TICTOC Working Group Doug Arnold
Internet Draft Meinberg-USA
Intended status: Standards Track Heiko Gerstung
Meinberg
Expires: May 22, 2014 Nov 23, 2013

Enterprise Profile for the Precision Time Protocol With Mixed Multicast and Unicast Messages

draft-ietf-tictoc-ptp-enterprise-profile-01.txt

Abstract

This document describes a profile for the use of the Precision Time Protocol in an IPV4 or IPv6 Enterprise information system environment. The profile uses the End to End Delay Measurement Mechanism, allows both multicast and unicast Delay Request and Delay Response Messages.

Table of Contents

1. Introduction 2
2. Conventions used in this document 3
3. Technical Terms 4
4. Problem Statement 5
5. Network Technology 6
6. Time Transfer and Delay Measurement 7
 6.1. Unicast Delay Measurement Mode 7
 6.2. Multicast Delay Measurement Mode 7
 6.3. Hybrid delay Measurement Mode 8
7. Default Message Rates 8
8. Requirements for Master Clocks 8
9. Requirements for Slave Clocks 10
10. Requirements for Transparent Clocks 10
11. Management and Signaling Messages 10
12. Forbidden PTP Options 11
13. Interoperation with Other PTP Profiles 11
14. Security Considerations 11
15. IANA Considerations 12
16. References 12
 16.1. Normative References 12
 16.2. Informative References 12
17. Acknowledgments 13
18. Authors addresses 13

1. Introduction

The Precision Time Protocol ("PTP"), standardized in IEEE 1588, has been designed in its first version (IEEE 1588-2002) with the goal to minimize configuration on the participating nodes. Network communication was based solely on multicast messages, which unlike NTP did not require that a receiving node ("slave clock") in [IEEE1588] needs to know the identity of the time sources in the network (the Master Clocks).

The so-called "Best Master Clock Algorithm" ([IEEE1588] Clause 9.3), a mechanism that all participating PTP nodes must follow, set up strict rules for all members of a PTP domain to determine which node shall be the active sending time source (Master Clock). Although the multicast communication model has advantages in smaller networks, it complicated the application of PTP in larger networks, for example in environments like IP based telecommunication networks or financial data centers. It is

considered inefficient that, even if the content of a message applies only to one receiver, it is forwarded by the underlying network (IP) to all nodes, requiring them to spend network bandwidth and other resources like CPU cycles to drop the message. The second revision of the standard (IEEE 1588-2008) is the current version (also known as PTPv2) and introduced the possibility to use unicast communication between the PTP nodes in order to overcome the limitation of using multicast messages for the bi-directional information exchange between PTP nodes. The unicast approach avoided that, in PTP domains with a lot of nodes, devices had to throw away up to 99% of the received multicast messages because they carried information for some other node. PTPv2 also introduced so-called "PTP profiles" ([IEEE1588] Clause 19.3). This construct allows organizations to specify selections of attribute values and optional features, simplifying the configuration of PTP nodes for a specific application. Instead of having to go through all possible parameters and configuration options and individually set them up, selecting a profile on a PTP node will set all the parameters that are specified in the profile to a defined value. If a PTP profile definition allows multiple values for a parameter, selection of the profile will set the profile-specific default value for this parameter. Parameters not allowing multiple values are set to the value defined in the PTP profile. A number of PTP features and functions are optional and a profile should also define which optional features of PTP are required, permitted or prohibited. It is possible to extend the PTP standard with a PTP profile by using the TLV mechanism of PTP (see [IEEE1588] Clause 13.4), defining an optional Best Master Clock Algorithm and a few other ways. PTP has its own management protocol (defined in [IEEE1588] Clause 15.2) but allows a PTP profile specify an alternative management mechanism, for example SNMP.

2. Conventions used in this document

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC-2119 [RFC2119].

In this document, these words will appear with that interpretation only when in ALL CAPS. Lower case uses of these words are not to be interpreted as carrying RFC-2119 significance.

3. Technical Terms

Acceptable Master Table: A PTP Slave Clock may maintain a list of masters which it is willing to synchronize to.

