TIME TRAVELS: A CLOSER LOOK AT PTP

Precision Time Protocol is a system requirement for the latest IP media standards, including AES67 audio and SMPTE ST2110 video (as well as other popular formats such as Dante audio). This white paper describes key factors that must be considered for successful PTP deployments, including system architectures, PTP device configuration, and multicast network considerations. A few troubleshooting tips are also provided.

PTP – Why it's necessary

The IEEE 1588 Precision Timing Protocol (PTP) specifies a methodology for synchronizing devices to a single, shared clock across packet-based networks, including Ethernet switches and IP routers. This standard, which was originally developed for laboratory and industrial installations, is now fundamental for IP video/audio signal transport using SMPTE ST 2110 and IP audio signals using AES67, among other media applications. One of the main uses of 1588 is for creating a common timebase for multiple video cameras (GENLOCK), and also to synchronize video and audio devices to a uniform clock (time code).

PTP Message Flow

The PTP synchronization process works by exchanging a series of messages between a master device and a slave device. During this process, the slave device can determine the exact amount of network delay from the master as well as determining if its clock is offset from the master. Knowing these two values gives the slave enough data to correct its internal clock to exactly match the clock of the master. This process is repeated periodically to ensure that each slave device maintains synchronization with the master. **Figure 1**, on the next page, shows an overview of this process, which is described in the following paragraphs.

1) The master, which knows the correct time, begins the process by periodically generating a Sync message. The sync message is typically multicast to all of the devices on a particular subnet. It contains header information that identifies the master and includes a precise time stamp of when that message left the master's network interface. Depending on the hardware and software configuration of the master, the Sync information can be generated either as one message (one-step) or the master may issue two messages in rapid succession (two-step).

a) Two-step operation is required when the master is unable to include a hardware timestamp directly inside an outbound Sync message and therefore must send a Sync Follow-up message that contains the precise hardware timestamp.

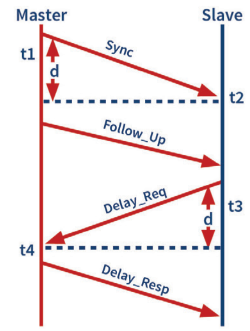
b) One-step messages are used when the master has the ability to include a precise timestamp derived from the hardware within the Sync message. Note that the selection of one-step or two-step operation is made as a configuration selection for both the master and the slave devices.

In either case, once the slave receives a Sync (and optionally the Sync Follow-up) message, the slave will be able to determine precisely when the Sync message arrived at its hardware interface and to read the message(s) to determine the precise timestamp when the Sync message left the master's hardware interface. The difference between those two timestamps represents the one way delay from the master to the slave, provided that there is no offset between the slave and master clocks.

2) The next step in the process is for the slave to send a Delay Request message back to the master. Essentially, this is a message is very similar to the Sync message, except that the header data fields are changed to show the slave device as the originator of the message.

A Delay Request message can be sent back using the same multicast group address as the original Sync message, or it can be unicast directly to the Master's IP address. In the former case, all of the other devices that are members of the multicast group will receive the message and have to parse it to determine that the message did not originate from the master. In the latter case, the slave device has to know the master's correct (non-multicast) IP address; in some cases this address needs to be set as a configuration

FIGURE 1. PTP MESSAGE FLOW SEQUENCE



$$\text{Delay (d)} = \frac{(t2 - t1) + (t4\text{-}t3)}{2}$$

$$\text{Offset} = \frac{(t2 - t1) - (t4\text{-}t3)}{2}$$

Message	Description	Slave Knows
Sync	Message to begin process. Slave measures arrival time precisely.	t2
Follow_Up	Message contains t1, which is exact time that Sync was sent.	t1, t2
Delay_Req	Slave sends Delay Request, and measures exact time of t3.	t1, t2, t3
Delay_Resp	Master sends exact time that Delay Request was sent at t4.	t1, t2, t3, t4

parameter of the slave device (which could make changing a master device more difficult). The option of using multicast or unicast technology to send Delay Request messages is typically a configuration parameter in the setup of slave and master devices.

