Synchronous Ethernet and IEEE 1588v2 Technology White Paper

Issue 01

Date 2012-10-30

Huawei Technologies Co., Ltd.

Address:	Huawei Industrial Base
	Bantian, Longgang
	Shenzhen 518129
	People's Republic of China
Website:	http://enterprise.huawei.com

About This Document

Keywords

BITS, GPS, IEEE 1588v2, synchronization

Abstract

Although the trend is toward all-IP networks and services, carriers are confronted with many difficulties in migrating from the legacy network to the next-generation network. Synchronization of the clock frequency and time is a key problem that carriers need to solve. This document describes clock synchronization on the synchronous Ethernet and the key technologies of IEEE 1588v2.

List of Acronyms

Acronym	Full Name	Description
1588v2	IEEE 1588v2	A clock synchronization algorithm drafted by the Institute of Electrical and Electronics Engineers (IEEE). The algorithm provides a standard for clock synchronization based on data packet transmission.
ACR	Adaptive Clock Recovery	
PDH	Plesiochronous Digital Hierarchy	
PSN	Packet Switched Networks	
SDH	Synchronous Digital Hierarchy	
TDM	Time Division Multiplex	

Contents

1 Introduction

1.1 Background

Today's trend is toward all-IP networks and services, including the bearer network. However, carriers are confronted with many difficulties in migrating from SDH-based legacy networks to all-IP Ethernet bearer networks. Supporting traditional TDM services on the new network is a key problem. TDM technology is mainly used for voice services and clock synchronization. For voice service, VoIP technology is widely used on the IP bearer network, but clock synchronization technologies are not as mature.

1.2 Definition of Synchronization

Synchronization on a network involves frequency and time synchronization (also called phase synchronization).

- Frequency synchronization

  Frequency synchronization is also referred to as clock synchronization. It keeps consistency between the frequencies or phases of clock signals to ensure that all devices on the communication network work at the same rate.

  Information is coded into discrete pulses through pulse code modulation (PCM) and transmitted on the data communication network. If the clock frequencies of two digital switching devices are different, or the digital bit streams transmitted between the two devices are added with phase drift and jitter due to interference and damage during the transmission, the bits in the buffer of the digital switching system may be lost or duplicated, resulting in slip in the bit stream.

- Time synchronization

  Time synchronization means to adjust the internal clock of the local device according to the received time information. Time synchronization and frequency synchronization follow similar principles. That is, both the frequency and phase of the clock are adjusted, and the phase of the clock is represented by a number, that is, a time point. Differing from frequency synchronization, time synchronization adjusts the clock discontinuously by receiving discontinuous time information. The adjustment of the Phase-Locked Loop (LLP) of the device clock is periodical.

Time synchronization involves two major functions: time service and time keeping. Time service is to adjust the clock according to the standard time. By adjusting the clock at irregular intervals, a device synchronizes its phase with the standard reference time. Time keeping refers to frequency synchronization. It ensures that the difference between the time on the local device and the standard reference time is within a reasonable range during the interval of clock adjustment.

- Difference between phase synchronization and frequency synchronization

Figure 1-1 Diagram of frequency synchronization

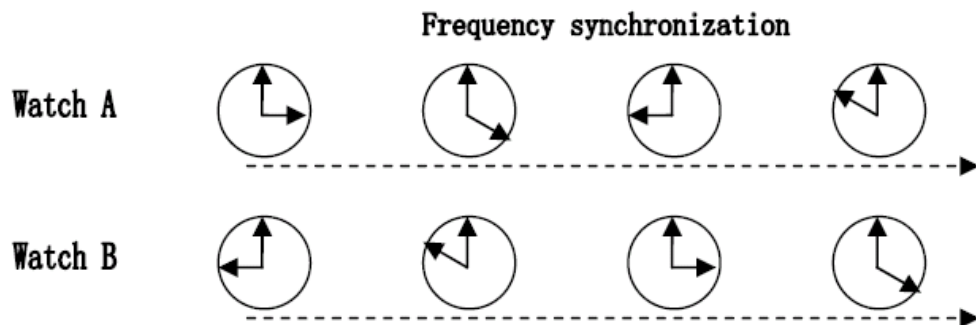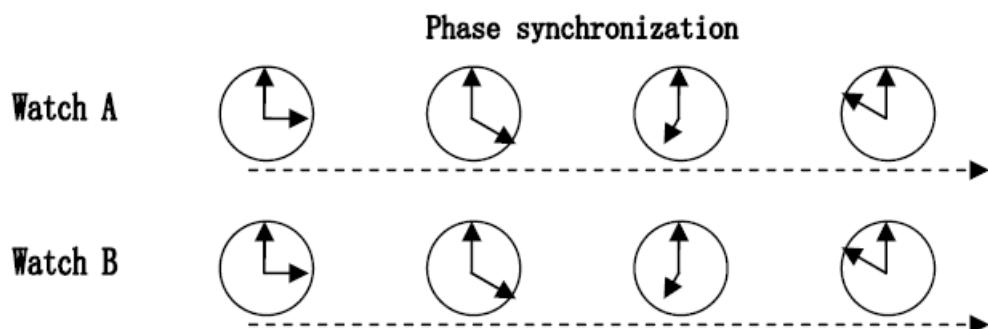**Figure 1-2** Diagram of phase synchronization

The preceding figures illustrate the difference between frequency synchronization and phase synchronization. In Figure 1-1, watch A and watch B show different time but they keep a constant time difference, for example, 6 hours. This is called frequency synchronization. In Figure 1-2, watch A and watch B always keep the same time. This is called phase synchronization.

