *serviceDataUnit* element of a frame.

Unless specifically specified otherwise, reserved fields are reserved for the purpose of allowing extended features to be defined in future revisions of this working paper. For devices conforming to this version of this working paper, nonzero reserved fields are not generated; values within reserved fields (whether zero or nonzero) are to be ignored.

3.7.3 Field value conventions

This working paper describes values of fields. For clarity, names can be associated with each of these defined values, as illustrated in Table 3.5. A symbolic name, consisting of upper case letters with underscore separators, allows other portions of this working paper to reference the value by its symbolic name, rather than a numerical value.

Table 3.5—*wrap* field values

Value	Name	Description
0	STANDARD	Standard processing selected
1	SPECIAL	Special processing selected
2,3	—	Reserved

Unless otherwise specified, reserved values allow extended features to be defined in future revisions of this working paper. Devices conforming to this version of this working paper do not generate nonzero reserved values, and process reserved fields as though their values were zero.

A field value of TRUE shall always be interpreted as being equivalent to a numeric value of 1 (one), unless otherwise indicated. A field value of FALSE shall always be interpreted as being equivalent to a numeric value of 0 (zero), unless otherwise indicated.

3.8 Bit numbering and ordering

Data transfer sequences normally involve one or more cycles, where the number of bytes transmitted in each cycle depends on the number of byte lanes within the interconnecting link. Data byte sequences are shown in figures using the conventions illustrated by Figure 3.2, which represents a link with four byte lanes. For multi-byte objects, the first (left-most) data byte is the most significant, and the last (right-most) data byte is the least significant.

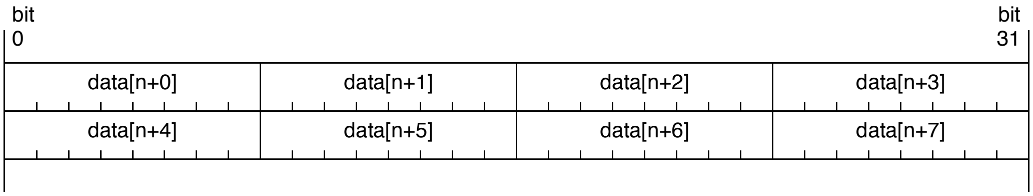


Figure 3.2—Bit numbering and ordering

Figures are drawn such that the counting order of data bytes is from left to right within each cycle, and from top to bottom between cycles. For consistency, bits and bytes are numbered in the same fashion.

NOTE—The transmission ordering of data bits and data bytes is not necessarily the same as their counting order; the translation between the counting order and the transmission order is specified by the appropriate reconciliation sublayer.

3.9 Byte sequential formats

Figure 3.3 provides an illustrative example of the conventions to be used for drawing frame formats and other byte sequential representations. These representations are drawn as fields (of arbitrary size) ordered along a vertical axis, with numbers along the left sides of the fields indicating the field sizes in bytes. Fields are drawn contiguously such that the transmission order across fields is from top to bottom. The example shows that *field1*, *field2*, and *field3* are 1-, 1- and 6-byte fields, respectively, transmitted in order starting with the *field1* field first. As illustrated on the right hand side of Figure 3.3, a multi-byte field represents a sequence of ordered bytes, where the first through last bytes correspond to the most significant through least significant portions of the multi-byte field, and the MSB of each byte is drawn to be on the left hand side.

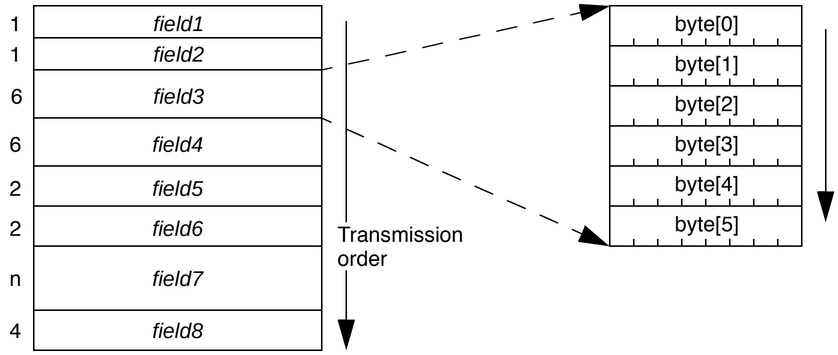


Figure 3.3—Byte sequential field format illustrations

NOTE—Only the left-hand diagram in Figure 3.3 is required for representation of byte-sequential formats. The right-hand diagram is provided in this description for explanatory purposes only, for illustrating how a multi-byte field within a byte sequential representation is expected to be ordered. The tag "Transmission order" and the associated arrows are not required to be replicated in the figures.

3.10 Ordering of multibyte fields

In many cases, bit fields within byte or multibyte objects are expanded in a horizontal fashion, as illustrated in the right side of Figure 3.4. The fields within these objects are illustrated as follows: left-to-right is the byte transmission order; the left-through-right bits are the most significant through least significant bits respectively.

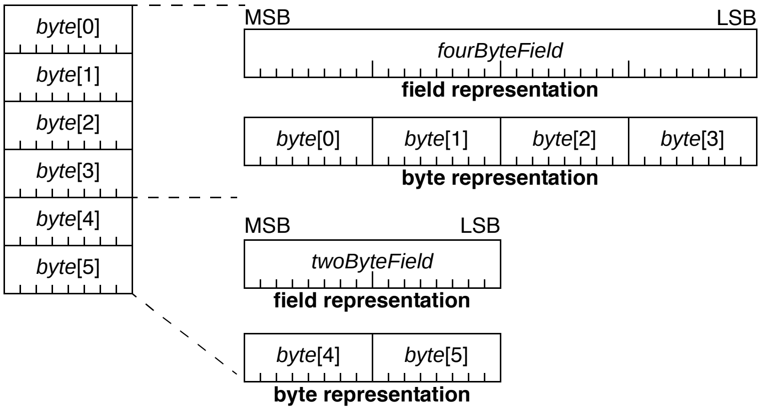


Figure 3.4—Multibyte field illustrations

The first *fourByteField* can be illustrated as a single entity or a 4-byte multibyte entity. Similarly, the second *twoByteField* can be illustrated as a single entity or a 2-byte multibyte entity.

> NOTE—The following text was taken from 802.17, where it was found to have benefits:
> The details should, however, be revised to illustrate fields within an AVB frame header *serviceDataUnit*.

To minimize potential for confusion, four equivalent methods for illustrating frame contents are illustrated in Figure 3.5. Binary, hex, and decimal values are always shown with a left-to-right significance order, regardless of their bit-transmission order.

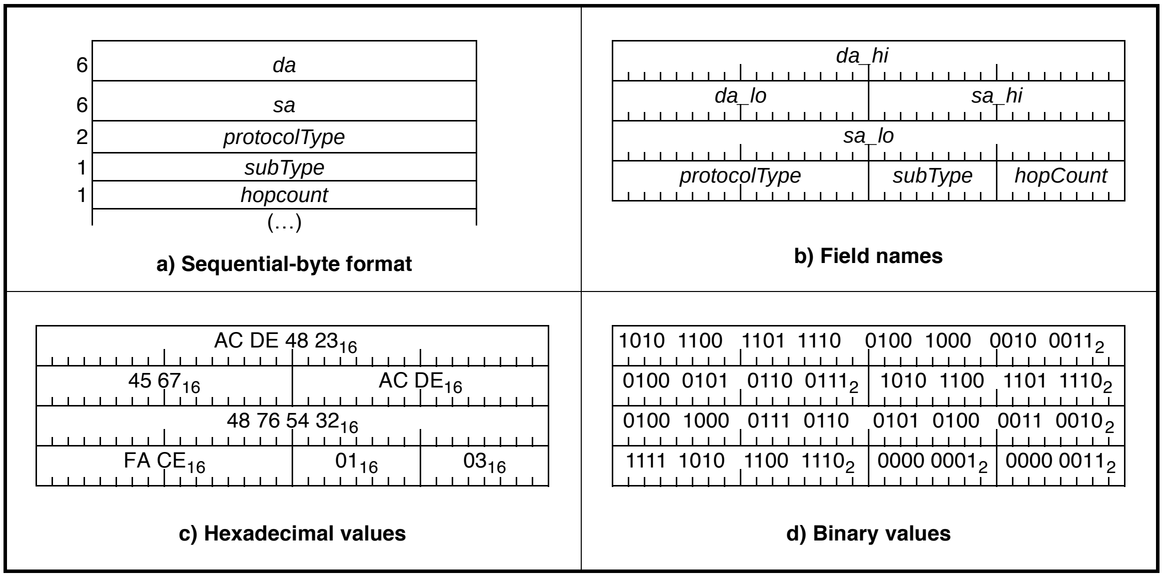
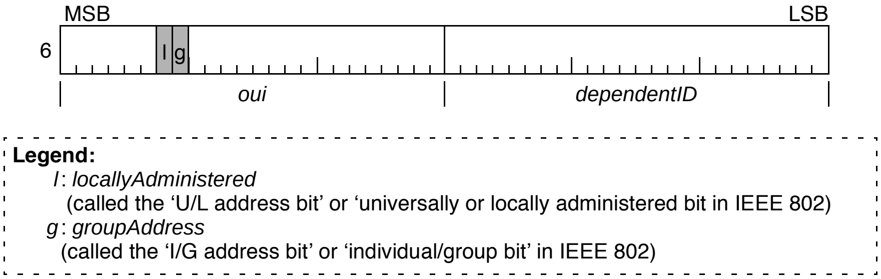


Figure 3.5—Illustration of fairness-frame structure

3.11 MAC address formats

The format of MAC address fields within frames is illustrated in Figure 3.6.



Legend:
l: *locallyAdministered*
(called the ‘U/L address bit’ or ‘universally or locally administered bit in IEEE 802)
g: *groupAddress*
(called the ‘I/G address bit’ or ‘individual/group bit’ in IEEE 802)

Figure 3.6—MAC address format

3.11.1 ***oui*:** A 24-bit organizationally unique identifier (OUI) field supplied by the IEEE/RAC for the purpose of identifying the organization supplying the (unique within the organization, for this specific context) 24-bit *dependentID*. (For clarity, the *locallyAdministered* and *groupAddress* bits are illustrated by the shaded bit locations.)

3.11.2 ***dependentID*****:** An 24-bit field supplied by the *oui*-specified organization. The concatenation of the *oui* and *dependentID* provide a unique (within this context) identifier.

To reduce the likelihood of error, the mapping of OUI values to the *oui*/*dependentID* fields are illustrated in Figure 3.7. For the purposes of illustration, specific OUI and *dependentID* example values have been assumed. The two shaded bits correspond to the *locallyAdministered* and *groupAddress* bit positions illustrated in Figure 3.6.

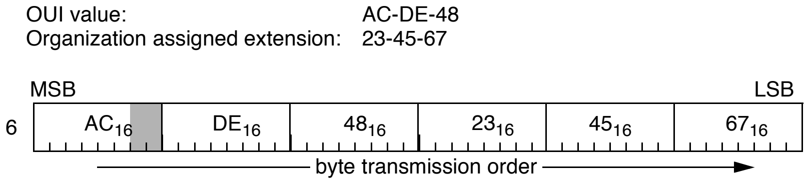


Figure 3.7—48-bit MAC address format

3.12 Informative notes

Informative notes are used in this working paper to provide guidance to implementers and also to supply useful background material. Such notes never contain normative information, and implementers are not required to adhere to any of their provisions. An example of such a note follows.

NOTE—This is an example of an informative note.

3.13 Conventions for C code used in state machines

Many of the state machines contained in this working paper utilize C code functions, operators, expressions and structures for the description of their functionality. Conventions for such C code can be found in Annex E.

4. Abbreviations and acronyms

NOTE—This clause should be skipped on the first reading (continue with Clause 5).
This text has been lifted from the P802.17 draft standard, which has a relative comprehensive list. Abbreviations/acronyms are expected to be added, revised, and/or deleted as this working paper evolves.

This working paper contains the following abbreviations and acronyms:

AP	access point
AV	audio/video
AVB	audio/video bridging
AVB network	audio/video bridged network
BER	bit error ratio
BMC	best master clock
BMCA	best master clock algorithm
CRC	cyclic redundancy check
FIFO	first in first out
IEC	International Electrotechnical Commission
IEEE	Institute of Electrical and Electronics Engineers
IETF	Internet Engineering Task Force
ISO	International Organization for Standardization
ITU	International Telecommunication Union
LAN	local area network
LSB	least significant bit
MAC	medium access control
MAN	metropolitan area network
MSB	most significant bit
OSI	open systems interconnect
PDU	protocol data unit
PHY	physical layer
PLL	phase-locked loop
RFC	request for comment
RPR	resilient packet ring
VOIP	voice over internet protocol

5. Architecture overview

5.1 Application scenarios

5.1.1 Garage jam session

As an illustrative example, consider AVB usage for a garage jam session, as illustrated in Figure 5.1. The audio inputs (microphone and guitar) are converted, passed through a guitar effects processor, two bridges, mixed within an audio console, return through two bridges, and return to the ear through headphones.