3) Each time a master receives a Delay Request message it must send out a Delay Response message to that device. Included in this message is the precise timestamp of when the Delay Request message was received at the master's input as well as the identity of the device that sent the Delay Request message. Note that in systems with higher quantities of devices, the processing of a large number of delay request messages can become a burden for the master.

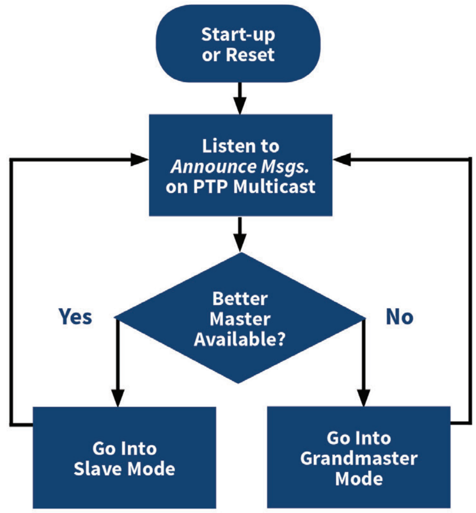
Once the slave device has received the Delay Response message, it has all the information necessary to calculate the delay from the master and to determine whether the slave's internal clock has any offset (positive or negative) from the master. The slave then uses these two values to correct its internal clock so that it is perfectly aligned to the master's clock. Very tight accuracy levels are frequently achieved in this process; slave clocks can often be aligned within a few nanoseconds of the master clock.

Distributing the Clocks

For a PTP network to operate properly, a few general network requirements need to be in place. First and foremost, the PTP synchronization process described above requires a bidirectional network capable of transmitting IP packets from the master to the slave and back. This enables all devices to receive Sync, Follow-Up, and Delay Response messages and to transmit Delay Requests.

Each master device, or device that could potentially become a master, must have the ability to receive Announce messages from every other master or potential master device. Announce messages are sent on a regular basis from the active master and all other devices that receive those messages should not attempt to become masters unless they determine that they would be a better master than the current master. All devices that could potentially become the master must execute a common Best Master Clock algorithm which is defined in the IEEE 1588 specification. Once a best master clock has been selected all other potential master devices must stop sending announce messages and should receive their timing from the common best master. This process is illustrated in **Figure 2**.

FIGURE 2. PTP MASTER CLOCK SELECTION PROCESS



Comparing two potential grandmaster clocks requires an A vs. B comparison based on multiple criteria. **Table 1** shows an ordered list of the criteria that are used each time two clocks are compared. The process works by going through each individual criteria starting from the top of the list and comparing clock A vs. clock B. if the two clocks are matched identically for a given criteria then the comparison continues by going to the next item on the list. As soon as one clock is found to be preferred to the other clock, the process stops and that clock becomes the grandmaster. The other clock automatically stops sending announce messages and remains silent as long as it continues to receive announce messages periodically from the selected grandmaster.

Within the list of criteria shown in **Table 1**, two priorities can be configured by the network administrator. Priority 1 is an eight bit field that can be used to adjust specific devices to be high or low priority; note that this criteria (since it is at the top of the list) is more important than all of the others. Priority 2 (which is shown in the fifth row of Figure 3) is a secondary 8-bit priority field that can be used to break ties between clocks that match in all of categories one through four. This criterion would be a good place to define which clock of a matched pair should get preference in a working system. Also note that criterion 6 is the ultimate tiebreaker: as long as the two clocks do not have identical MAC addresses (which they shouldn't) there will always be one clock selected over the other.