1.3 Requirements for Clock Synchronization on the Communication Network

1.3.1 Requirements of Different Services

- TDM service on the traditional fixed network

  The TDM service on the traditional fixed network mainly refers to the voice service. If the clocks on two ends of the bearer network are inconsistent, the accumulate difference between the clocks will cause clock slip. ITUT G.823 defines the requirements and test standards for the TDM service on the fixed network, which are called the traffic interface standards.

- IP RAN

  Wireless communications have the most rigid requirements for clock frequency synchronization. The frequencies of BTSs must be synchronized to limit the frequency difference within a tolerable range; otherwise, services may be interrupted when wireless terminals switch between BTSs. Different from applications on the fixed TDM network, the clock here refers to the wireless radio clock, which requires more accurate frequency synchronization.

  Currently, a number of wireless access standards are available with different requirements for the accuracy of clock synchronization, as shown in the following table.

1. Requirements of wireless access standards for clock synchronization

Wireless Access Standard	Frequency Synchronization Accuracy	Phase Synchronization Accuracy
GSM	0.05 ppm	NA
WCDMA	0.05 ppm for BTS	NA
TD-SCDMA	0.05 ppm	3 us
CDMA2000	0.05 ppm	3 us
WiMax FDD	0.05 ppm	NA
WiMax TDD	0.05 ppm	1 us
LTE	0.05 ppm	Time synchronization preferred

Generally speaking, the European standard represented by GSM and WCDMA adopts the asynchronous technology for stations, which requires only frequency synchronization with an accuracy of 0.05 ppm or 50 ppb. Whereas the synchronous technology represented by CDMA and CDMA2000 requires phase synchronization.

Currently, the Global Positioning System (GPS) is the only solution to phase synchronization. As GPS also offers a solution to frequency synchronization, CDMA networks do not need additional synchronization functions.

To save costs and avoid military risks, GSM and WCDMA networks do not use the GPS. Therefore, the clock needs to be provided by the bearer network. On the legacy network, frequency synchronization is implemented through PDH or SDH. On the all-IP network, the clock needs to be provided by the IP network.

The IP radio access network (RAN) is a new application and the ITUT is now drafting applicable standards for IP RAN in G.8261. Currently, it has been a consensus that frequency synchronization should comply with the traffic interface standards in ITUT G.823 and have an accuracy of 50 ppb.

1.3.2 Requirements of the Dedicated Clock Synchronization Network (BITS)

The traditional communication network has an independent clock delivery network in addition to the service bearer network. The clock delivery network delivers clock signals through PDH or SDH. As defined by the ITUT, clock delivery should comply with the timing interface standards in G.823.

2 Synchronous Ethernet

2.1 Introduction to Synchronous Ethernet

Synchronous Ethernet is a technology used to recover the clock from the code stream on the Ethernet link.

As an asynchronous system, the Ethernet can function properly without a high-precision clock. Consequently, most Ethernet devices do not provide high-precision clocks. Nevertheless, this does not mean that the Ethernet is unable to provide a high-precision clock. Actually, the Ethernet uses serial code streams to transmit clock signals at the physical layer just as the SDH does; and the receiver must be able to recover the clock. Otherwise, communication fails. In other words, the Ethernet itself is capable of transmitting clock signals, but this capability is not actually used.

From a technical perspective, the clock obtained from the physical layer of the Ethernet is even more precise than that obtained using the SDH technology. For a clock to be obtained from the code stream of a line, there must be sufficient clock transitions in the code stream. In other words, the code stream cannot contain too many consecutive 1s or 0s. By performing a random scrambling, SDH significantly lowers the possibility of consecutive 1s and 0s. Nevertheless, consecutive 1s and 0s do occur. The physical layer codes of the Ethernet are in the form of 4B/5B (FE) and 8B/10B (GE). This means that for every 4 bits, one additional bit is added to eliminate the possibility of four consecutive 1s or 0s. This facilitates recovery of the clock.