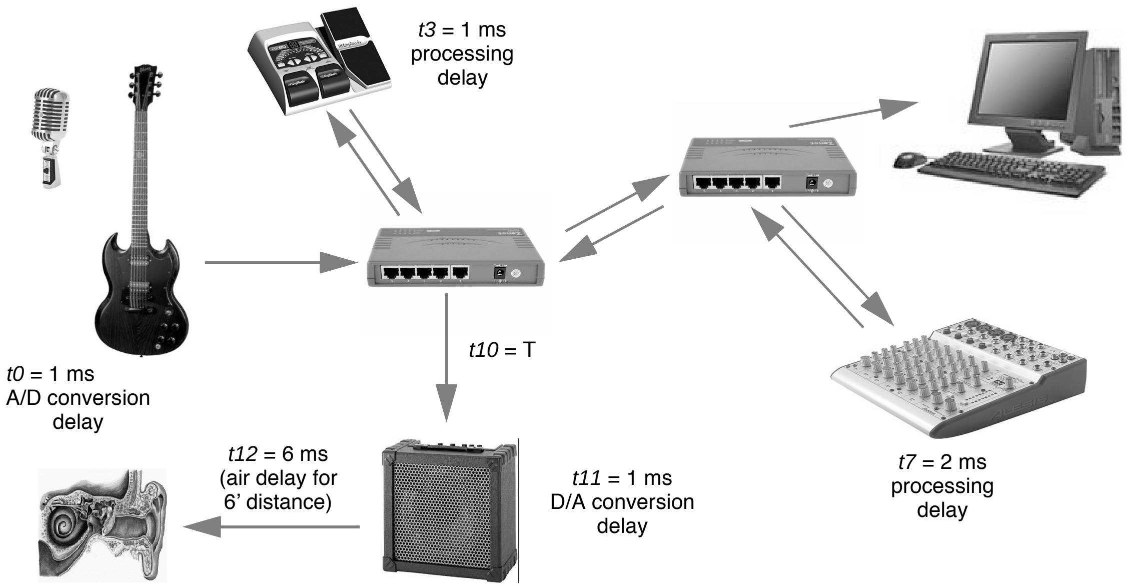


Figure 5.1—Garage jam session

Using Ethernet within such systems has multiple challenges: low-latency and tight time-synchronization. Tight time synchronization is necessary to avoid cycle slips when passing through multiple processing components and (ultimately) to avoid overrun/underrun at the the final D/A converter's FIFO. The challenge of low-latency transfers is being addressed in other forums and is outside the scope of this draft.

5.1.2 Looping topologies

Bridged Ethernet networks currrently have no loops, but bridging extensions are contemplating looping topologies. To ensure longevity of this standard, the time-synchronization protocols are tolerant of looping topologies that could occur (for example) if the dotted-line link were to be connected in Figure 5.2.

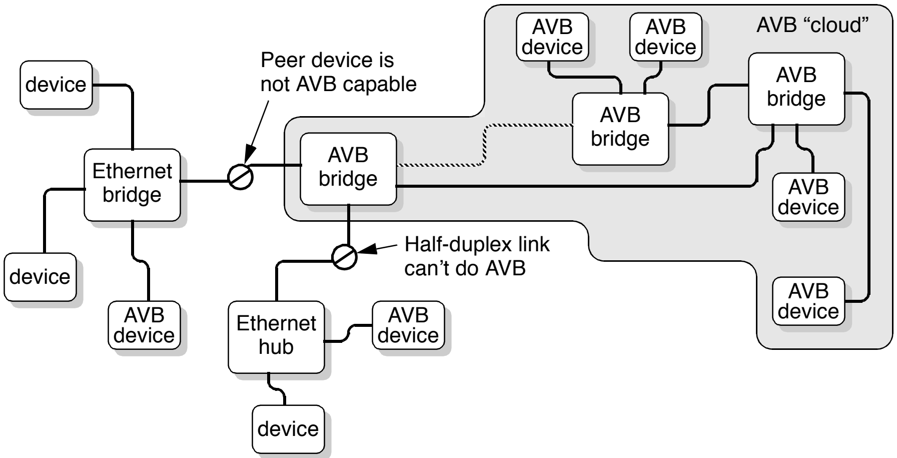


Figure 5.2—Possible looping topology

Separation of AVB devices is driven by the requirements of AVB bridges to support subscription (bandwidth allocation) and pacing of time-sensitive transmissions, as well as time-of-day clock-synchronization.

5.2 Design methodology

5.2.1 Assumptions

This working paper specifies a protocol to synchronize independent timers running on separate stations of a distributed networked system, based on concepts specified within IEEE Std 1588-2002. Although a high degree of accuracy and precision is specified, the technology is applicable to low-cost consumer devices. The protocols are based on the following design assumptions:

a) Each end station and intermediate bridges provide independent clocks.

b) All clocks are accurate, typically to within ±100PPM.

c) Details of the best time-synchronization protocols are physical-layer dependent.

5.2.2 Objectives

With these assumptions in mind, the time synchronization objectives include the following:

a) Precise. Multiple timers can be synchronized to within 10's of nanoseconds.

b) Inexpensive. For consumer AVB devices, the costs of synchronized timers are minimal.
(GPS, atomic clocks, or 1PPM clock accuracies would be inconsistent with this criteria.)

c) Scalable. The protocol is independent of the networking technology. In particular:

   1) Cyclical physical topologies are supported.
   2) Long distance links (up to 2 kM) are allowed.

d) Plug-and-play. The system topology is self-configuring; no system administrator is required.

5.2.3 Strategies

Strategies used to meet these objectives include the following:

a) Precision is achieved by calibrating and adjusting *timeOfDay* clocks.
   1) Offsets. Offset value adjustments eliminate immediate clock-value errors.
   2) Rates. Rate value adjustments reduce long-term clock-drift errors.
b) Simplicity is achieved by the following:
   1) Concurrence. Most configuration and adjustment operations are performed concurrently.
   2) Feed-forward. PLLs are unnecessary within bridges, but possible within applications.
   3) Frequent. Frequent (nominally 100 Hz) interchanges reduces needs for overly precise clocks.

5.3 Time-synchronization facilities

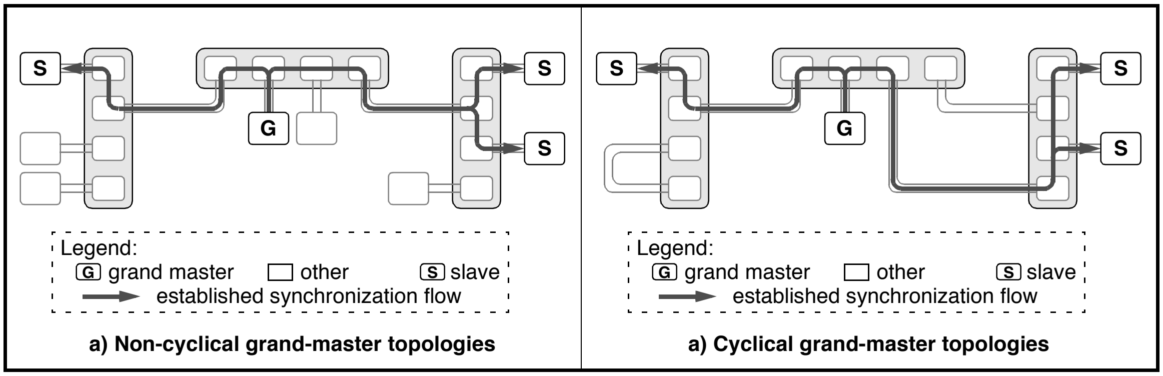
In concept, the clock-synchronization protocol starts with the selection of the reference-timer station, called a grand-master station (oftentimes abbreviated as grand-master). Every AVB-capable station is grand-master capable, but only one is selected to become the grand-master station within each network. To assist in the grand-master selection, each station is associated with a distinct preference value; the grand-master is the station with the “best” preference values. Thus, time-synchronization services involve two subservices, as listed below and described in the following subclauses.

a) Selection. Looping topologies are isolated (from a time-synchronizatin perspective) into a spanning tree. The root of the tree, which provides the time reference to others, is the grand master.
b) Distribution. Synchronized time is distributed through the the grand-master’s spanning tree.

5.3.1 Grand-master selection

5.3.1.1 Grand-master responsibilities

Clock synchronization involves streaming of timing information from a grand-master timer to one or more slave timers. Although primarily intended for non-cyclical physical topologies (see Figure 5.3a), the synchronization protocols also function correctly on cyclical physical topologies (see Figure 5.3b), by activating only a non-cyclical subset of the physical topology.



a) Non-cyclical grand-master topologies

a) Cyclical grand-master topologies

Figure 5.3—Timing information flows

5.3.1.2 Grand-master behavior

As part of the grand-master selection process, stations forward the best of their observed preference values to neighbor stations, allowing the overall best-preference value to be ultimately selected and known by all. The station whose preference value matches the overall best-preference value ultimately becomes the grand-master.

The grand-master station observes that its precedence is better than values received from its neighbors, as illustrated in Figure 5.4a. A slave stations observes its precedence to be worse than one of its neighbors and forwards the best-neighbor precedence value to adjacent stations, as illustrated in Figure 5.4b. To avoid cyclical behaviors, a *hopsCount* value is associated with preference values and is incremented before the best-precedence value is communicated to others.

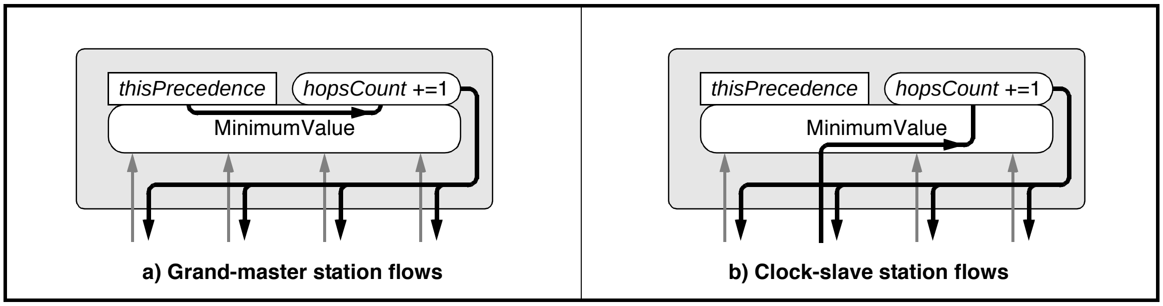


Figure 5.4—Grand-master precedence flows

5.3.1.3 Grand-master precedence

Grand-master precedence is based on the concatenation of multiple fields, as illustrated in Figure 5.5. The *port* value is used within bridges, but is not transmitted between stations.

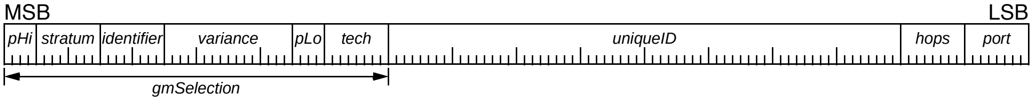
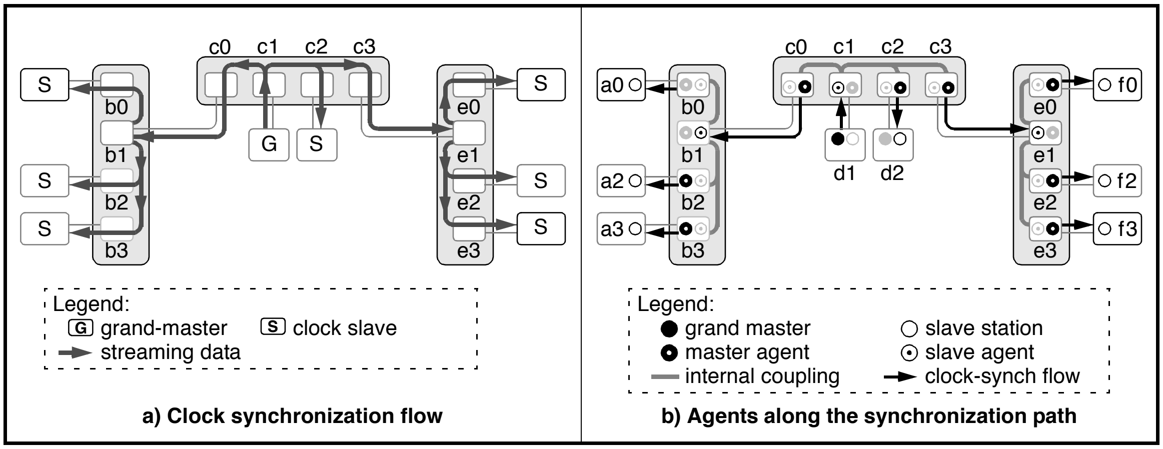


Figure 5.5—Grand-master precedence

This format is similar to the format of the spanning-tree precedence value, but a wider *uniqueID* is provided for compatibility with interconnects based on 64-bit station identifiers.