Criteria 2, 3 and 4 are related to the quality of a particular clock and the different levels are defined in the IEEE 1588 specification. Without going into excessive amounts of detail, a clock that is connected to a better quality source, one that has more accurate internal

TABLE 1. PTP MASTER CLOCK SELECTION CRITERIA

Clock Selection Criteria (in order of precedence)	Characteristic	Details
Priority Field (user provisioned)	Lower number higher priority	8-bit field, 128 for master capable, 255 for slave-only
Clock Class	Class of clock status	GPS, Free Running, Holdover
Clock Accuracy	Numberated range of accuracy to UTC	Example 25-100 ns
Clock Variance	Frequency stability	Log scale statistic representing jitter and wander of the clock oscillator
Priority 2 Field (user provisioned)	Lower number higher priority	8-bit field, can be used to indicate primary/ backup clocks
Source port ID	Ethernet MAC Address	Last resort if all other criteria are identical

electronics, and one that is more stable will be selected over other clocks. Note that some of the values used in these fields made the specific to a particular profile of clock usage. In particular, SMPTE has defined several additional clock classes (criterion 2) that go beyond what is in the basic IEEE standard.

Transparent and Boundary Clocks

Inside PTP networks, Ethernet switches provide connections between devices, including those between grandmasters and slave devices. These switches can be operated in one of three different modes – Boundary, Transparent, or non-PTP-aware. The latter state, which may be used on small networks when there is no other choice, is not recommended. When switches are not PTP-aware, the transit time for PTP messages can vary, and this will affect the accuracy of the downstream clocks. Each time a PTP Sync or Delay Request message is delayed for even a short time, it can cause an inaccurate delay or offset value to be calculated, thereby perturbing the slave's clock and affecting it's stability. Even with moderate traffic loading, an unmanaged Ethernet switch can be a source of instability in a PTP distribution system.

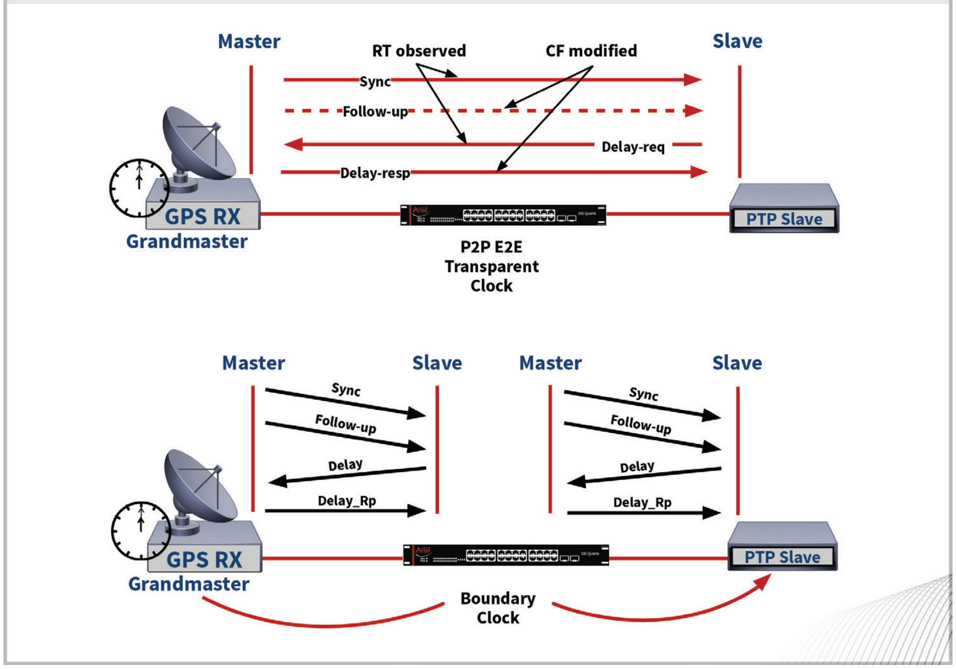
Both Transparent Clocks (TCs) and Boundary Clocks (BCs) are used in networks to properly distribute PTP messages to multiple devices. In fact, both types of devices may be present at different layers of the PTP clock hierarchy as shown in **Figure 4**. With properly configured and functioning TCs and BCs, it is possible to get thousands of devices all synchronized to a common master clock with accuracies in the nanosecond range.