Figure 2-1 Synchronization mechanism of synchronous Ethernet
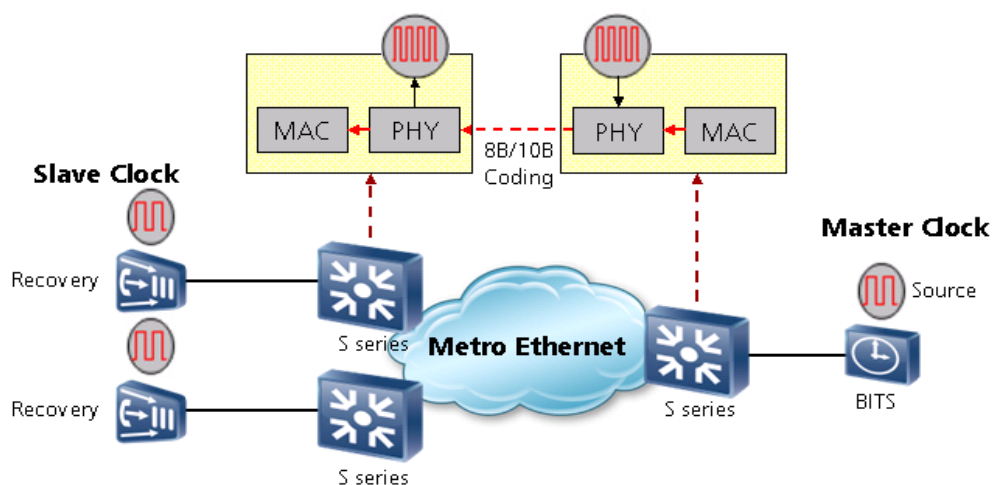
The sender switch transmits the high-precision clock to the Ethernet PHY chip. The Ethernet PHY chip then sends data with the high-precision clock to the receiver switch. The PHY chip of the receiver switch extracts the clock from the data code stream. Clock precision does not decrease in this process. This is the basic mechanism of synchronous Ethernet.

2.2 Technological Implementation

The solution of transmitting the clock through data code streams at the physical layer of the Ethernet provided by ITU-T Q13/SG15 is called Sync-E.

The code stream transmitted at the physical layer of the Ethernet contains sufficient clock transitions. Clock transitions occur even when no data is being transmitted. This nature of the Ethernet enables the Ethernet to recover the clock from the link. The sender sends high-precision clock signals, and the receiver recovers the clock signals from the link. The clock signals pass through a jitter attenuation circuit to reduce jitters generated on the link and then can be used as the clock in the system.

The code formats are:

- 10Base-T Manchester code: does not support clock recovery.
- 100Base-TX MLT-3 code: supports clock recovery.
- 100BASE-FX 4B5B code: supports clock recovery.
- 1000Base-X 8B10B code: supports clock recovery.
- 1000Base-T PAM5 code: supports clock recovery.
- 10GBase-R 64B66B code: supports clock recovery.

The Ethernet and TDM network both support clock recovery. On the Ethernet, however, low-precision clock signals are transmitted to save costs. The clock recovery function of the Ethernet is ignored, but it does not mean that the Ethernet cannot transmit high-precision clock.

Implementation of synchronous Ethernet is quite simple. A device needs a clock module (or clock board) to provide a high-precision system clock for all Ethernet interface cards. The

PHY component of each Ethernet interface sends out data through this clock. PHY component of the receiver Ethernet interface recovers the clock, divides the frequency, and sends it to the clock board. The clock board checks the quality of the clocks reported by interfaces and selects the most precise one to synchronize the system clock with it.

To select the clock source correctly, devices on the network also transmit the Synchronous State Message (SSM) with the clock information. On the SDH network, the quality level of the clock is transmitted through the outband overhead byte in the SDH frame. There is no outband tunnel on the Ethernet. Therefore, the downstream device sends SSM messages to notify the downstream device of the quality level of a clock.

2.2.1 Introduction to the SSM Message

- SSM, used to transmit the quality level of a clock

  The SSM was initially defined by the ITU-T to identify the quality level (QL) of the synchronization source on the SDH network. As specified by the ITU-T, the four spare bits in one of the five Sa bytes in a 2 Mbps bit stream are used to transmit the SSM value. The SSM and certain clock source selection protocols can improve synchronization performance and prevent timing loops on the network. In this way, the clock can be synchronized in different network structures and reliability of the synchronization network is also improved.
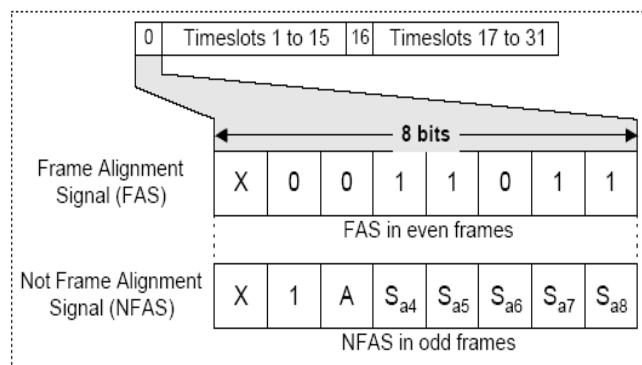
1. Codes of SSM quality levels

Quality Level	SSM Code	Code Value
QL-PRC (G.811)	0010	0x02
QL-SSU-A (G.812 I, V)	0100	0x04
QL-SSU-B (G.812 VI)	1000	0x08
QL-SEC (G.813 I)	1011	0x0b
QL-DNU	1111	0x0f