5.3.2 Synchronized-time distribution

Clock-synchronization information conceptually flows from a grand-master station to clock-slave stations, as illustrated in Figure 5.6a. A more detailed illustration shows pairs of synchronized clock-master and clock-slave components, as illustrated in Figure 5.6b. The active clock agents are illustrated as black-and-white components; the passive clock agents are illustrated as grey-and-white components.



a) Clock synchronization flow

b) Agents along the synchronization path

Figure 5.6—Hierarchical flows

Internal communications distribute synchronized time from clock-slave agents b1, c1, and e1 to the other clock-master agents on bridgeB, bridgeC, and bridgeE respectively. Within a clock-slave, precise time synchronization involves adjustments of timer value and rate-of-change values.

Time synchronization yields distributed but closely-matched *timeOfDay* values within stations and bridges. No attempt is made to eliminate intermediate jitter with bridge-resident jitter-reducing phase-lock loops (PLLs,) but application-level phase locked loops (not illustrated) are expected to filter high-frequency jitter from the supplied *timeOfDay* values

5.3.3 MAC service model

The MAC service model assumes the presence of one or more time-synchronized AVB ports communicating with a higher-level MAC client, as illustrated in Figure 5.7. The receive portion of each port provides grand-master precedence information, which assists in the grand-master selection process. The transmit portion of each port provides the higher-level client with a local time and (in response) receives status containing the grand-master precedence and the client's synchronized version of the grand-master time. All components are assumed to have access to a common free-running (not adjustable) local timer. There is not necessarily a one-to-one correspondence between the primitives and formal procedures and the interfaces in any particular implementation.

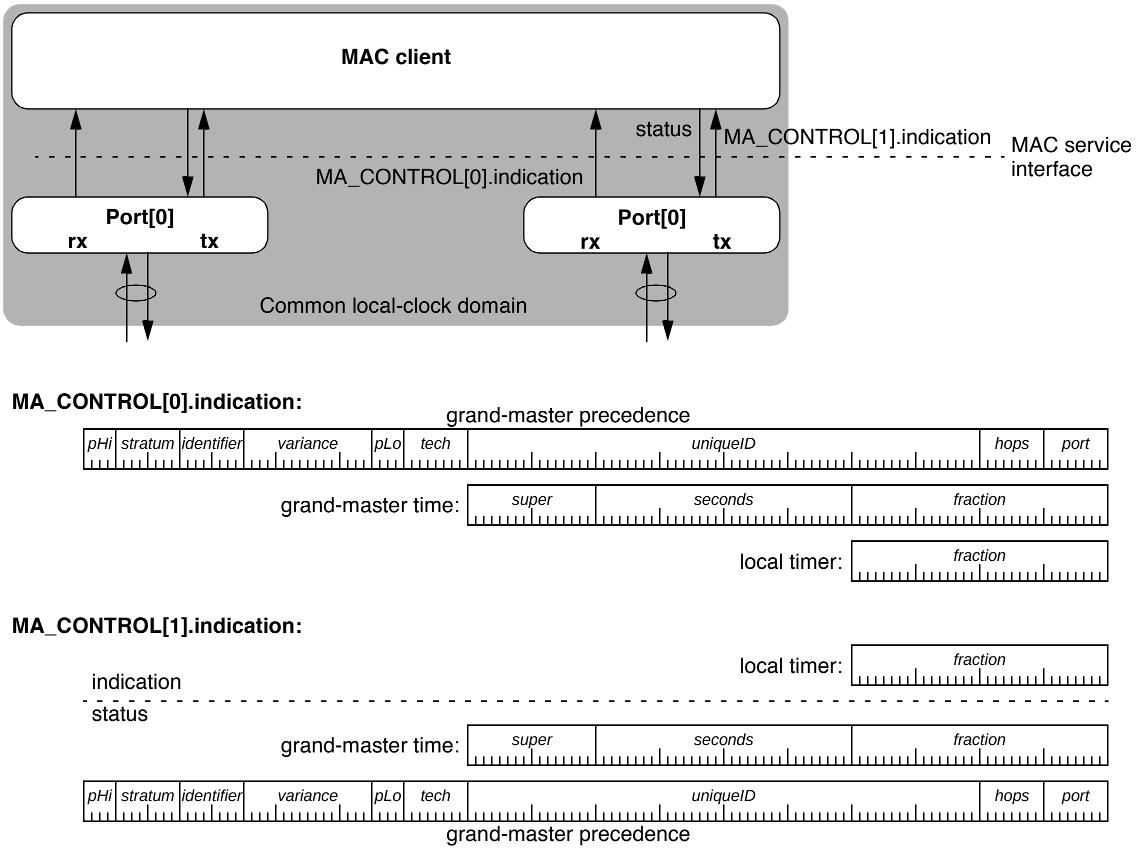


Figure 5.7 — AVB service interface model

5.4 Client processing details

TBD.

5.5 Key distinctions from IEEE Std 1588

Although based on the concepts of IEEE Std 1588, this draft is different in multiple ways:

a) All bridges are boundary clocks, since they compensate their received time by pass-through delays, by setting the client time on received frames, then transmitting the current client time when frames are transmitted. There are no transparent (in 1588 terminology) bridges.

b) The processing by the client is distinguished by the processing performed in the MAC, thus isolating the client from that multiple (and sometimes strange) wireless and PON MACs protocols.

c) To simplify computations, time is uniformly represented as a simple 80-bit scaled signed integer.

d) For Ethernet, a higher update frequency of 100 Hz is assumed. This reduces timeouts for failed grand masters, and worst-case times for clear the network of rogue packets, while also reducing timer-value drifts between updates.

e) For Ethernet, only one frame type simplifies the protocols and reduces transient-recovery times.

6. Medium access control (MAC) service and reference model

6.1 Overview

This clause provides an overview of the MAC sublayer and the reconciliation sublayer (see shaded portions of Figure 6.1), specifies the services provided by the MAC sublayer, including the MAC control sublayer and the MAC datapath sublayer, and provides a reference model for the MAC sublayer. Higher-layer clients can include the logical link control (LLC) sublayer, bridge relay entity, or other users of ISO/IEC LAN international standard MAC services. The services are described in an abstract way and do not imply any particular implementations or any exposed interfaces. There is not necessarily a one-to-one correspondence between the primitives and formal procedures and the interfaces in any particular implementation.

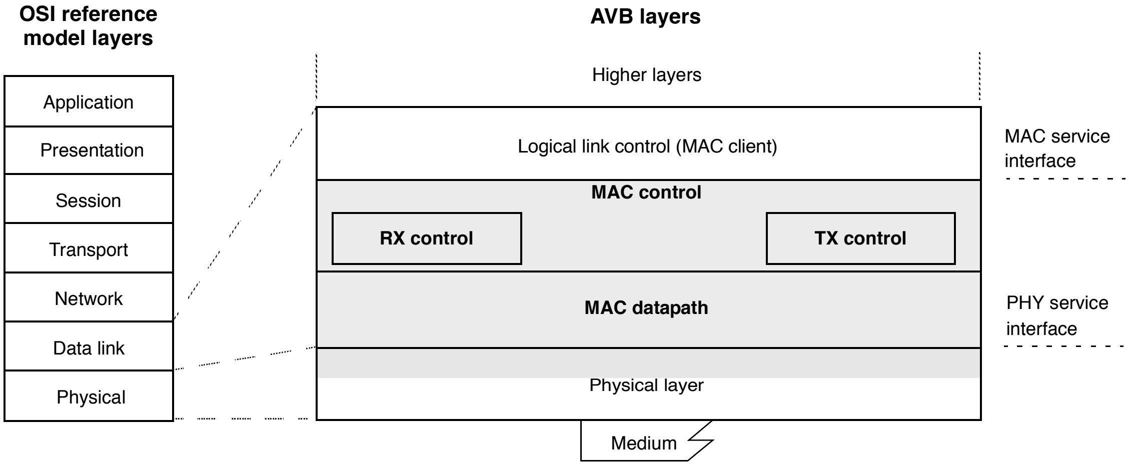


Figure 6.1 — AVB service and reference model relationship to the ISO/IEC OSI reference model

6.2 Overview of MAC services

The services provided by the MAC sublayer are listed below. For each port, these services are invoked periodically at modest processing rates (100Hz, for Ethernet MACs).

a) The receive port provides grand-master precedence and time, as well as local time, to the client.

b) The transmit port provides the client with a recent-past local time and (in response) the client provides the grand-master precedence and time as status.

6.3 MAC services to the client layer

The services of a layer or sublayer are the set of capabilities that it offers to the next higher (sub)layer. The services specified in this standard are described by abstract service primitives and parameters that characterize each service. This definition of a service is independent of any particular implementation.

The following two service primitives are defined for the client interfaces, and shall be implemented.

— MA_CONTROL[0].indication
— MA_CONTROL[1].indication

6.3.1 MA_CONTROL[0].indication

6.3.1.1 Function

The MA_CONROL[0].indication primitive provides the client with an association of local time and the grand-master time (as received from the closer-to grand-master station), as well as the grand-master precedence.

6.3.1.2 Semantics of the service primitive

The semantics of the primitives are as follows:

```
MA_CONTROL[0].indication
(
    local_time,
    gm_precedence,
    gm_time
)
```

The parameters of the MA_CONROL[0].indication are described below.

local_time
Provides the value of a local time when the value of gm_time was sampled.

gm_precedence
Provides the recently observed grand-master predecence, or NULL if such as precedence has not been received within a physical-layer dependent timeout interval. Fields within this component include the following (more-significant fields are listed first):

pHi A 4-bit user-selectable overriding precedence.
stratum An 8-bit identifier that identifies the type of grand-master.
identifier An 8-bit identifier that identifies ….
variance A 16-bit value that identifies the grand-master clock quality.
pLo A 4-bit user-selectable tie-breaking precedence.
tech An 8-bit value that, when appended with the uniqueID, is globally unique.
uniqueID An 8-bit value that, when prepended with the tech, is globally unique.
hops An 8-bit value that counts visible repeaters from the grand master.
port An 8-bit value that identifies the port that received the gm_precedence value.

gm_time
Provides the value of the distributed grand-master time, compensated for intermediate delays. Fields within this component include the following (more-significant fields are listed first):

super A signed 16-bit value representing 2^{32} seconds.
seconds An unsigned 32-bit value representing seconds.
fraction An unsigned 32-bit value representing fractions-of-seconds.

6.3.1.3 When generated

The MA_CONROL[0].indication primitive is invoked by the MAC entity whenever new knowledge of time or the grand-master precedence is generated.

6.3.1.4 Effect of receipt

The receipt of the MA_CONROL[0].indication primitive causes the client entity to adjust its image of grand-master time and rate. Details are TBD.

6.3.2 MA_CONTROL[1].indication

6.3.2.1 Function

The MA_CONROL[1].indication primitive provides the client with a recent value of local time; the client returns the associated grand-master time (as synchronized to the closer-to grand-master station), as well as the grand-master precedence.

6.3.2.2 Semantics of the service primitive

The semantics of the primitives are as follows:

```
MA_CONTROL[1].indication
(
    local_time,
)

MA_CONTROL[1].status
(
    gm_precedence,
    gm_time
)
```

The parameters of the MA_CONROL[1].indication are described below.

local_time
Provides the value of a local time when the value of gm_time was sampled.

The returned values for the MA_CONROL[1].indication are described below.

gm_precedence
Provides the client's recently observed best grand-master precedence. See xx for details.
gm_time
Provides the value of the distributed grand-master time, compensated for intermediate delays. See xx for details.

6.3.2.3 When generated

The MA_CONROL[1].indication primitive is periodically invoked by the MAC entity whenever new knowledge of time and/or the grand-master precedence is needed to update attached clock-slave neighbors.

6.3.2.4 Effect of receipt

The receipt of the MA_CONROL[0].indication primitive causes the client entity to returns the desired status values.

7. Ethernet duplex-cable time synchronization

7.1 Design methodology

7.1.1 Assumptions

Support of duplex-link Ethernet is based on concepts specified within IEEE Std 1588-2002. Although a high degree of accuracy and precision is specified, the technology is applicable to low-cost consumer devices. The protocols are based on the following design assumptions:

a) Point-to-point transmit/receive duplex connections are provided.

b) Transmit/receive propagation delays within duplex cables are well matched.

7.1.2 Strategies

Strategies used to meet these objectives include the following:

a) Precision is achieved by calibrating and adjusting *timeOfDay* clocks.
   1) Offsets. Offset value adjustments eliminate immediate clock-value errors.
   2) Rates. Rate value adjustments reduce long-term clock-drift errors.

b) Simplicity is achieved by the following:
   1) Symmetric. Clock-master/clock-slave computations are similar (only slave results are saved).
   2) Periodic. Messages are sent periodically, rather than in timely response to other requests.
   3) Frequent. Frequent (typically 1 kHz) interchanges reduces needs for precise clocks.

c) Balanced functionality.
   1) Low-rate. Complex computations are infrequent and can be readily implemented in firmware.
   2) High-rate. Frequent computations are simple and can be readily implemented in hardware.