Transparent clocks sit in the path that PTP (and all other) network traffic passes through and do as much as they can to make themselves invisible to the traffic (**Figure 3**). This is accomplished by first calculating the amount of time in nanoseconds that a PTP message spends in transit going through a TC device (in other works, the exact latency of the Ethernet switch). This value is then added to the correction field that is contained in PTP message headers. When the message reaches its destination, the receiving device can compensate for the TC latency using the value of this field.

Boundary clocks act as both a slave and a master in order to propagate PTP messages to large populations of slave devices (**Figure 3**). An Ethernet switch that is set up to be a BC will normally have one port that is configured as a PTP slave, and that port will be in the same network as the grandmaster or another master clock. All or some of the other ports in the switch can then be configured as BC Master ports, meaning that devices that are connected to these ports can synchronize directly to this switch. The primary benefit of BCs in a network is to reduce the amount of traffic coming from slave devices to a master. Remember that master devices will regularly receive Delay Request messages from every one of their slave devices, and must send an individual Delay Response message to each one. As the number of slave devices that depend on a single master grows, the amount of message processing time also grows. If this burden gets too large, the master device will have trouble handing the load. Hence, BCs are a common feature of large PTP networks. from slave devices to a master.

Remember that master devices will regularly receive Delay Request messages from every one of their slave devices, and must send an individual Delay Response message to each one. As the number of slave devices that depend on a single master grows, the amount of message processing time also grows. If this burden gets too large,

FIGURE 3. TRANSPARENT AND BOUNDARY CLOCK FUNCTIONS IN A PTP NETWORK



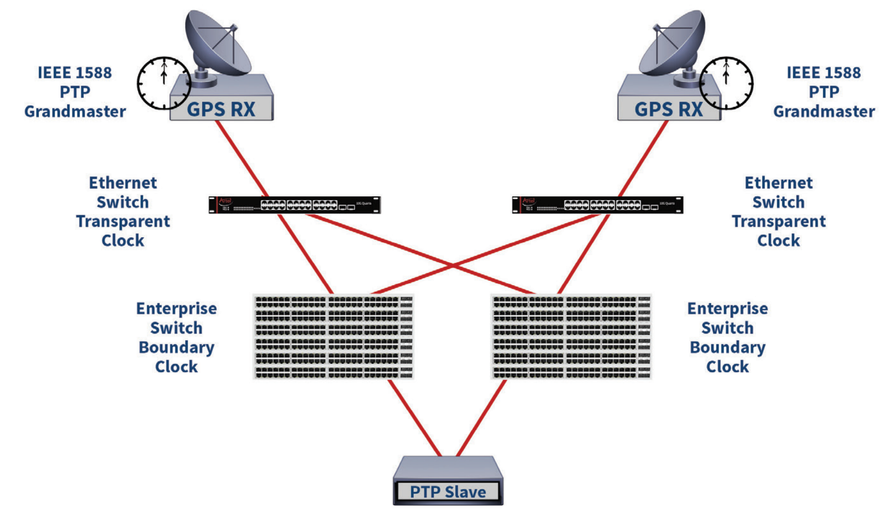
the master device will have trouble handing the load. Hence, BCs are a common feature of large PTP networks.

Redundant Systems

Most modern media networks use redundant hardware for critical system infrastructure. This practice, of course, should extend to the PTP clock distribution system as well as to the large scale Ethernet switches used to interconnect devices within these systems. Having fully redundant Ethernet backbones makes a good deal of sense and can be accommodated with the correct PTP architecture. **Figure 4** shows a simplified diagram of a media network with redundant Ethernet switches along with redundant PTP timing sources.