- Transmission of the SSM message in Sync-E

Figure 2-2 Transmission of the SSM message in Sync-E
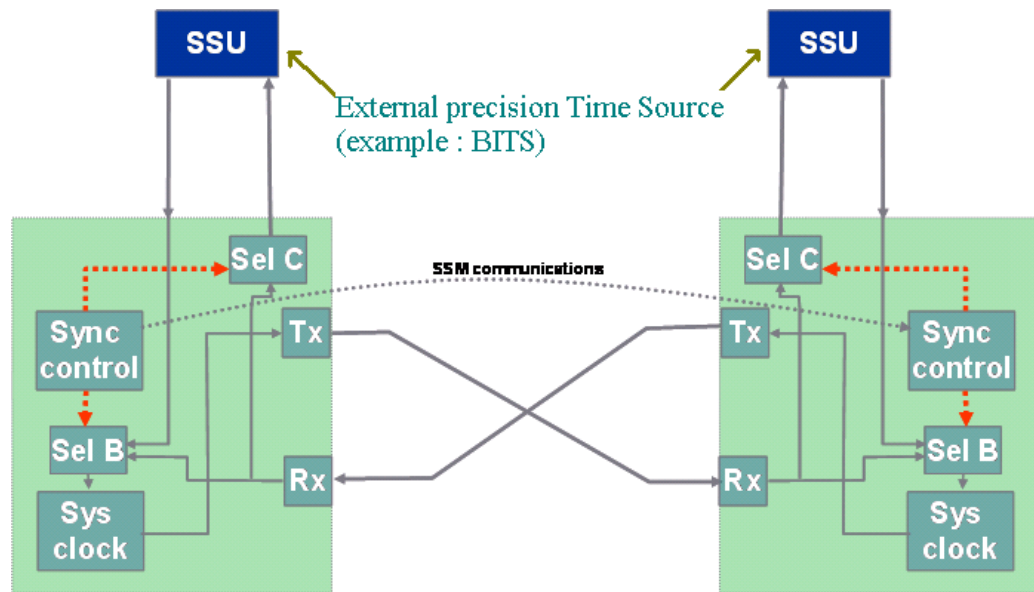
The ESMC messages are transmitted through the IEEE 802.3 Organization Specific Slow Protocol (OSSP). In addition, ESMC messages and ESSM messages are transmitted separately so that Sync-E interfaces and non-Sync-E interfaces are compatible with each other. ESMC messages are classified into two types: event-driven message and periodic information message. The event-driven message is an important message for the Sync-E clock and it must contain the QL information. The periodic information message can carry more status information such as the source ID, source information, and number of hops for sync trail from source. The periodic information message (heartbeat message) has a Revision field, which indicates the difference with the last message. The frequency for sending these messages is configurable. Generally, the acceptable frequency is 1 frame per second (fps). When to send the messages and how to process the message have not been defined yet, but the frame format of the message has been defined as follows:

- The frame is a slow protocol frame and uses the format of a standard Ethernet frame.
- The frame uses the unique destination MAC address 0x0180C2000002.
- The value of the Length/Type field is the same as that in the slow protocol frame (value 0x8809).
- The frame has a subtype field 0x0A to identify the organizational protocol.

2.3 Advantages and Disadvantages of Sync-E

Advantages:

- The quality of the recovered clock is 0.01 ppm, which is close to the quality of the SDH clock.
- Sync-E is hardly affected by the PSN network and therefore has high reliability.
- The architecture of the clock system is similar to the architecture of the SDH solution and is mature.

Disadvantages:

- Sync-E needs to be deployed on all the nodes on the network, which requires high expenses in network reconstruction.
- Sync-E does not support time synchronization (phase synchronization).

<div align="right">

3 IEEE 1588v2

</div>

3.1 Development of IEEE 1588v2

The Ethernet became the IEEE 802.3 standard in 1985. In 1995, the data transmission rate is increased from 10 Mbit/s to 100 Mbit/s. During the increase of transmission rate, the computer and network industry were trying to solve the problem of timing synchronization on the Ethernet. The Network Time Protocol (NTP) is developed to ensure software-based timing synchronization between network devices. However, the NTP protocol cannot provide the precision required by measurement tools and industrial control.

To solve the preceding problem, the precise network clock synchronization committee was constituted in the end of 2000. In 2001, with the support of the National Institute of Standards and Technology (NIST), the committee drafted the related standard, which has been used as the IEEE 1588 standard since the end of 2002.

In the communications industry, the clock signal transmission technology of the PSN develops fast. The revised IEEE 1588 standard was issued in June 2006 and the IEEE 1558v2 was revised in 2007.

3.2 Introduction to IEEE 1588v2

IEEE 1588v2 is the IEEE Standard for a Precision Clock Synchronization Protocol for Networked Measurement and Control Systems. It is used for synchronization of the time and clock frequency. It defines the Precision Time Protocol (PTP) to provide sub-microsecond precision synchronization of time and clock frequency. The PTP protocol will be described in later sections.

The PTP defines the following types of clock devices:

- Ordinary clock (OC): has only one physical port to the network. The OC can be used as the grandmaster clock or slave clock. It can send and receive PTP messages and supports the synchronization hierarchy determining mechanism. In addition, it supports the Delay mechanism and Pdelay mechanism.