7.2 Time-synchronization operation

7.2.1 Periodic packet transmissions

Time-sychronization involves periodic not-necessarily synchronized packet transmissions between adjacent stations, as illustrated in Figure 7.1a. The transmitted frame contains the following information:

select—specifies grand-master precedence
global—an estimation of the grand-master time
local—a sampling of the station-local time
delta—derived parameters from the neighbor, returned in a following cycle.

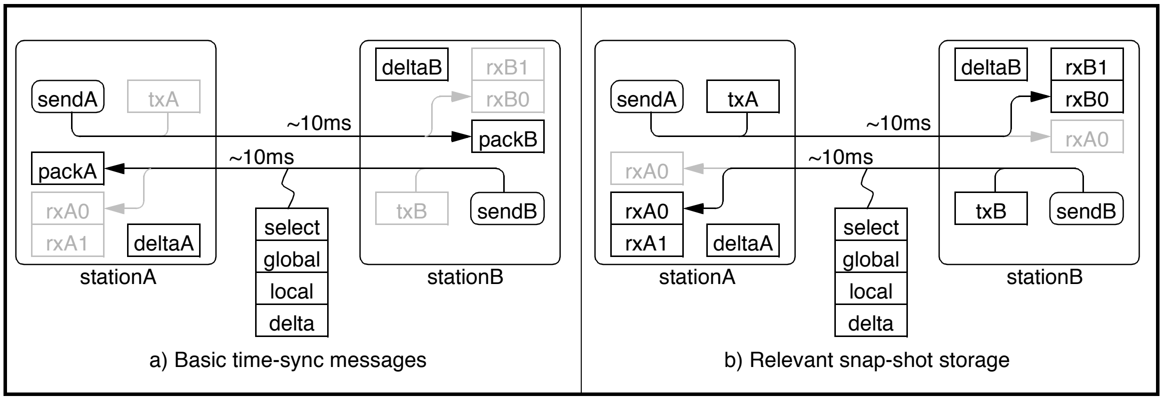


Figure 7.1—Timer snapshot locations

Snapshots are taken when packets are transmitted (illustrated as *txA* and *txB*) and received (illustrated as *rxA* and *rxB*), as illustrated in Figure 7.2b. The transmitted stopshot *txA* is placed into the next frame that is transmitted, as *packetA.local*, along with grand-master time *packetA.global* sampled at this time. The transmitted stopshot *txB* is similarly placed into the next frame that is transmitted, as *packetB.local*, along with grand-master time *packetB.global* sampled at this time.

The receive snapshot is double buffered, in that the value of *rxB0* is copied to *rxB1* when the *rxB0* snapshot is taken. Similarly, the value of *rxA0* is copied to *rxA1* when the *rxA0* snapshot is taken.

The computed value of *deltaA* is the difference between the received *packetFromB.local* value and the previous *rxA* snapshot, as specified by Equation 7.1. Similarly, *deltaB* (the value transmitted from stationB to stationA) is specified by Equation 7.2.

$$deltaA = rxA1 - packetFromB.local; \tag{7.1}$$
$$deltaB = rxB1 - packetFromA.local; \tag{7.2}$$

The value of the intermediate span delay is readily derived from these values. At stationA and stationB, these computations are specified by Equation 7.3 and Equation 7.4, respectively.

$$cableDelayComputedAtA = (deltaA + packetFromB.delta)/2; \tag{7.3}$$
$$cableDelayComputedAtB = (packetFromA.delta + deltaB)/2; \tag{7.4}$$

7.2.2 Clock-slave client-supplied information

The clock-slave client-supplied information involves transmission of information from the port to the client (see Clause 6). For stationA, this information is listed below.

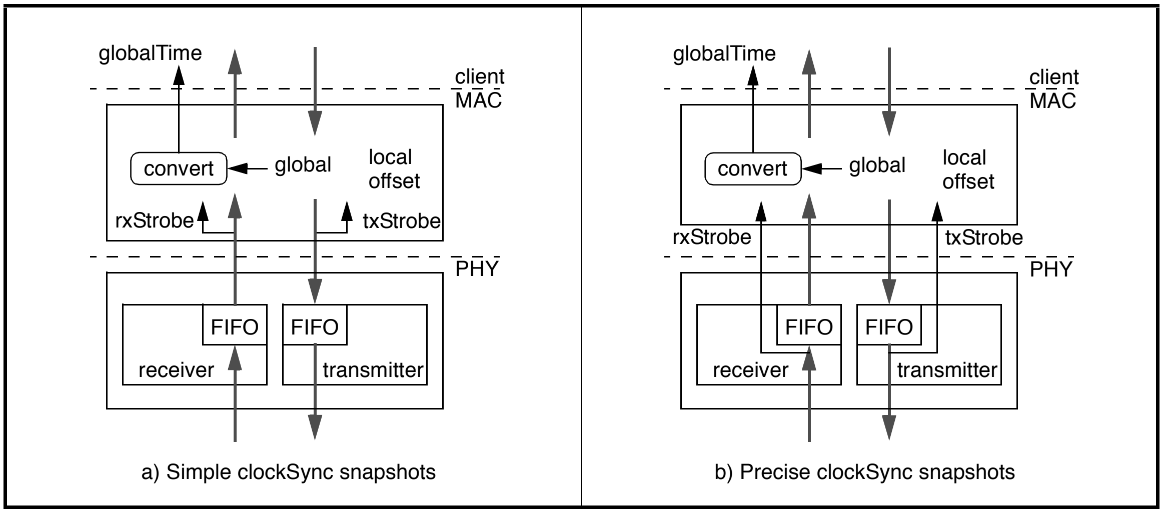
local_time The value of *rxA1*.
gm_time The sum of two components, *packetFromB.global*+*deltaA*, where:
packetFromB.global—the value received from the clock-master stationB.
deltaA—the most-recent computed value.

For stationB this information is listed below.

local_time The value of *rxB1*.
gm_time The value of two components, *packetFromA.global*+*deltaB*, where:
packetFromA.global—the value received from clock-master stationA.
deltaB—the most-recent computed value.

7.2.3 Timer snapshot locations

Mandatory jitter-error accuracies are sufficiently loose to allow transmit/receive snapshot circuits to be located with the MAC, as illustrated in Figure 7.2a. Vendors may elect to further reduce timing jitter by latching the receive/transmit times within the PHY, where the uncertain FIFO latencies can be best avoided.



a) Simple clockSync snapshots

b) Precise clockSync snapshots

Figure 7.2—Timer snapshot locations

7.2.4 Rate-difference effects

If the absence of rate adjustments, significant *timeOfDay* errors can accumulate between send-period updates, as illustrated on the left side of Figure 7.3. The 2 µs deviation is due to the cumulative effect of clock drift, over the 10 ms send-period interval, assuming clock-master and clock-slave crystal deviations of −100 PPM and +100 PPM respectively.

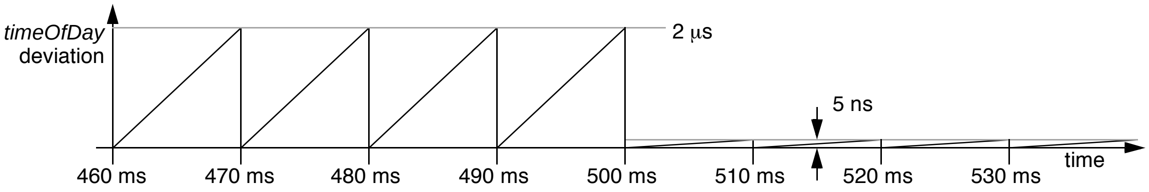


Figure 7.3—Rate-adjustment effects

While this regular sawtooth is illustrated as a highly regular (and thus perhaps easily filtered) function, irregularities could be introduced by changes in the relative ordering of clock-master and clock-slave transmissions, or transmission delays invoked by asynchronous frame transmissions. Tracking peaks/valleys or filtering such irregular functions are thought unlikely to yield similar *timeOfDay* deviation reductions.

The differences in rates could easily be reduced to less than 1 PPM, assuming a 200 ms measurement interval (based on a 100 ms slow-period interval) and a 100 ns arrival/departure sampling error. A clock-rate adjustment at time 100 ms could thus reduce the clock-drift related errors to less than 5 ns. At this point, the timer-offset measurement errors (not clock-drift induced errors) dominate the clock-synchronization error contributions.

The preceding discussion illustrates the need for timer-rate (as well as timer-offset) adjustments. To avoid PLL-chain whiplash type of effects, the rates are computed as follows:

a) Each clock-slave station computes the rate difference between its clock and its clock master.
b) Each clock-slave station adds its rate difference to the cumulative rate difference; that result is passed to its attached clock-slave stations.

Detailed descriptions of these rate-adjustment behaviors are TBD.

7.2.5 Clock-synchronization intervals

Clock synchronization involves synchronizing the clock-slave clocks to the reference provided by the grand clock master. Tight accuracy is possible with matched-length duplex links, since bidirectional messages can cancel the cable-delay effects.

Clock synchronization involves the processing of periodic events. Multiple time periods are involved, as listed in Table 7.1. The clock-period events trigger the update of free-running timer values; the period affects the timer-synchronization accuracy and is therefore constrained to be small.

The send-period events trigger the interchange of timeSync frames between adjacent stations. While a smaller period (1 ms or 100 µs) could improve accuracies, the larger value is intended to reduce costs by allowing computations to be executed by inexpensive (but possibly slow) bridge-resident firmware.

The slow-period events trigger the computation of timer-rate differences. The timer-rate differences are computed over two slow-period intervals, but recomputed every slow-period interval. The larger 100 ms (as opposed to 10 ms) computation interval is intended to reduce errors associated with sampling of clock-period-quantized slow-period-sized time intervals.

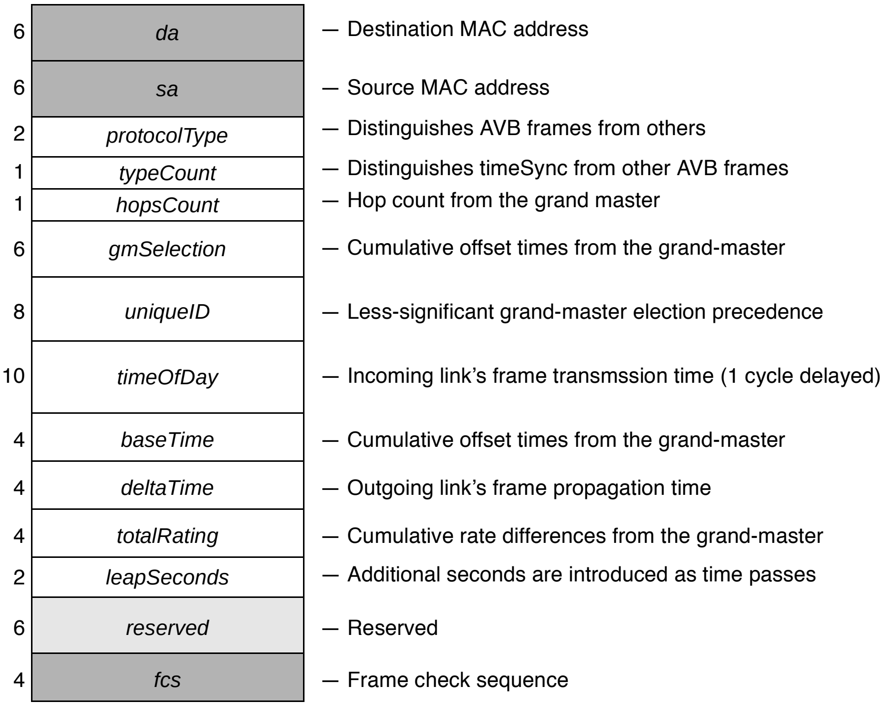
Table 7.1—Clock-synchronization intervals

Name	Time	Description
clock-period	< 20 ns	Resolution of timer-register value updates
send-period	10 ms	Time between sending of periodic timeSync frames between adjacent stations
slow-period	100 ms	Time between computation of clock-master/clock-slave rate differences

7.3 timeSync frame format

7.3.1 timeSync fields

Clock synchronization (timeSync) frames facilitate the synchronization of neighboring clock span-master and clock span-slave stations. The frame, which is normally sent once each isochronous cycle, includes time-snapshot information and the identity of the network's clock master, as illustrated in Figure 7.4. The gray boxes represent physical layer encapsulation fields that are common across Ethernet frames.



6	*da*	— Destination MAC address
6	*sa*	— Source MAC address
2	*protocolType*	— Distinguishes AVB frames from others
1	*typeCount*	— Distinguishes timeSync from other AVB frames
1	*hopsCount*	— Hop count from the grand master
6	*gmSelection*	— Cumulative offset times from the grand-master
8	*uniqueID*	— Less-significant grand-master election precedence
10	*timeOfDay*	— Incoming link's frame transmssion time (1 cycle delayed)
4	*baseTime*	— Cumulative offset times from the grand-master
4	*deltaTime*	— Outgoing link's frame propagation time
4	*totalRating*	— Cumulative rate differences from the grand-master
2	*leapSeconds*	— Additional seconds are introduced as time passes
6	*reserved*	— Reserved
4	*fcs*	— Frame check sequence

Figure 7.4—timeSync frame format

7.3.1.1 ***da*****:** A 48-bit (destination address) field that specifies the station(s) for which the frame is intended. The *da* field contains either an individual or a group 48-bit MAC address (see 3.11), as specified in 9.2 of IEEE Std 802-2001.