In this example network, two separate grandmasters are available to generate redundant PTP signals, although only one will be active at any time. Low port count transparent Ethernet switches are used to distribute the grandmaster outputs to redundant enterprise-class Ethernet switches which may act as boundary clocks. In the event of a switchover from the active to the standby network, having a grandmaster distributed across both systems will allow the slave devices to continue to operate without a clock or master discontinuity. Note that it is extremely important that the grandmaster devices are able to see messages from each other and thereby permit Announce messages from the active grandmaster to propagate to the standby grandmaster. Careful network setup may be required prevent Ethernet traffic loops from forming.

FIGURE 4. DISTRIBUTING PTP MESSAGES ACROSS A REDUNDANT NETWORK



Configuring PTP Domains and Profiles

A PTP domain is used to identify masters and slaves that are going to work together. Only devices that are in the same domain are able to operate together. Masters and slaves that receive PTP messages from other domains will ignore them. The domain number is configured into every PTP device and it is carried in the header of every PTP message. This feature of the IEEE 1588 specification allows one network to support multiple different PTP systems simultaneously, but it is hard to conceive of an application where that would be useful in a modern media network. (Note that in version 1 of the IEEE 1588 specification that the domains were named instead of numbered; however, for the current version of 1588 only numbers are used.)

A PTP profile defines many of the key operational characteristics of a PTP network. Different profiles can be used for different applications, allowing system performance to be fine-tuned to match unique requirements. For example, PTP technology can be used to synchronize operations of an electric utility power grid that stretches over hundreds of kilometers, generating long message transmission times. A suitable profile for that application could use different polling times and message timeouts as compared to a media network inside a single building. Fortunately, PTP profiles can be defined to meet performance requirements of each application independently. Message timing is one area where PTP profiles come into play. A PTP profile will define the Announce message frequency, the Sync message frequency, and the Delay Request message frequency. More frequent messages are desired for some applications, because they lead to quicker convergence of slave clocks to the masters, but they have the drawback of increasing the amount of PTP message traffic on the network.

Carefully balancing the frequency of messages against the convergence time is essential for having a well-

TABLE 2. PROFILE COMPATIBILITY BETWEEN AES AND SMPTE STANDARDS

Parameter Name	AES 67 Range	SMPTE 2059 Range	Proposed Value
domainNumber	0 to 255	0 to 127	0
logAnnounceInterval	0 to 4	-3 to 1	0 (1sec)
announceReceiptTimeout	2 to 10	2 to 10	3 (3sec)
logSincInterval	-4 to 1	-7 to -1	-3 (.125 sec)
logMinDelayReqInterval (based on logSyncInterval = -3)	-3 to 2	-3 to 2	-3 (.125 sec)

functioning, stable network. Interestingly, the AES 67 standard and the SMPTE ST 2059 standard defined two different ranges of values for these various message frequencies . A study was done in 2016 to find a common operating point for both standards and the results of that study are shown in **Table 2** above.