- Boundary clock (BC): has multiple physical ports to the network. Each physical port is similar to the physical port of the OC and can be connected to multiple sub-domains. The BC can be used as the intermediate transit device.

- End-to-end transparent clock (E2E TC): has multiple ports. The E2E TC can forward all PTP messages and can correct the residential time of PTP event messages. The ports of

the E2E TC do not have a protocol engine and do not determine the master and slave status of the clock.

- Peer-to-peer transparent clock (P2P TC): has multiple ports and each port has a Pdelay processing module to support the Pdelay mechanism. The ports of the P2P TC do not have a protocol engine and do not determine the master and slave status of the clock.

- PTP management device: has multiple ports, including the management port used to send PTP Management messages.

    PTP also defines the following messages for synchronization and control between devices:

- Event message: It is the timing message. That is, a device must generate an accurate timestamp when sending or receiving an event message. Event messages are classified into the following types:
    - Sync
    - Delay_Req
    - Pdelay_Req
    - Pdelay_Resp

- General message: A device does not generate an accurate timestamp when sending or receiving a general message. General messages are classified into the following types:
    - Announce
    - Follow_Up
    - Delay_Resp
    - Pdelay_Resp_Follow_Up
    - Management
    - Signaling

- Functions of PTP messages are as follows:
    - The Sync, Delay_Req, Follow_Up, and Delay_Resp messages are used to generate and transmit time information to synchronize the ordinary clock or boundary clock.
    - The Pdelay_Req, Pdelay_Resp, and Pdelay_Resp_Follow_Up messages are used to measure the delay of the path between two clock ports.
    - The Announce messages are used to establish the synchronization hierarchy.
    - The Management messages are used to query and update the PTP data set maintained by the clock device.
    - The Signaling messages are used for other purposes, for example, coordinate the frequencies of sending unicast packets on the master and slave ports.

IEEE 1588v2 involves three key technologies: best master clock (BMC) algorithm, master-slave synchronization mechanism, and TC model.

IEEE 1588v2 uses accurate timestamps to synchronize the time and frequency through handshake.

Figure 3-1 Working principle of IEEE 1588v2
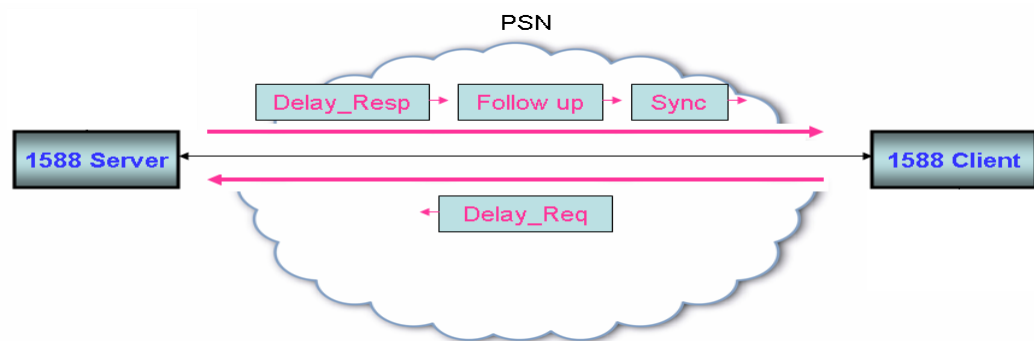
3.3 Implementation of IEEE 1588v2

3.3.1 Master-Slave Determining Mechanism – BMC Algorithm

Network devices may reference different time sources. Before clock synchronization, the master and slave status of the devices must be determined. The objectives are:

- At the network layer:
    - Determine the grandmaster clock device connected to the best clock source.
    - Determine the path between each clock device and the grandmaster clock device to prevent timing loops.
- At the device layer:
    - Determine the status of each port, which can be master, slave, or passive.

The calculation process of the BMC algorithm is as follows:

- Each port calculates the best message Erbest. A port compares the priorities of the received Announce messages and selects the one with the highest priority as the Erbest. After the calculation, each port of a clock device obtains an Erbest.
- Each device calculates the best message Ebest. A device compares the priorities of the Erbest messages reported by its ports and selects the one with the highest priority as the Ebest.
- Each port calculates the recommended status. A port compares the Ebest and defaultDS with its own Erbest to determine the recommended status of itself.