Alternate Master: A PTP Master Clock, which is not the Best Master, may act as a master with the Alternate Master flag set on the messages it sends.

Announce message: Contains the master clock properties of a Master clock. Used to determine the Best Master.

Best Master: A clock with a port in the master state, operating consistently with the Best Master Clock Algorithm.

Best Master Clock Algorithm: A method for determining which state a port of a PTP clock should be in. The algorithm works by identifying which of several PTP Master capable clocks is the best master. Clocks have priority to become the acting Grandmaster, based on the properties each Master Clock sends in its Announce Message.

Boundary Clock: A device with more than one PTP port. Generally boundary clocks will have one port in slave state to receive timing and then other ports in master state to re-distribute the timing.

Clock Identity: In IEEE 1588-2008 this is an 64-bit number assigned to each PTP clock which must be unique. Often the

Ethernet MAC address is used since there is already an international infrastructure for assigning unique numbers to each device manufactured.

End to End Delay Measurement Mechanism: A network delay measurement mechanism in PTP facilitated by an exchange of messages between a Master Clock and Slave Clock.

Grandmaster: the primary master clock within a domain of a PTP system

IEEE 1588: The timing and synchronization standard which defines PTP, and describes The node, system, and communication properties necessary to support PTP.

Master clock: a clock with at least one port in the master state.

NTP: Network Time Protocol, defined by RFC 5905, see [NTP].

Ordinary Clock: A clock that has a single Precision Time Protocol (PTP) port in a domain and maintains the timescale used in the domain. It may serve as a master clock, or be a slave clock.

Peer to Peer Delay Measurement Mechanism: A network delay measurement mechanism in PTP facilitated by an exchange of messages between adjacent devices in a network.

Preferred Master: A device intended to act primarily as the Grandmaster of a PTP system, or as a back up to a Grandmaster.

PTP: The Precision Time Protocol, the timing and synchronization protocol define by IEEE 1588.

PTP port: An interface of a PTP clock with the network. Note that there may be multiple PTP ports running on one physical interface, for example a unicast slave which talks to several Grandmaster clocks in parallel.

PTPv2: Refers specifically to the second version of PTP defined by IEEE 1588-2008.

Rogue Master: A clock with a port in the master state, even though it should not be in the master state according to the Best Master Clock Algorithm, and does not set the alternate master flag.

Slave clock: a clock with at least one port in the slave state, and no ports in the master state.

Slave Only Clock: An Ordinary clock which cannot become a Master clock.

TLV: Type Length Value, a mechanism for extending messages in networked communications.

Transparent Clock. A device that measures the time taken for a PTP event message to transit the device and then updates the message with a correction for this transit time.

Unicast Discovery: A mechanism for PTP slaves to establish a unicast communication with PTP masters using a configures table of master IP addresses and Unicast Message Negotiation.

Unicast Negotiation: A mechanism in PTP for Slave Clocks to negotiate unicast Sync, announce and Delay Request Message Rates from a Master Clock.

4. Problem Statement

This document describes a version of PTP intended to work in large enterprise networks. Such networks are deployed, for example, in financial corporations. It is becoming increasingly common in such networks to perform distributed time tagged measurements, such as one-way packet latencies and cumulative delays on software systems spread across multiple computers. Furthermore there is

often a desire to check the age of information time tagged by a different machine. To perform these measurements it is necessary to deliver a common precise time to multiple devices on a network. Accuracy currently required in the Financial Industry range from 100 microseconds to 500 nanoseconds to the Grandmaster. This profile does not specify timing performance requirements, but such requirements explain why the needs cannot always be met by NTP, as commonly implemented. Such accuracy cannot usually be achieved with a traditional time transfer such as NTP, without adding non-standard customizations such as hardware time stamping, and on path support. These features are currently part of PTP, or are allowed by it. Because PTP has a complex range of features and options it is necessary to create a profile for enterprise networks to achieve interoperability between equipment manufactured by different vendors.

Although enterprise networks can be large, it is becoming increasingly common to deploy multicast protocols, even across multiple subnets. For this reason it is desired to make use of multicast whenever the information going to many destinations is the same. It is also advantageous to send information which is unique to one device as a unicast message. The latter can be essential as the number of PTP slaves becomes hundreds or thousands.