7.3.1.2 ***sa*****:** A 48-bit (source address) field that specifies the local station sending the frame. The *sa* field contains an individual 48-bit MAC address (see 3.11), as specified in 9.2 of IEEE Std 802-2001.

7.3.1.3 ***protocolType*****:** A 16-bit field contained within the payload that identifies the format and function of the following fields.

7.3.1.4 *typeCount*: An 8-bit field that identifies the format and function of the following fields (see 7.3.2).

7.3.1.5 *hopsCount*: An 8-bit field that identifies the maximum number of hops between the talker and associated listeners.

7.3.1.6 *gmSelection*: A 6-byte field that has specifies precedence in the grand-master selection protocols (see 7.3.3).

7.3.1.7 *uniqueID*: A 64-bit field that uniquely identifies the contending grand clock master (see 7.3.4).

7.3.1.8 *timeOfDay*: An 80-bit field that specifies the time within the source station when the previous timeSync frame was transmitted (see 7.3.5).

7.3.1.9 *baseTime*: A 32-bit field that specifies the fractions-off-second offset time within the source station (see 7.3.6).

7.3.1.10 *deltaTime*: A 32-bit field that specifies the differences between timeSync receive and transmit times, as measured in fractions-of-second on the opposing link (see 7.3.6).

7.3.1.11 *totalRating*: A 32-bit field that specifies a scaled version of the *diffRate* value. The *diffRate* value transmitted by station *n* represents the rate difference between station *n* and the grand master.

7.3.1.12 *leapSeconds*: A 16-bit field that specifies the number of seconds that should be added to the *timeOfDay* value, when converting between xx and yy values. (On IEEE-1588, this is called the *UTCOffset* field.)

7.3.1.13*fcs*: A 32-bit (frame check sequence) field that is a cyclic redundancy check (CRC) of the frame.

7.3.2 *typeCount* subfields

The 8-bit *typeCount* field provides a subtype identifier and a sequence number for detecting lost-frame transmissions, as illustrated in Figure 7.5.

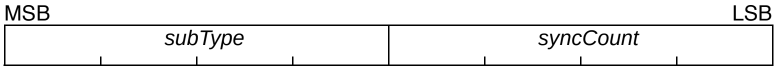


Figure 7.5—*typeCount* format

7.3.2.1 *subCount*: A 4-bit field that distinguishes the timeSync frame from other frames with the same *protocolType* field.

7.3.2.2 *syncCount*: A 4-bit field that is incremented on each *timeSync*-frame transmission.

7.3.3 *gmSelection* subfields

The 48-bit *gmSelection* field provides precedence information of the current grand master, as illustrated in Figure 7.6.

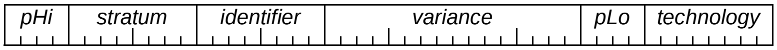


Figure 7.6—*gmSelection* format

7.3.3.1 *pHi*: A 4-bit field that can be configured by the user and overrides the remaining *gmSelection*-resident precedence fields.

7.3.3.2 *stratum*: An 8-bit precedence-selection field defined by the like-named IEEE-1588 field.

7.3.3.3 *identifier*: An 8-bit precedence-selection field defined by the like-named IEEE-1588 field.

7.3.3.4 *variance*: A 16-bit precedence-selection field defined by the like-named IEEE-1588 field.

7.3.3.5 *pLo*: A 4-bit field that can be configured by the user and overrides the remaining *gmSelection*-resident precedence fields.

7.3.3.6 *technology*: An 8-bit field that identifies the format of the 8-byte *uniqueID* field.

7.3.4 *uniqueID* subfields

The 64-bit *uniqueID* field is a unique identifier. For stations that have a uniquely assigned 48-bit *macAddress*, the 64-bit *uniqueID* field is derived from the 48-bit MAC address, as illustrated in Figure 7.7.

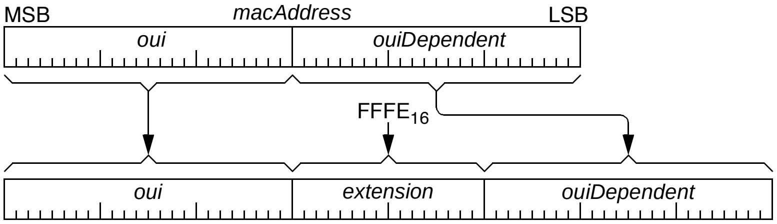


Figure 7.7—*uniqueID* format

7.3.4.1 *oui*: A 24-bit field assigned by the IEEE/RAC (see xx).

7.3.4.2 *extension*: A 16-bit field assigned to encapsulated EUI-48 values.

7.3.4.3 *ouiDependent*: A 24-bit field assigned by the owner of the *oui* field (see xx).

7.3.5 timeOfDay subfield formats

Time-of-day values within a frame are based on seconds and fractions-of-second values, consistent with IETF specified NTP[B7] and SNTP[B8] protocols, as illustrated in Figure 7.8.

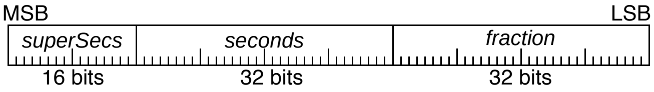


Figure 7.8—Complete seconds timer format

7.3.5.1 ***superSecs*****:** A 16-bit field that extends the range of the *seconds* field.

7.3.5.2 ***seconds*****:** A 32-bit field that specifies time in seconds.

7.3.5.3 ***fraction*****:** A 32-bit field that specified time offset within the *second*, in units of 2^{-32} second.

The concatenation of thes fields specifies a 96-bit *timeOfDay* value, as specified by Equation 7.5.

$$time = superSecs*2^{32} + seconds + (fraction / 2^{32}) \tag{7.5}$$

7.3.6 Local time formats

The local-time values within a frame are based on a fractions-of-second value, as illustrated in Figure 7.9. The 32-bit *fraction* field specifies the time offset within the *second*, in units of 2^{-32} second.

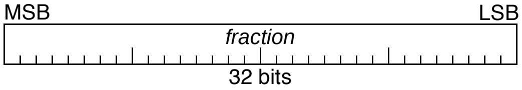


Figure 7.9—Complete seconds timer format

Annexes

Annex A

(informative)

Bibliography

[B1] IEEE 100, The Authoritative Dictionary of IEEE Standards Terms, Seventh Edition.[1]

[B2] IEEE Std 802-2002, IEEE Standards for Local and Metropolitan Area Networks: Overview and Architecture.

[B3] IEEE Std 801-2001, IEEE Standard for Local and Metropolitan Area Networks: Overview and Architecture.

[B4] IEEE Std 802.1D-2004, IEEE Standard for Local and Metropolitan Area Networks: Media Access Control (MAC) Bridges.

[B5] IEEE Std 1394-1995, High performance serial bus.

[B6] IEEE Std 1588-2002, IEEE Standard for a Precision Clock Synchronization Protocol for Networked Measurement and Control Systems.

[B7] IETF RFC 1305: Network Time Protocol (Version 3) Specification, Implementation and Analysis, David L. Mills, March 1992[2]

[B8] IETF RFC 2030: Simple Network Time Protocol (SNTP) Version 4 for IPv4, IPv6 and OSI, D. Mills, October 1996.

[1] IEEE publications are available from the Institute of Electrical and Electronics Engineers, 445 Hoes Lane, P.O. Box 1331, Piscataway, NJ 08855-1331, USA (http://standards.ieee.org/).

[2] IETF publications are available via the World Wide Web at http://www.ietf.org.

Annex B

(informative)

Bridging to IEEE Std 1394

To illustrate the sufficiency and viability of the AVB time-synchronization services, the transformation of IEEE 1394 packets is illustrated.

B.1 Hybrid network topologies

B.1.1 Supported IEEE 1394 network topologies

This annex focuses on the use of AVB to bridge between IEEE 1394 domains, as illustrated in Figure B.1. The boundary between domains is illustrated by a dotted line, which passes through a SerialBus adapter station.

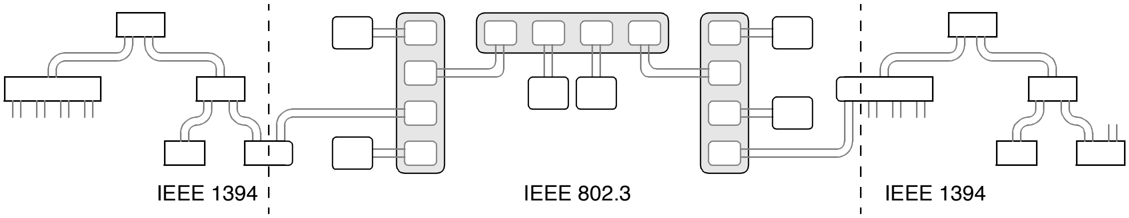


Figure B.1—IEEE 1394 leaf domains

B.1.2 Unsupported IEEE 1394 network topologies

Another approach would be to use IEEE 1394 to bridge between IEEE 802.3 domains, as illustrated in Figure B.2. While not explicitly prohibited, architectural features of such topologies are beyond the scope of this working paper.

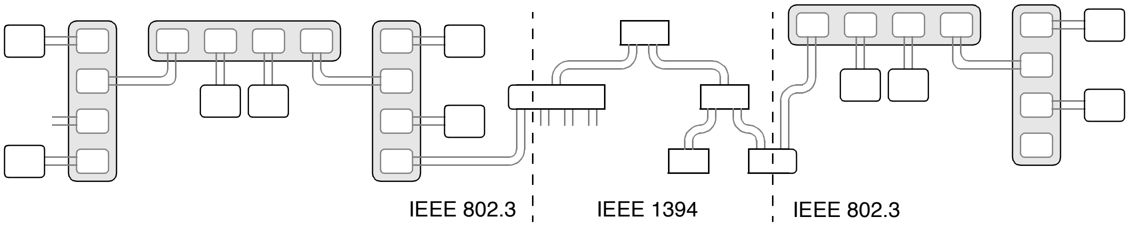
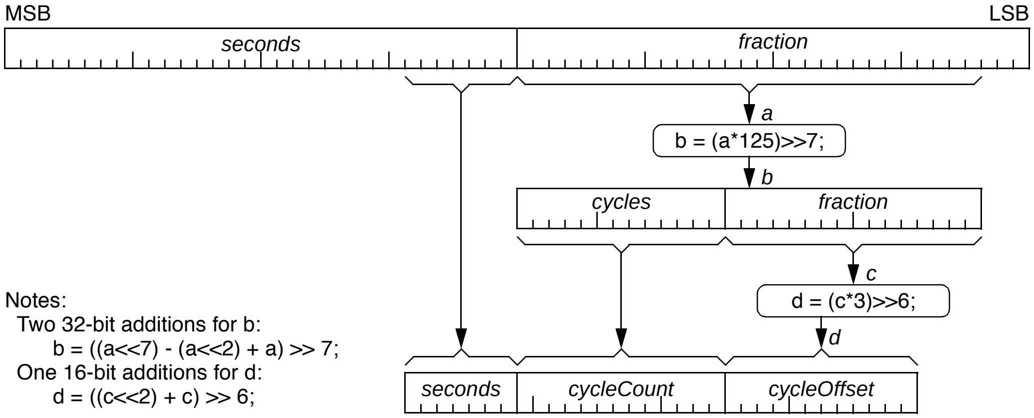


Figure B.2—IEEE 802.3 leaf domains

B.1.3 Time-of-day format conversions

The difference between AVB and IEEE 1394 time-of-day formats is expected to require conversions within the AVB-to-1394 adapter. Although multiplies are involved in such conversions, multiplications by constants are simpler than multiplications by variables. For example, a conversion between AVB and IEEE 1394 involves no more than two 32-bit additions and one 16-bit addition, as illustrated in Figure B.3.



Notes:
Two 32-bit additions for b:
b = ((a<<7) - (a<<2) + a) >> 7;
One 16-bit additions for d:
d = ((c<<2) + c) >> 6;

Figure B.3—Time-of-day format conversions

B.1.4 Grand-master precedence mappings

Compatible formats allow either an IEEE 1394 or IEEE 802.3 stations to become the network's grand-master station. While difference in format are present, each format can be readily mapped to the other, as illustrated in Figure B.4:

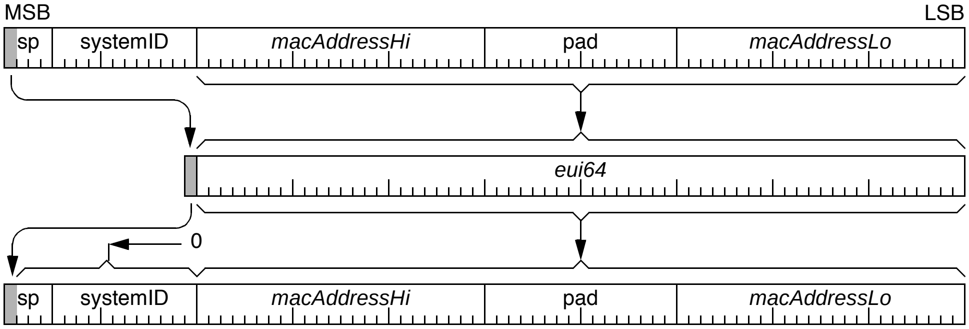


Figure B.4—Grand-master precedence mapping

Annex C

(informative)

Review of possible alternatives

C.1 Clock-synchronization alternatives

NOTE—This tables has not been reviewed for considerable time and is thus believed to be inaccurate. However, the list is being maintained (until it can be updated) for its usefulness as talking points.