3.3.2 Process of Determining the Master-Slave Hierarchy

**Figure 3-2** Process of determining the master-slave hierarchy
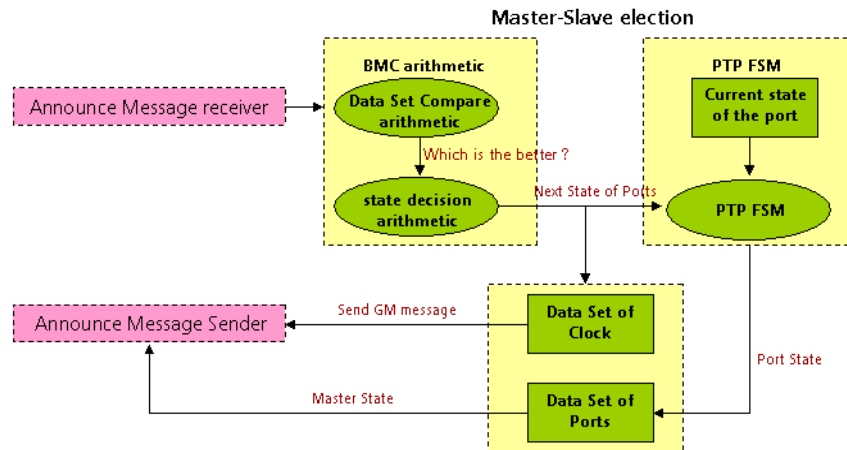
- Sending and receiving Announce messages: The OC or BC devices exchange their own grandmaster clock parameters. The information in the Announce messages is from the data set of each device.

- Running the BMC algorithm: When receiving Announce messages, a device runs the BMC algorithm to compare the grandmaster clock parameters in the Announce messages. Then the device selects the port that receives the Announce message with the best grandmaster parameters as the grandmaster clock or announces itself as the grandmaster clock. The device also calculates the recommended status of each port.

- Updating the data set: According to the recommended status calculated through the BMC algorithm, each device updates the data set.

- Determining the port status: The state machine of each port determines the status of the interface according to the recommended status calculated through the BMC algorithm and the actual status of the port. Then the master-slave hierarchy is established in the entire domain.

3.3.3 Master-Slave Synchronization Mechanism

The master-slave synchronization mechanism involves time synchronization and frequency synchronization.

Figure 3-3 Master-slave time synchronization mechanism
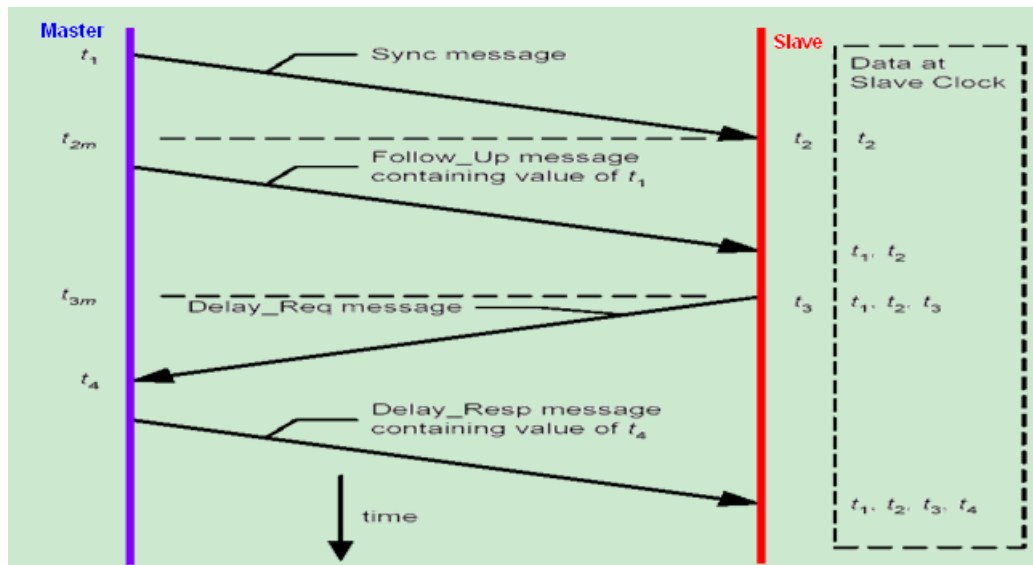
As shown in the preceding figure, the time synchronization process is as follows:

1. The master clock sends a Sync message at t1.
2. The slave clock receives the Sync message at t2.
3. The slave clock sends a Req message at t3.
4. The master clock receives the Req message at t4.
5. The master clock sends a Resp message containing t4 to the slave clock.

Then, t2 - t1 = Delay + Offset

t4 - t3 = Delay - Offset
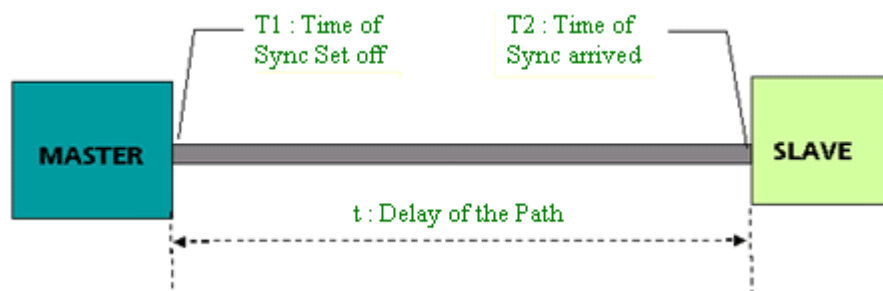
That is:

Offset = [(t2 - t1) - (t4 - t3)]/2

Delay = [(t2 - t1) + (t4 - t3)]/2

The frequency synchronization process is as follows.