PTP devices operating in these networks need to be robust. This includes the ability to ignore PTP messages which can be identified as improper, and to have redundant sources of time.

5. Network Technology

This PTP profile SHALL operate only in networks characterized by UDP [RFC768] over either IPv4 [RFC791] or IPv6 [RFC2460], as described by Annexes D and E in [IEEE1588] respectively. If a network contains both IPv4 and IPv6, then they SHALL be treated as separate communication paths. Clocks which communicate using IPv4 can interact with clocks using IPv6 if there is an intermediary device which simultaneously communicates with both IP versions. A boundary clock might perform this function, for example. A PTP domain SHALL use either IPv4 or IPv6 over a communication path, but not both. The PTP system MAY include switches and routers. These devices MAY be transparent clocks, boundary clocks, or neither, in any combination. PTP Clocks MAY be Preferred Masters, Ordinary Clocks, or Boundary Clocks. The ordinary clocks may be Slave Only Clocks, or be master capable.

Note that clocks SHOULD always be identified by their clock ID and not the IP or Layer 2 address. This is important in IPv6 networks since Transparent clocks are required to change the source address of any packet which they alter. In IPv4 networks some clocks might be hidden behind a NAT, which hides their IP addresses from the rest of the network. Note also that the use of NATs may place limitations on the topology of PTP networks, depending on the port forwarding scheme employed. Details of implementing PTP with NATs are out of scope of this document.

Similar to NTP, PTP makes the assumption that the one way network delay for Sync Messages and Delay Response Messages are the same. When this is not true it can cause errors in the transfer of time from the Master to the Slave. It is up to the system integrator to design the network so that such effects do not prevent the PTP system from meeting the timing requirements. The details of network asymmetry are outside the scope of this document. See for example, [G8271].

6. Time Transfer and Delay Measurement

Master clocks, Transparent clocks and Boundary clocks MAY be either one-step clocks or two-step clocks. Slave clocks MUST support both behaviors. The End to End Delay Measurement Method MUST be used.

Note that, in IP networks, Sync messages and Delay Request messages exchanged between a master and slave do not necessarily traverse the same physical path. Thus, wherever possible, the network SHOULD be traffic engineered so that the forward and reverse routes traverse the same physical path. Traffic engineering techniques for path consistency are out of scope of this document.

In all three of the delay measurement modes, Sync messages MUST be sent as PTP event multicast messages (UDP port 319) to the PTP primary IP address. Two step clocks SHALL send Follow-up messages as PTP general messages (UDP port 320). Announce messages MUST be sent as multicast messages (UDP port 320) to the PTP primary address. The PTP primary IP address is 224.0.1.129 for IPv4 and FF0X:0:0:0:0:0:0:181 for Ipv6, where X can be a value between 0x0 and 0xF, see [IEEE1588] Annex E, Section E.3.

6.1. Unicast Delay Measurement Mode

All Delay Request PTP event messages and Delay Response PTP general messages MUST be sent as unicast messages. Unicast Discovery and Unicast Message Negotiation options MUST NOT be utilized.

6.2. Multicast Delay Measurement Mode

All Delay Request PTP event messages and Delay Response general messages MUST be sent as multicast messages.

6.3. Hybrid Delay Measurement Mode

Delay Request Messages MAY be sent as either multicast or unicast PTP event messages. Master clocks SHALL respond to multicast Delay Request messages with multicast Delay Response PTP general messages. Master clocks SHALL respond to unicast Delay Request PTP event messages with unicast Delay Response PTP general messages.

The default mode for Enterprise Profile PTP Master Clocks is Hybrid Delay Measurement Mode. This allow for the use of Ordinary clocks which do not support the Enterprise Profile, as long as they are slave Only Clocks.

7. Default Message Rates

The Sync, Announce and Delay Request default message rates SHALL each be once per second. The Sync and Delay Request message rates MAY be set to other values, but not less than once every 128 seconds, and not more than 128 messages per second. The Announce message rate SHALL NOT be changed from the default value. The Announce Receipt Timeout Interval SHALL be three Announce Intervals for Preferred Masters, and four Announce Intervals for all other masters.