A comparison of the AVB and IEEE 1588 time-synchronization proposals is summarized in Table C.1.

Table C.1Protocol comparison

Properties		**Row**	**Descriptinos**	
state			**AVB-SG**	**1588**
timeSync MTU <= Ethernet MTU		1	yes	
No cascaded PLL whiplash		2	yes	
Number of frame types		3	1	> 1
Phaseless initialization sequencing		4	yes	no
Topology		5	duplex links	general
Grand-master precedence parameters		6	spanning-tree like	special
Rogue-frame settling time, per hop		7	10 ms	1 s
Arithmetic complexity	numbers	8	64-bit binary	2 x 32-bit binary
	negatives	9	2’s complement	signed
Master transfer discontinuities	rate	10	gradual change	
	offset limitations	11	duplex-cable match sampling error	
Firmware friendly	no delay constraints	12	yes	
	n-1 cycle sampling	13	yes	
Time-of-day value precision	offset resolution	14	233 ps	
	overflow interval	15	136 years	

Row 1: The size of a timeSync frame should be no larger than an Ethernet MTU, to minimize overhead.
AVB-SG: The size of a timeSync frame is an Ethernet MTU.
1588: The size of a timeSync frame is (to be provided).

Row 2: Cascaded phase-lock loops (PLLs) can yield undesirable whiplash responses to transients.
AVB-SG: There are no cascaded phase-lock loops.
1588: There are multiple initialization phases (to be provided).

Row 3: There number of frame types should be small, to reduce decoding and processing complexities.
AVB-SG: Only one form of timeSync frame is used.
1588: Multiple forms of timeSync frames are used (to be provided).

Row 4: Multiple initialization phases adds complexity, since miss-synchronized phases must be managed.
AVB-SG: There are no distinct initialization phases.
1588: There are multiple initialization phases (to be provided).

Row 5: Arbitrary interconnect topologies should be supported.
AVB-SG: Topologies are constrained to point-to-point full-duplex cabling.
1588: Supported topologies include broadcast interconnects.

Row 6: Grand-master selection precedence should be software configurable, like spanning-tree parameters.
AVB-SG: Grand-master selection parameters are based on spanning-tree parameter formats.
1588: Grand-master selection parameters are (to be provided).

Row 7: The lifetime of rogue frames should be minimized, to avoid long initialization sequences.
AVB-SG: Rogue frame lifetimes are limited by the 10 ms per-hop update latencies.
1588: Rogue frame lifetimes are limited by (to be provided).

Row 8: The time-of-day formats should be convenient for hardware/firmware processing.
AVB-SG: The time-of-day format is a 64-bit binary number.
1588: The time-of-day format is a (to be provided).

Row 9: The time-of-day negative-number formats should be convenient for hardware/firmware processing.
AVB-SG: The time-of-day format is a 2's complement binary number.
1588: The time-of-day format is a (to be provided).

Row 10: The rate discontinuities caused by grand-master selection changes should be minimal.
AVB-SG: Smooth rate-change transitions with a 2.5 second time constant is provided.
1588: (To be provided).

Row 11: The time-of-day discontinuities caused by grand-master selection changes should be minimal.
AVB-SG: Maximum time-of-day errors are limited by cable-length asymmetry and time-snapshot errors.
1588: (To be provided).

Row 12: Firmware friendly designs should not rely on fast response-time processing.
AVB-SG: Response processing time have no significant effect on time-synchronization accuracies.
1588: (To be provided).

Row 13: Firmware friendly designs should not rely on immediate or precomputed snapshot times.
AVB-SG: Snapshot times are never used within the current cycle, but saved for next-cycle transmission.
1588: (To be provided).

Row 14: The fine-grained time-of-day resolution should be small, to faciliate accurate synchronization.
AVB-SG: The 64-bit time-of-day timer resolution is 233 ps, less than expected snapshot accuracies.
1588: (To be provided).

Row 15: The time-of-day extent should be suffiently large to avoid overflows within one's lifetime.
AVB-SG: The 64-bit time-of-day timer overflows once every 136 years.
1588: (To be provided).

Annex D

(informative)

Time-of-day format considerations

To better understand the rationale behind the ‘extended binary’ timer format, other formats are evaluated and compared within this annex.

D.1 Possible time-of-day formats

D.1.1 Extended binary timer formats

The extended-binary timer format is used within this working paper and summarized herein. The 64-bit timer value consist of two components: a 32-bit *seconds* and 32-bit *fraction* fields, as illustrated in Figure 4.1.

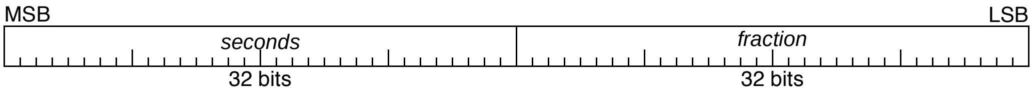


Figure 4.1—Complete seconds timer format

The concatenation of 32-bit *seconds* and 32-bit *fraction* field specifies a 64-bit *time* value, as specified by Equation D.1.

$$time = seconds + (fraction / 2^{32}) \tag{D.1}$$

Where:

seconds is the most significant component of the time value (see Figure 4.1).

fraction is the less significant component of the time value (see Figure 4.1).

D.1.2 IEEE 1394 timer format

An alternate “1394 timer” format consists of *secondCount*, *cycleCount*, and *cycleOffset* fields, as illustrated in Figure D.2. For such fields, the 12-bit *cycleOffset* field is updated at a 24.576MHz rate. The *cycleOffset* field goes to zero after 3171 is reached, thus cycling at an 8kHz rate. The 13-bit *cycleCount* field is incremented whenever *cycleOffset* goes to zero. The *cycleCount* field goes to zero after 7999 is reached, thus restarting at a 1Hz rate. The remaining 7-bit *secondCount* field is incremented whenever *cycleCount* goes to zero.

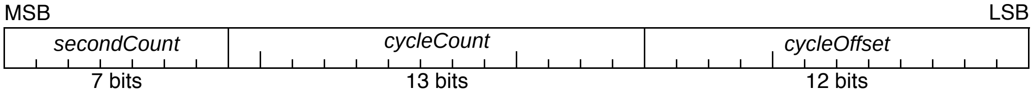


Figure D.2—IEEE 1394 timer format

D.1.3 IEEE 1588 timer format

IEEE 1588 timer format consists of seconds and nanoseconds fields components, as illustrated in Figure D.3. The nanoseconds field must be less than 10^9; a distinct *sign* bit indicates whether the time represents before or after the epoch duration.

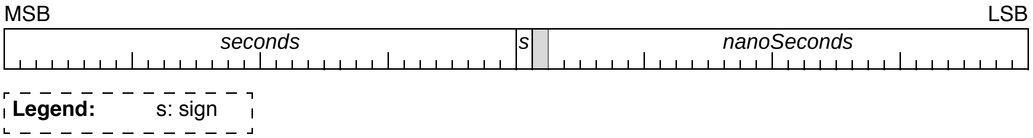


Figure D.3—IEEE 1588 timer format

D.1.4 EPON timer format

The IEEE 802.3 EPON timer format consists of a 32-bit scaled nanosecond value, as illustrated in Figure D.4. This clock is logically incremented once each 16 ns interval.

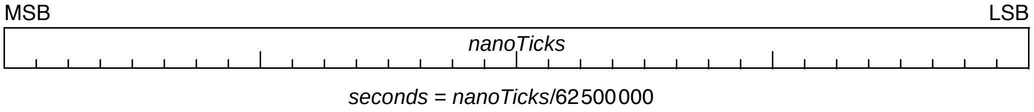


Figure D.4—EPON timer format

D.1.5 Compact seconds timer format

An alternate “compact seconds” format could consist of 8-bit *seconds* and 24-bit *fraction* fields, as illustrated in Figure D.5. This would provided similar resolutions to the IEEE 1394 timer format, without the complexities associated with its binary coded decimal (BCD) like encoding.

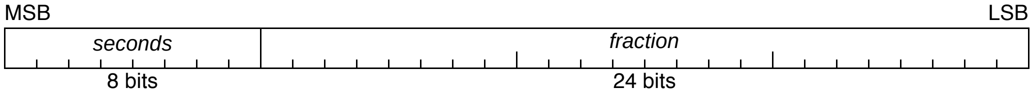


Figure D.5—Compact seconds timer format

D.1.6 Nanosecond timer format

An alternate “nanosecond” format could consists of 2-bit *seconds* and 30-bit *nanoSeconds* fields, as illustrated in Figure D.6.

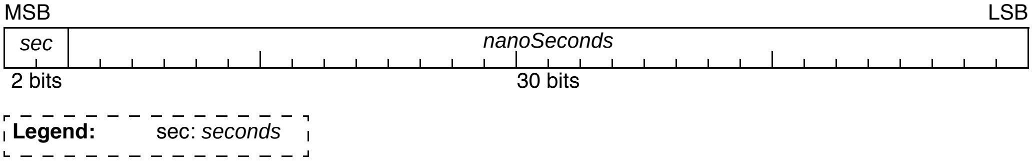


Figure D.6—Nanosecond timer format

D.2 Time format comparisons

To better understand the relative benefits of different time formats, the relevant properties are summarized in Table D.1. Counter complexity is not included in the comparison, since the digital logic complexity is comparable for all formats.

Table D.1—Time format comparison

Name	**Subclause**	**Range**	**Precision**	**Arithmetic**	**Seconds**	**Defined standards**
Column	—	1	2	3	4	5
extended binary	TBD	136 years	232 ps	Good	Good	RFC 1305 NTP, RFC 2030 SNTPv4
IEEE 1394	D.1.2	128 s	30 ns	Poor	Good	IEEE 1394
IEEE 1588	D.1.3	272 years	1 ns	Fair	Good	IEEE 1588
IEEE 802 (EPON)	D.1.4	69 s	16 ns	Good	Poor	IEEE 802.3
compact seconds	D.1.5	256 s	60 ns	Best	Good	—
nanoseconds	D.1.6	4 s	1 ns	Best	Poor	—

Column 1: A desirable property is the support of a wide range of second values, to eliminate the need for defining/coordinating/implementing auxiliary seconds-synchronization protocols. The 136-year range of the extended binary format is sufficient for this purpose.

Column 2: A desirable property is a fine-grained resolution, sufficient to measure each bit-transmission times. The ‘extened binary’ provides the most precision; exceeds the resolution of expected cost-effective time-capture circuits.

Column 3: Computation of time differences involves the subraction of two timer-snapshot values. Subtraction of ‘extended binary’ numbers involving standard 64-bit binary arithmetic; no special field-overlow compensations are required. Only the less precise ‘compact seconds’ and nanoseconds formats are simpler, due to the reduced 32-bit size of the timer values.

Column 4: Time values must oftentimes be compared to externally provided values (e.g., timers extracted from GPS or stratum-clock sources). For these purposes, the availability of a seconds component is desired. The ‘extended binary’ format provides a seconds component that can be easily extracted or such purposes.