**Figure 3-4** Master-slave frequency synchronization mechanism

1. The slave device calculates the corrected time when each Sync message arrives at the slave port based on the time of the master device, that is, the set off time of the Sync message plus the delay of the path.

2. The slave device calculates the interval between each two Sync messages based on the corrected time, and the interval between each two Sync messages based on the arriving timestamp on the slave port. Then, the slave device calculates the proportion factor between the two types of interval.

3. The slave clock adjusts its clock frequency according to the proportion factor.

   Proportion factor = [(T1N+tN) - (T10+t0)]/[T2N - T20]

When a PTP message traverses the protocol stack on a node, the timestamp of the message is selected among the reference points defined by the protocol stack (A, B, C in the following figure). The closer a reference point is to the actual physical connection point, the smaller the timing error is. Point A in the following figure is the best reference point.

Figure 3-5 Reference points of the timestamp
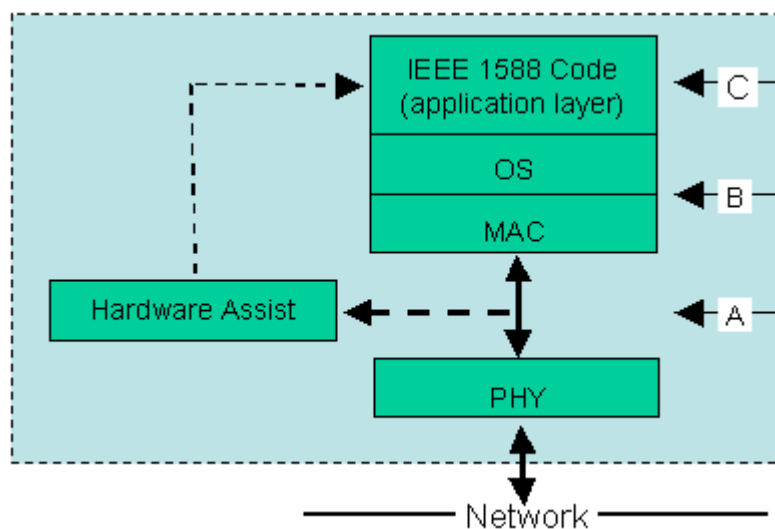
 NOTE

> IEEE 1588 cannot measure the asymmetric path delays between the master and slave devices. The asymmetric path delays can only be measured by other means and provided for IEEE 1588.

What is asymmetric correction?

1. When the communication path is symmetric, the path delay is the same in both directions of the path. In this case, the average path delay can be used to calculate the unidirectional path delay.

2. When the communication path is asymmetric, the path delays in the two directions of the path are different. In this case, asymmetric correction needs to be performed on the average path delay to obtain the unidirectional path delay.

The delayAsymmetry parameter needs to be used to calculate the actual path delay from the master to the slave or from the responder to the requester.

tms = <meanPathDelay> + delayAsymmetry,

tsm = <meanPathDelay> - delayAsymmetry.
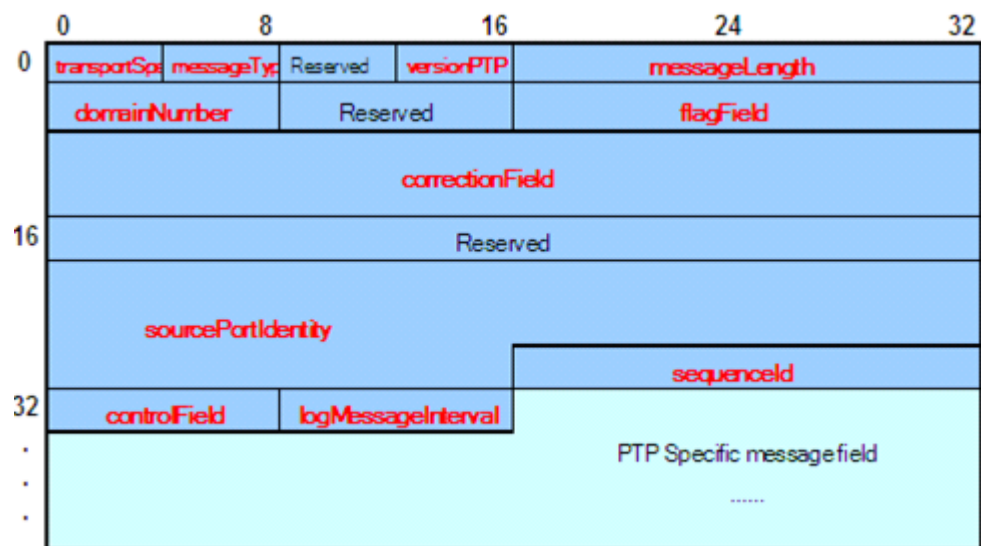
3.4 Format of PTP Messages

All PTP messages consist of a header, a body, and a suffix.