8. Requirements for Master Clocks

Master clocks SHALL obey the standard Best Master Clock Algorithm from [IEEE1588]. Clocks which are master capable MAY act as Alternate masters according to [IEEE1588]. Preferred Master Clocks SHOULD operate as Alternate Masters when they are not the Best Master. Using multiple masters can mitigate rogue master and man-in-the-middle attacks such as delay attacks, packet interception / manipulation attacks. Assuming the path to each master is different, an attacker would have to attack more than one path simultaneously.

A port of a clock SHALL NOT be in the master state unless the clock has a current value for the number of UTC leap seconds. A clock with a port in the master state SHOULD indicate the maximum adjustment to its internal clock within one sync interval. The maximum phase adjustment is indicated in the Enterprise Profile announce TLV field for Maximum Phase Adjustment.

The Announce Messages SHALL include a TLV which indicates that the clock is operating in the Enterprise Profile. The TLV shall have the following structure:

TLV Type (Enumeration16): ORGANIZATION_EXTENSION value = 0003 hex

Length Field (UInteger16): value = 12. The number of TLV octets

Port Number (UInteger16): The Port Number of the port transmitting the TLV. The all-ones Port Number, with value FFFFhex, is used to indicate that the identified profile is applicable to all ports on the clock.

Organization Unique Identifier (3 Octets): The Organization ID value for IETF assigned by IEEE = 00005Ehex

IETF Profile number (UInteger16): value = 1

Revision number (UInteger8): value = 1

Delay Measurement Mode (Enumeration8):
- Value 0 = Multicast Delay Measurement Mode
- Value 1 = Unicast Delay Measurement Mode
- Value 2 = Hybrid Delay Measurement Mode
- Values 3-255 = reserved for future use.

Maximum Absolute Phase Adjustment Value within one sync interval (UInteger16): value

Maximum Phase Adjustment Units (Enumeration8):
- Value 0 = unknown
- Value 1 = seconds
- Value 3 = milliseconds
- Value 6 = microseconds
- Value 9 = nanoseconds
- Value 12 = picoseconds
- Value 15 = femtoseconds
- All other values reserved for future use

The TLV represents the Delay Request Mode of the Master, not necessarily the Grandmaster. So for example, one port of a Boundary clock could be set to Unicast Delay Measurement Mode, while the rest of the network is Hybrid Delay Measurement Mode.

9. Requirements for Slave Clocks

Slave clocks MUST be able to operate properly in a network which contains Alternate Masters. Slaves SHOULD make use of information from the Alternate Masters in their clock control subsystems. Slave Clocks MUST be able to operate properly in the presence of a Rogue Master. Slaves SHOULD NOT Synchronize to a Rogue Master. Slaves MAY use an Acceptable Master Table. If the Best Master is not an Acceptable Master, but an Alternate Master is an Acceptable Master, then the Slave SHOULD synchronize to the acceptable Alternate Master. Note that IEEE 1588-2008 requires slave clocks to support both two-step or one-step Master clocks. See [IEEE1588], section 11.2.

Since Announce messages are sent as multicast messages in all mode, slaves can obtain the IP addresses of master from the Announce messages. Note that the IP source addresses of Sync and Follow-up messages may have been replaced by the source addresses of a transparent clock, so slaves MUST send Delay Request messages to the IP address in the Announce message. Sync and Follow-up messages can be correlated with the Announce message using the clock ID, which is never altered by Transparent clocks.

10. Requirements for Transparent Clocks

Transparent clocks SHALL NOT change the transmission mode of an Enterprise profile PTP message. For example a Transparent clock SHALL NOT change a unicast message to a multicast message. Transparent clocks SHALL NOT alter the Enterprise Profile TLV of an Announce message, or any other part of an Announce message.

11. Management and Signaling Messages

PTP Management messages MAY be used. Any PTP management message which is sent with the targetPortIdentity.clockIdentity set to all 1s (all clocks) MUST be sent as a multicast message. Management messages with any other value of for the Clock Identity is intended for a specific clock and MUST be sent as a unicast message. Similarly, if any signaling messages are used they MUST also be sent as unicast messages whenever the message is intended for a specific clock.