Index

B

baseTime
- *See* clockSync frame

C

clockSync frame
- *da* 42
- *sa* 42
- *protocolType* 42
- *typeCount* 43
  - *subType* 43
  - *syncCount* 43
- *hopsCount* 43
- *gmSelection* 43
  - *pHi* 44
  - *stratum* 44
  - *identifier* 44
  - *variance* 44
  - *pLo* 44
  - *technlology* 44
- *uniqueID* 43
  - *oui* 44
  - *extension* 44
  - *ouiDependent* 44
- *timeOfDay* 43
  - *superSecs* 45
  - *seconds* 45
  - *fraction* 45
- *baseTime* 43
- *deltaTime* 43
- *totalRating* 43
- *leapSeconds* 43
- *fcs* 43

D

da
- *See* clockSync frame

deltaTime
- *See* clockSync frame

E

extension
- *See* clockSync frame
- *See uniqueID*subfields

F

fcs
- *See* clockSync frame

fraction
- *See* clockSync frame
- *See timeOfDay* subfields

G

gmSelection
- *See* clockSync frame

gmSelection subfields
- *pHi* 44
- *stratum* 44
- *identifier* 44
- *variance* 44
- *pLo* 44
- *technology* 44

H

hopsCount
- *See* clockSync frame

I

identifier
- *See* clockSync frame
- *See gmSelection*subfields

L

leapSeconds
- *See* clockSync frame

O

oui
- *See* clockSync frame
- *See uniqueID*subfields

ouiDependent
- *See* clockSync frame
- *See uniqueID*subfields

P

pHi
- *See* clockSync frame
- *See gmSelection*subfields

pLo
- *See* clockSync frame
- *See gmSelection*subfields

protocolType
- *See* clockSync frame

S

sa
- *See* clockSync frame

seconds
- *See* clockSync frame
- *See timeOfDay* subfields

stratum
- *See gmSelection*subfields

subCount
- *See* clockSync frame

subType
- *See typeCount*subfields

superSecs
- *See* clockSync frame
- *See timeOfDay*subfields

syncCount
- *See* clockSync frame
- *See typeCount*subfields

T

technology
- *See* clockSync frame
- *See gmSelection*subfields

timeOfDay
- *See* clockSync frame

timeOfDay subfields
- *superSecs* 45
- *seconds*.. 45
- *fraction*.. 45

totalRatinge
- *See* clockSync frame

typeCount
- *See* clockSync frame

typeCount subfields
- *subType*... 43
- *syncCount* 43

U

uniqueID
- *See* clockSync frame

uniqueID subfields
- *extension* 44
- *oui* 44
- *ouiDependent*... 44

V

variance
- *See* clockSync frame
- *See gmSelection*subfields

Annex E

(informative)

C-code illustrations

NOTE—This annex is provided as a placeholder for illustrative C-code. Locating the C code in one location (as opposed to distributed throughout the working paper) is intended to simplify its review, extraction, compilation, and execution by critical reviewers.
Also, placing this code in a distinct Annex allows the code to be conveniently formatted in 132-character landscape mode. This eliminates the need to truncate variable names and comments, so that the resulting code can be better understood by the reader.

This Annex provides code examples that illustrate the behavior of AVB entities. The code in this Annex is purely for informational purposes, and should not be construed as mandating any particular implementation. In the event of a conflict between the contents of this Annex and another normative portion of this standard, the other normative portion shall take precedence.