- Header: It is the general packet header consisting of 34 bytes. The header is the forepart of the 10 types of PTP messages.
- Body: The contents of the body identify the type of a PTP message.
- Suffix: This field can be filled with 0s or TLVs to expand the functions of the PTP protocol. For the TLVs that can be added to the suffix, see the definition of the TLV entities.

Figure 3-6 Format of PTP messages

Figure 3-7 PTP message format – general header

The general header of a PTP message contains the following fields:

- transportSpecific

- 0: indicates that the PTP message is used by IEEE 1588.
- 1: indicates that the PTP message is used by 802.1as.
- messageType: indicates the type of the PTP message.
  - 0: Sync message
  - 1: Delay_Req message
  - 2: Pdelay_Req message
  - 3: Pdelay_Resp message
  - 4-7: reserved
  - 8: Follow_Up message
  - 9: Delay_Resp message
  - A: Pdelay_Resp_Follow_Up message
  - B: Announce message
  - C: Signaling message
  - D: Management message
  - E-F: reserved
- versionPTP: indicates the version of IEEE 1588.
- messageLength: indicates the length of the PTP message.
- domainNumber: indicates the domain that the sender of the PTP message belongs to.
- correctionField: is the correction value expressed in ns multiplied by 216. For example, if the correction value is 2.5 ns, the value of this field is 0x28000.
- sourcePortIdentity: indicates the clock ID and the port that sends the PTP message.
- sequenceId: indicates the sequence number of the PTP message and the relationship between consecutive PTP messages.
- controlField: depends on the message type and is used for compatibility with IEEE 1588v1.
- logMessageInterval: indicates the interval for sending PTP messages, depending on the message type.

3.5 Problems Requiring Consideration in PTP Deployment

3.5.1 Influence of the Physical Topology on the Synchronization Precision

Theoretically, the PTP protocol can run in any networking environment. In a complicated system, however, there are usually some edge bridges or edge routers that do not support PTP. These devices cause relatively large delay and delay jitter. As a result, clock synchronization cannot achieve the expected precision.

To improve the timing performance, you can use a BC device to replace the edge router or ordinary bridge. In a linear topology, you can use a TC device to replace the ordinary bridge.

3.5.2 Synchronization Precision

- Delay jitter of the protocol stack

  The simplest implementation of PTP is to use it as a common application on the top of the network protocol stack. Timestamps are generated on the application layer. Therefore, the delay jitter in the protocol stack causes errors in timestamps. Typical errors of timestamps range from 100 us to several milliseconds.

  If timestamps are generated at the interrupt level rather than the application level, the delay jitter can be reduced to about 10 us. The actual jitter depends on the application at the interrupt level and the service traffic on the network.

  The hardware assistant technology can minimize the delay jitter in the protocol stack by generating timestamps at the physical layer of the protocol stack. The delay jitter can be controlled in an order of ns. The error of the timestamp is caused by the phase locking feature of the PHY chip used to recover the clock from data streams and synchronize data.

- Delay jitter during forwarding

  Network components introduce jitter when transmitting messages, which reduces the accuracy of the offsetFromMaster and meanPathDelay values.

  High-priority services have smaller delay jitter when being forwarded. On the bridge and router that can differentiate service priorities, you can set the priority of the PTP event messages to be higher than the priority of other data packets, reducing the jitter of the PTP event messages.

- Accuracy of timestamps

  The clock that is used to generate timestamps required by PTP must retain the ideal accuracy. The local clock frequency of the TC or slave clock is usually different from the clock frequency of the grandmaster clock, which affects the accuracy of timestamps.

3.5.3 Implementation of the PTP System

To ensure the PTP system with the best performance, ensure that:

- All PTP nodes in the entire domain use the same transport protocol.
- All PTP nodes in the entire domain use the same BMC algorithm.
- All nodes in the domain use the same status configuration items.
- All nodes in the domain use the same delay measurement mechanism.
- The attributes and configurable data set members of all nodes in the system use the same default values.
- The attributes of all nodes in the system have the same value range, that is, the same maximum and minimum values.
- All nodes in the domain use the same PTP template.
- All nodes in the domain use the same PTP version.

3.5.4 Performance

The best clock synchronization can be obtained if the following conditions are met:

- The path delay between the master and slave clocks is the symmetric in the two directions.

- If the delays in two directions are asymmetric in the timestamp mechanism or protocol channel and the asymmetry cannot be ignored, asymmetric correction should be performed.

- The path delay between the master and slave node is constant within an interval for sending Delay_Req messages.

- The timestamp used in PTP should be generated as close as possible to the physical layer. The timestamp is transmitted through the Follow_Up message of the master clock or the Pdelay_Resp_Follow_Up message of the PTP TC.

- The clock calculation capacity of the protocol is high enough and the number of clocks is sufficient to meet the timing requirement. Otherwise, the synchronization performance is degraded when the system cannot process PTP messages because there are not sufficient resources. For example, the BC and OC devices need sufficient resources to process Delay_Req messages.

- The internal oscillator has the qualified stability and precision.

4 References

- Technical Communication on Synchronous Ethernet and IEEE 1588v2
- Clock Transparent Transmission White Paper
- Introduction to IEEE 1588v2