12. Forbidden PTP Options

Clocks operating in the Enterprise Profile SHALL NOT use peer to peer timing for delay measurement. Clocks operating in the Enterprise Profile SHALL NOT use Unicast Message Negotiation to determine message rates. Slave clocks operating in the Enterprise Profile SHALL NOT use Unicast Discovery to establish connection to Master clocks. Grandmaster Clusters are NOT ALLOWED. Clocks operating in the Enterprise Profile SHALL NOT use Alternative Time Scales.

13. Interoperation with Other PTP Profiles

Clocks operating in the Enterprise profile will not interoperate with clocks operating in the Power Profile [C37.238], because the

Enterprise Profile requires the End to End Delay Measurement Mechanism and the Power Profile requires the Peer to Peer Delay Measurement Mechanism.

Clocks operating in the Enterprise profile will not interoperate with clocks operating in the Telecom Profile for Frequency Synchronization[G8265.1], because the Enterprise Profile forbids Unicast Message Negotiation, and Unicast Discovery. These features are part of the default manner of operation for the Telecom Profile for Frequency Synchronization.

Clocks operating in the Enterprise profile will interoperate with clocks operating in the default profile if the default profile clocks operate on IPv4 or IPv6 networks, use the End to End Delay Measurement Mechanism, and use management messages in unicast when those messages are directed at a specific clock. If any of these requirements are not met than Enterprise Profile clocks will not interoperate with Default Profile Clocks. The Default Profile is described in Annex J of [IEEE1588].

Enterprise Profile Clocks will interoperate with clocks operating in other profiles if the clocks in the other profiles obey the rules of the Enterprise Profile. These rules MUST NOT be changed to achieve interoperability with other profiles.

14. Security Considerations

Protocols used to transfer time, such as PTP and NTP can be important to security mechanisms which use time windows for keys and authorization. Passing time through the networks poses a security risk since time can potentially be manipulated.

The use of multiple simultaneous masters, using the IEEE 1588 Alternate Master option can mitigate problems from rogue masters and man-in-the-middle attacks. See sections 9 and 10. Additional security mechanisms are outside the scope of this document.

15. IANA Considerations

There are no IANA requirements in this specification.

16. References

16.1. Normative References

[IEEE1588] IEEE std. 1588-2008, "IEEE Standard for a Precision Clock Synchronization for Networked Measurement and Control Systems." July, 2008.

[RFC768] Postel, J., "User Datagram Protocol," RFC 768, August, 980.

[RFC791] "Internet Protocol DARPA Internet Program Protocol Specification," RFC 791, September, 1981.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

[RFC2460] Deering, S., Hinden, R., "Internet Protocol, Version 6 (IPv6) Specification," RFC 2460, December, 1998.

16.2. Informative References

[C37.238] IEEE std. C37.238-2011, "IEEE Standard Profile for Use of IEEE 1588 Precision Time Protocol in Power System Applications," July 2011.

[G8265.1] ITU-T G.8265.1/Y.1365.1, "Precision time protocol telecom profile for frequency synchronization," October 2011.

[G8271] ITU-T G.8271/Y.1366, "Time and Phase

Synchronization Aspects of Packet Networks" February, 2012.

[NTP] Mills, D., Martin, J., Burbank, J., Kasch, W., "Network Time Protocol Version 4: Protocol and Algorithms Specification," RFC 5905, June 2010.

17. Acknowledgments

The authors would like to thank members of IETF for reviewing and providing feedback on this draft.

This document was prepared using 2-Word-v2.0.template.dot.

18. Authors' Addresses

Doug Arnold
Meinberg USA
228 Windsor River Rd
Windsor, CA 95492
USA

Email: doug.arnold@meinberg-usa.com

Heiko Gerstung
Meinberg Funkuhren GmbH & Co. KG
Lange Wand 9
D-31812 Bad Pyrmont
Germany

Email: Heiko.gerstung@meinberg.de