The syntax used for the following code examples conforms to ANSI X3T9-1995.

```
// ***********************************************************************************************************************************
//                                                                                                 1         1         1         1
//       1         2         3         4         5         6         7         8         9         0         1         2         3
//345678901234567890123456789012345678901234567890123456789012345678901234567890123456789012345678901234567890123456789012345678901234

#include <assert.h>
#include <stdio.h>

// unsigned char      uint8_t;                                                 // 1-byte unsigned integer
// unsigned short     uint16_t;                                                // 2-byte unsigned integer
// unsigned int       uint32_t;                                                // 4-byte unsigned integer
// unsigned long long uint64_t;                                                // 8-byte unsigned integer

// signed char        int8_t;                                                  // 1-byte signed integer
// signed short       int16_t;                                                 // 2-byte signed integer
// signed int         int32_t;                                                 // 4-byte signed integer
// signed long long   int64_t;                                                 // 8-byte signed integer

#define OPTION_FAST 0                                                          // 1 if precedence-change is very-quick
#define OPTION_BASE 0                                                          // 1 if baseTimer hardware is provided
#define DIFF_SCALE ((double)4096 * ((uint64_t)1 << 31))                        // Changes <200PPM to a 32-bit signed integer
#define EXTRACT_CORE(a, b) (((a) << 32) | ((b) >> 32))                         // Extract seconds and fraction component
#define FULL_SCALE (0x7FFFFFFF)                                                // Biggest 32-bit positive integer
#define LIMIT(a, b, c) MAX(MIN((a), (b)), (c))                                 // Force base/bounds constraints
#define MAX(a, b) ((a) < (b) ? (b) : (a))                                      // Maximum value definition
#define MIN(a, b) ((a) > (b) ? (b) : (a))                                      // Minimum value definition
#define MINIMUM_WIDE(a, b) (CompareWide((a), (b)) < 0 ? (a) : (b))
#define ONES64 ~((uint64_t)0)                                                  // 64-bit all-ones value
#define SCALE64 ((double)16 * (1 << 30) * (1 << 30))                           // Floating-point equivalent of (1<<64)
#define BASE_PORT(siPtr) (siPtr->portPtr)                                      // A pointer to the first station port
#define NEXT_PORT(piPtr) (piPtr->portPtr)                                      // A pointer to the next station port
#define GRAND 2                                                                // An indication of grand-master mode
#define SLAVE 1                                                                // If not grand-master, slave mode


// The grand-master precedence check is based on concatenated fields, as follows:
//
//   MSB                                                                                     LSB
//   |                    hi                     |                      lo                     |
//   +-----------------------------------------------------------------------------------------+
//   |   0000    systemTag                     eui64                       00   hops   portTag |
//   +----16----’----16----’--------------------64-----------------------’--8--’--8--’----16---+
//
// If hops == ONES, this value is considered VOID and has the worse precedence
// Otherwise, the best precedence corresponds to the smallest of two tested values.
//

#if (CPU_TYPE == BIG)
typedef struct
{
    uint64_t hi;                                 // more-significant portion
    uint64_t lo;                                 // less-significant portion
} DoubleData;
typedef struct
```

```
{
    unsigned fill16:16;
    unsigned systemTag:16;
    unsigned uniqueHi:32;
    unsigned uniqueLo:32;
    unsigned fill08:8;
    unsigned hopsCount:8;
    unsigned portLevel:4;
    unsigned portNumber:12;
} DoubleInfo;
#else
typedef struct
{
    uint64_t lo;                                 // less-significant portion
    uint64_t hi;                                 // more-significant portion
} DoubleData;
typedef struct
{
    unsigned portNumber:12;
    unsigned portLevel:4;
    unsigned hopsCount:8
    unsigned fill08:8;
    unsigned uniqueLo:32;
    unsigned uniqueHi:32;
    unsigned systemTag:16;
    unsigned fill16:16;
} DoubleInfo;
#endif

typedef union
{
    DoubleData data;                             // As 64-bit data values
    DoubleInfo info;                             // As data fields
} PrecedenceInfo;

typedef struct _PortInfo
{
    struct _PortInfo *portPtr;                   // Points to the next linked port
    unsigned portLevel:4;                        // Relative priority number of ports
    unsigned portNumber:12;                      // Port number
    DoubleData portPrecedence;                   // Incoming frame parameters

    uint8_t  skipCount;                          // Number of 10ms intervals
    uint32_t cableDelay;                         // The cable delay, from local master
    uint32_t linkOffset;                         // The cable difference, from local master
    uint64_t deltaTime;                          // For inclusion in transmitted frames

                                                 // Best if captured accurately by the PHY
    uint64_t latchRxFlexTime;                    // Snapshot of flexTimer, on clockSync arrival,
                                                 // available for this clockSync reception.
    uint64_t latchTxFlexTime;                    // Snapshot of flexTimer, on clockSync departure,
                                                 // available for next clockSync transmission

    uint64_t savedRxFlexTime;                    // Previous latchRxFlexTime value
    uint64_t savedRxFlexData;                    // Previous clockSync.lastFlexTime value
```

```
                                                 // If OPTION_BASE is 1, baseTimer snapshots:
    uint32_t latchRxBaseTime;                    // Captured clockSync base-time arrival
    uint32_t latchTxBaseTime;                    // Captured clockSync base-time departure (delayed)
} PortInfo;

typedef struct
{                                                // Customized per-station components
    PortInfo *portPtr;                           // Points to a linked-list of ports
    double nominalFrequency;                     // Nominal clock frequency, in Hz
    int8_t clockDeviation;                       // Deviation in parts-per-million
    uint64_t eui64;                              // 64-bit extended unique identifier
    unsigned systemLevel:4;                      // Relative priority number of ports
    unsigned systemNumber:12;                    // Port number

    unsigned selectCount;                        // Grand-master selection count

    uint8_t  skipCount;                          // Number of 10ms intervals
    uint32_t myDiffRate;                         // The rate difference, from upstream neighbor
    uint32_t diffRate;                           // The rate difference, from grand-master
    uint32_t linkOffset;                         // The cable difference, from local master
    uint64_t deltaTime;                          // For inclusion in transmitted frames

    DoubleData thisPrecedence;                   // The precedence of this station
    DoubleData bestPrecedence;                   // The best observed precedence
    int16_t bestPort;                            // Selected clock-slave port

    uint64_t savedRxFlexTime;                    // Previous latchRxFlexTime value
    uint32_t savedRxBaseTime;                    // Previous latchRxBaseTime value

    uint64_t timeOfDay;                          // Offset and rate-compensated timer value
    uint64_t flexTimerHi;                        // Offset and rate adjustable 64-bit timer
    uint64_t flexTimerLo;                        // Offset and rate adjustable 64-bit timer
    uint64_t flexOffset;                         // Adjustable offset value for flexTimer
    uint64_t flexRate;                           // 40-bit adjustable rate for flexTimer

    uint64_t baseTimer;                          // Fixed-rate fixed-offset 64-bit timer
    uint64_t baseRate;                           // SCALE64/clockFrequency, pre-initialized

    uint32_t savedRxBaseTickTime;                // Saved values of savedRxBaseTime
    uint32_t savedRxBaseTickData;                // Saved values of clockSync.lastBaseTime;
} StationInfo;

typedef struct                                   // The clockSync frame, reserved-padded to
{                                                // the minimum 64-byte frame size.
    uint32_t  da_hi;                             // Ethernet’s 48-bit destination address
    uint16_t  da_lo;                             //  "
    uint16_t  sa_hi;                             // Ethernet’s 48-bit source address
    uint32_t  sa_lo;                             //  "
    uint16_t  protocolType;                      // Specifies format/meaning of following
    uint8_t   subType;                           // Refined format/meaning specification
    uint8_t   syncCount;                         // Sequence numbers for consistency checks
    uint8_t   hopsCount;                         // Hop counts from the grand master
    uint8_t   reserved;                          // A few reserved bytes, for 64-byte minimum
    uint16_t  systemTag;                         // Precedence for grand-master election
    uint64_t  uniqueID;                          // Identifier for grand-master election
    uint64_t  lastFlexTime;                      // flexTimer on last clockSync transmission
    uint64_t  deltaTime;                         // Time difference on opposing link
    uint64_t  offsetTime;                        // Cumulative grand-master offset differences
    uint32_t  diffRate;                          // Cumulate grand-master rate differences
    uint32_t  lastBaseTime;                      // baseTimer on last clockSync transmission
    uint32_t  fcs;                               // Frame check sequence
} ClockSyncFrame;

uint32_t   BaseTimerChange(uint64_t, uint64_t, double);
void       ClockSyncArrived(StationInfo *, PortInfo *);
void       ClockSyncDeparted(StationInfo *, PortInfo *);
void       ClockSyncReceive(StationInfo *, PortInfo *, ClockSyncFrame *, uint8_t);
void       ClockSyncTransmit(StationInfo *, PortInfo *, ClockSyncFrame *);
int        CompareWide(DoubleData, DoubleData);
DoubleData PrecedenceMerge(uint16_t, uint64_t, uint8_t, uint8_t, uint16_t);
void       TimerTick(StationInfo *);
int        UpdatePrecedence(StationInfo *, PortInfo *);


// Called with:
//   stationInfoPtr  -- the station information context
void
StationSetup(StationInfo *stationInfoPtr)
{
    PortInfo *portPtr;
    StationInfo *siPtr = stationInfoPtr;
    uint16_t systemTag;
    uint16_t count;

    assert(siPtr != NULL);
    siPtr->baseRate = SCALE64 / siPtr->nominalFrequency;                                           // The nominal frequency
    siPtr->systemLevel = 0X8;                                                                      // Mid-range default with
    systemTag = ((uint16_t)(siPtr->systemLevel) << 12) | siPtr->systemNumber;                      // systemNumber extension
    siPtr->thisPrecedence =                                                                        // This station precedence
     PrecedenceMerge(systemTag, siPtr->eui64, 0, 0, 0);                                               // has zero-valued hopsCount
    count = 0;
    for (portPtr = BASE_PORT(siPtr); portPtr != NULL; portPtr = NEXT_PORT(portPtr)) {              // Per-port initialization
        portPtr->portLevel = 0X8;                                                                  // Mid-range default with
        portPtr->portNumber = count;                                                               // port-number extension
        count += 1;
    }
}

// Called with:
//   stationInfoPtr  -- the station information context
//   portInfoPtr     -- the port information context
//   clockSyncPtr -- the contents of a clockSync frame
void
ClockSyncReceive(StationInfo *stationInfoPtr, PortInfo *portInfoPtr, ClockSyncFrame *clockSyncPtr, uint8_t rateAdjust)
{
    PortInfo *piPtr = portInfoPtr, *portPtr;
    PrecedenceInfo precedence;
    StationInfo *siPtr = stationInfoPtr;
    ClockSyncFrame *csPtr = clockSyncPtr;
```

```
uint32_t measuredDelta, receivedDelta, diffRate;
uint64_t rxDelta, txDelta, clockDelta, cableDelay;
double tempRate;
int8_t grand, slave;

assert(siPtr != NULL && piPtr != NULL);
if (csPtr != NULL && csPtr->hopsCount != 0XFF)                                                 // Compute the precedence,
    piPtr->portPrecedence = PrecedenceMerge(csPtr->systemTag,                                  // in the absence of timeouts
      csPtr->uniqueID, csPtr->hopsCount, piPtr->portLevel, piPtr->portNumber);
else                                                                                           // Compute the precedence,
    piPtr->portPrecedence.hi = piPtr->portPrecedence.lo = ~((uint64_t)0);                      // in the presence of timeouts

if (OPTION_FAST && UpdatePrecedence(siPtr, piPtr)) {
    for (portPtr = BASE_PORT(siPtr); portPtr != NULL; portPtr = NEXT_PORT(portPtr))
        siPtr->bestPrecedence = MINIMUM_WIDE(piPtr->portPrecedence, siPtr->bestPrecedence);
    siPtr->selectCount += 1;
}
if (csPtr == NULL)
    return;

rxDelta = piPtr->savedRxFlexTime - csPtr->lastFlexTime;                                        // Measured receive-link delay
txDelta = csPtr->deltaTime;                                                                    // Reported transmit-link delay
clockDelta = (txDelta - rxDelta)/2;                                                            // Local timer differences
cableDelay = (txDelta + rxDelta)/2;                                                            // Cable transmission delay

precedence.data = siPtr->bestPrecedence;
grand = (precedence.info.hopsCount == 0) ? GRAND : 0;                                          // Grand-master properties
slave = (precedence.info.portNumber == piPtr->portNumber) ? SLAVE : 0;                         // Slave port identification
switch(grand | slave) {
case GRAND:                                                                                    // Grand-master properties
case GRAND | SLAVE:                                                                            // override slave-port ID
    siPtr->diffRate = 0;                                                                       // Grand-master reference
    siPtr->flexRate = siPtr->baseRate;                                                         // runs at the base rate
    break;
case SLAVE:
    if (rateAdjust) {                                                                          // Low-rate adjustments
        measuredDelta = (siPtr->savedRxBaseTime - siPtr->savedRxBaseTickTime);                 // Clock-slave difference
        receivedDelta = (csPtr->lastBaseTime - siPtr->savedRxBaseTickData);                    // Clock-master difference
        siPtr->savedRxBaseTickTime = siPtr->savedRxBaseTime;                                   // Previous saved value
        siPtr->savedRxBaseTickData = csPtr->lastBaseTime;                                      // Previous saved value
        tempRate = DIFF_SCALE * ((double)(receivedDelta - measuredDelta)/receivedDelta);       // Local rate difference
        siPtr->myDiffRate = LIMIT(tempRate, FULL_SCALE, -FULL_SCALE);                          // Rate difference limits
    }
    siPtr->diffRate = diffRate =
     LIMIT(siPtr->myDiffRate + csPtr->diffRate, FULL_SCALE, -FULL_SCALE);                      // Rate-range limitation
    siPtr->flexOffset = csPtr->offsetTime + clockDelta + siPtr->linkOffset;                    // Offset compensation
    siPtr->flexRate = siPtr->baseRate + siPtr->baseRate * (diffRate / DIFF_SCALE);             // Rate compensation
    if (OPTION_BASE)
        siPtr->savedRxBaseTime = piPtr->latchRxBaseTime;
    else
        siPtr->savedRxBaseTime +=                                                              // Receiver’s baseTimer snapshot
         BaseTimerChange(siPtr->savedRxFlexTime, piPtr->latchRxFlexTime, diffRate);
    break;
default:
    break;
    }
    piPtr->cableDelay = cableDelay;                                                                // Local cable-delay knowledge
    piPtr->savedRxFlexTime = piPtr->latchRxFlexTime;                                               // Saved reference time
    piPtr->deltaTime = rxDelta;                                                                    // Saved for retransmission
}

// Called with:
//   stationInfoPtr  -- the station information context
//   portInfoPtr     -- the port information context
int
UpdatePrecedence(StationInfo *stationInfoPtr, PortInfo *portInfoPtr)
{
    PortInfo       *piPtr = portInfoPtr;
    DoubleData      pastPrecedence;
    PrecedenceInfo  precedence;
    StationInfo    *siPtr = stationInfoPtr;
                                                                                                   // Set grand-master precedence
    assert(siPtr != NULL && piPtr != NULL);
    precedence.data = pastPrecedence = siPtr->bestPrecedence;                                      // Compare precedence values
    if (piPtr->portNumber == precedence.info.portNumber)                                           // If this is the best port,
        siPtr->bestPrecedence = MINIMUM_WIDE(piPtr->portPrecedence, siPtr->thisPrecedence);        // update baseline precedence
    else                                                                                           // If this is not the best port,
        siPtr->bestPrecedence = MINIMUM_WIDE(piPtr->portPrecedence, siPtr->bestPrecedence);        // update overall precedence
    return(CompareWide(siPtr->bestPrecedence, pastPrecedence) != 0);                               // A precedence-change result
}

// Called with:
//   stationInfoPtr  -- the station information context
//   portInfoPtr     -- the port information context
//   clockSyncPtr    -- the contents of a clockSync frame
void
ClockSyncTransmit(StationInfo *stationInfoPtr, PortInfo *portInfoPtr, ClockSyncFrame *clockSyncPtr)
{
    ClockSyncFrame *csPtr = clockSyncPtr;
    PortInfo       *piPtr = portInfoPtr;
    PrecedenceInfo  precedence;
    StationInfo    *siPtr = stationInfoPtr;

    assert(siPtr != NULL && piPtr != NULL && csPtr != NULL);
    if (UpdatePrecedence(siPtr, piPtr))                                                            // If precedence has changed,
        siPtr->selectCount += 1;                                                                   // start fast transmissions

    // An absent baseTimer is emulated by properly scaling time differences,
    // measured from the last recorded received-clockSync event.
    //   - baseTime value was computed
    //   - a different normDiffRate value has taken effect
    if (!OPTION_BASE)
        piPtr->latchTxBaseTime = siPtr->savedRxBaseTime +                                          // Derived from latchTxFlexRate
          BaseTimerChange(siPtr->savedRxFlexTime, piPtr->latchTxFlexTime, siPtr->diffRate);

    precedence.data = siPtr->bestPrecedence;
    csPtr->hopsCount = precedence.info.hopsCount;                                                  // Increment hop-count value
    csPtr->systemTag = precedence.info.systemTag;                                                  // Supply systemTag values
    csPtr->uniqueID  = ((uint64_t)(precedence.info.uniqueHi) << 32) | precedence.info.uniqueLo;    // Unique number tie-breaker
    csPtr->lastFlexTime = piPtr->latchTxFlexTime;                                                  // Send last timer value
    csPtr->deltaTime    = piPtr->deltaTime;                                                        // Send received-link delay
    csPtr->lastBaseTime = piPtr->latchTxBaseTime;                                                  // Send last baseTimer value
    csPtr->offsetTime   = siPtr->flexOffset;                                                       // This station’s cumulative offset
    csPtr->diffRate     = siPtr->diffRate;                                                         // Send current diffRate value
}

// Called when a clockSync frame is received, to latch timer values.
// Latches timers are available when ClockSyncReceive() is called.
//
// Called with:
//   stationInfoPtr -- the station information context
//   portInfoPtr     -- the port information context
void
ClockSyncArrived(StationInfo *stationInfoPtr, PortInfo *portInfoPtr)
{
    PortInfo    *piPtr = portInfoPtr;
    StationInfo *siPtr = stationInfoPtr;

    assert(siPtr != NULL);
    piPtr->latchRxFlexTime = EXTRACT_CORE(siPtr->flexTimerHi, siPtr->flexTimerLo);                 // Latch seconds:fraction fields
    if (OPTION_BASE)                                                                               // If a baseTimer is present,
        piPtr->latchRxBaseTime = siPtr->baseTimer >> 32;                                           // latch its fraction field
}

// Called when a clockSync frame is transmitted, to latch timer values.
// Latches timers are available for the next ClockSyncTransmit() call.
//
// Called with:
//   stationInfoPtr -- the station information context
//   portInfoPtr     -- the port information context
void
ClockSyncDeparted(StationInfo *stationInfoPtr, PortInfo *portInfoPtr)
{
    PortInfo    *piPtr = portInfoPtr;
    StationInfo *siPtr = stationInfoPtr;

    assert(siPtr != NULL);
    piPtr->latchTxFlexTime = EXTRACT_CORE(siPtr->flexTimerHi, siPtr->flexTimerLo);                 // Latch seconds:fraction fields
    if (OPTION_BASE)                                                                               // If a baseTimer is present,
        piPtr->latchTxBaseTime = siPtr->baseTimer >> 32;                                           // latch its fraction field
}

// Called at a high clock rate (less than 20 ns) to update flexTimer and baseTimer (if present).
// This routine is intended to illustrate the computations involved in updating hardware timers;
// this code is _not_ expected to be incorporated into firmware.
//
//                                            +-----------------------------------------+
//              0000 0000 0000 0000(hex)      |       fraction           subfraction    | flexRate
//                       |                    +---------32---------.---------32---------+
//                       |                                         |
//   ____________________v____________________,____________________v____________________,
//  '-----------------------------------------'-----------------------------------------'
//  (                                   flexAdd:128                                     )
//   -----------------------------------------------------------------------------------
//                                          ^   |
//                                          |   v
//  +-----------------------------------------------------------------------------------+
//  |    superSeconds            seconds              fraction           subfraction    | flexTimer
//  +---------32---------'---------32---------'---------32---------'---------32---------+
//                        \_______________________________________/
//                                            |
//                                            v
//                       -----------------------------------------
//                      (               offsetAdd:64              )-->timeOfDay
//                       -----------------------------------------
//                                            ^
//                                            |
//                      +-----------------------------------------+
//                      |       seconds              fraction     | flexOffset
//                      +---------32---------'---------32---------+
//
//
//                                            +-----------------------------------------+
//                                            |       fraction           subfraction    | baseRate
//                                            +---------32---------.---------32---------+
//                                                                 |
//                                                                 v
//                                             -----------------------------------------
//                                            (               baseAdd:64                )
//                                             -----------------------------------------
//                                                               ^   |
//                                                               |   v
//                                            +-----------------------------------------+
//                                            |       fraction           subfraction    | baseTimer
//                                            +---------32---------'---------32---------+
//
//
// Called with:
//   stationInfoPtr -- the station information context
void
TimerTick(StationInfo *stationInfoPtr)
{
    StationInfo *siPtr = stationInfoPtr;
    int64_t pastTimerLo;

    assert(siPtr != NULL);
    pastTimerLo = siPtr->flexTimerLo;                                                               // Saved to detect overflows
    siPtr->flexTimerLo += siPtr->flexRate;                                                          // Addition of the subfractions
    siPtr->flexTimerHi += (pastTimerLo > siPtr->flexTimerLo) ? 1 : 0;                               // Propagate carry into seconds
    siPtr->timeOfDay = EXTRACT_CORE(siPtr->flexTimerHi, siPtr->flexTimerLo) + siPtr->flexOffset;    // Compensate for offset drifts
    if (OPTION_BASE)
        siPtr->baseTimer += siPtr->baseRate;
}

// Called with:
//   stationInfoPtr -- the station information context
//   currentTime    -- the current flexTimer value
//   diffRate       -- the scaled rate difference
uint32_t
```

```
BaseTimerChange(uint64_t lastTime, uint64_t nextTime, double diffRate)
{
    uint64_t delta;

    delta = nextTime - lastTime;                                                                   // Compute lapsed time
    delta -= delta * (diffRate / DIFF_SCALE);                                                      // Compensate by rate difference
    return(delta);                                                                                 // Return incremenal change
}

// Merge multiple fields into an 128-bit integer, for comparisons
// Called with:
//   systemTag      -- the 16-bit most-significant precedence subfield
//   uniqueID       -- the 64-bit unique identifier (EUI-64)
//   hopsCount      -- the hop-count distance from the grand master
//   portTag        -- the tag associated with the port
DoubleData
PrecedenceMerge(uint16_t systemTag, uint64_t uniqueID, uint8_t hopsCount, uint8_t portLevel, uint16_t portNumber)
{
    PrecedenceInfo result;

    result.info.fill16     = 0;
    result.info.systemTag  = systemTag;
    result.info.uniqueHi   = (uniqueID >> 32);
    result.info.uniqueLo   = uniqueID;
    result.info.fill08     = 0;
    result.info.hopsCount  = hopsCount;
    result.info.portLevel  = portLevel;
    result.info.portNumber = portNumber;
    return(result.data);
}

// Performs a comparison of 128-bit preceision unsigned values
// Called with:
//   a -- the first of two 128-bit values
//   b -- the final of two 128-bit values
int
CompareWide(DoubleData a, DoubleData b)
{
    if (a.hi != b.hi)
        return(a.hi > b.hi ? 1 : -1);
    if (a.lo != b.lo)
        return(b.lo > b.lo ? 1 : -1);
    return(0);
}
```