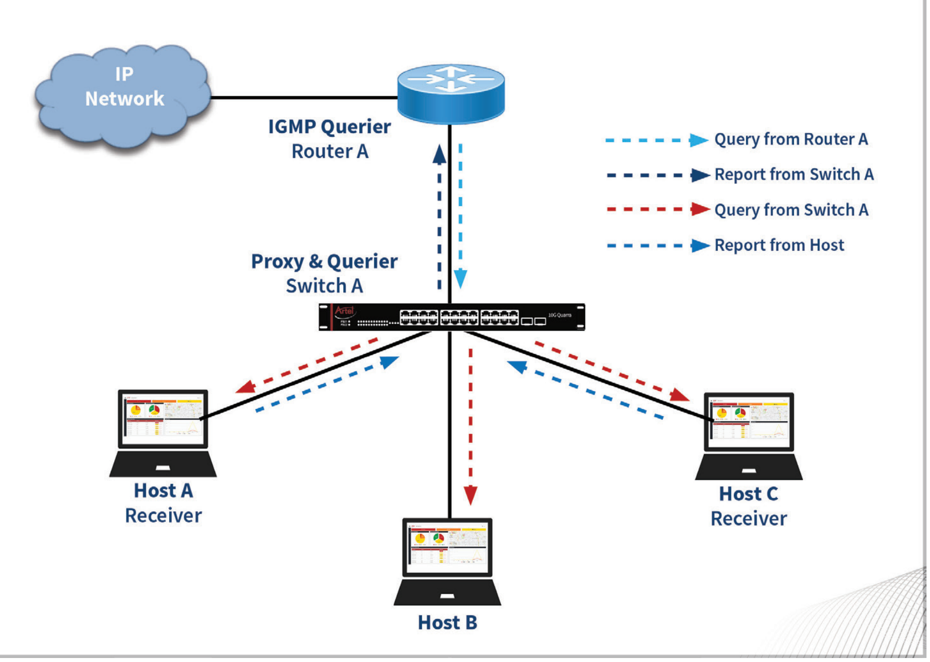
Network Considerations

A few network aspects that are somewhat peculiar to PTP systems deserve a quick mention. These are the multicast querier, the use of combined networks that have both legacy (non-PTP) devices and PTP-enabled ones, and the issue of port speed mismatches.

One of the great advances in multicasting technology was the development of IGMP snooping within Ethernet switches. This technology enables switches to listen to IGMP messages that flow between end devices and the network routers and then restrict delivery of multicast messages to only those switch ports which have devices that are participating in a specific multicast. For this technology to work, each end device must periodically generate and IGMP Membership Report message that the switch can observe. To facilitate this process, multicast networks typically include a device that will act as a querier that regularly sends out IGMP Membership Queries in order to stimulate end devices to send Membership Reports. On a pure layer 2 network, which is often the case for Media Networks, an IGMP querier is required. Only one querier should be active at any time in each subnet. Many modern multicast-aware switches have a built-in querier function that can be enabled on various ports using the appropriate switch configuration commands. **Figure 5** shows a network with multiple switches and an IGMP querier generating the necessary Membership Query messages.

During the transition from traditional video equipment to IP and PTP based video equipment there will undoubtedly be circumstances when legacy devices that do not support PTP timing need to be synchronized with new equipment. There are two ways of dealing with this situation. The first way involves using a dual function master (with both PTP and standard synch outputs) connected

FIGURE 5. LAYER 2 NETWORK WITH IGMP QUERIER



to a sync distribution network of the kind that have existed in broadcast studios for many years. The other way uses a small adapter connected to the PTP system that provides a local sync for each device that requires the legacy signal. Either method can be made to work, and both methods may be used during the transition from legacy equipment to PTP-enabled devices. Port speed mismatches occur when network connections that pass through switches or other devices that change Ethernet bit rates. For example, connecting a 100 Mbit device to a one gigabit port through the use of a format converter creates a port speed mismatch that results in a clock offset (phase error) of about four microseconds . The problem occurs because the speed of packet transit in one direction does not match the speed of packet transit in the other direction, thereby causing an offset in the PTP calculations. This can be corrected either by removing the offending device or by configuring the affected Ethernet switch ports to compensate for the timing offset.

Troubleshooting

Installing a PTP network can sometimes be a bit tricky, with many different configuration parameters that need to be set properly in order for the system to work. The good news is that once a network has been set up and is working properly, it will generally perform quite well for distributing accurate clock signals to all of the devices that are present on the network.

One of the first steps in troubleshooting a network is determining if a problem is occurring. One common failure syndrome is for a device to try to take over as a master, which is generally caused by the device not receiving Announce messages from the active master on the network. (Remember that PTP devices are supposed to take over as the PTP master whenever announcements from another master are not received, thus indicating that another master is not available.) A good way to observe whether or not this is happening is by looking at the traffic on the PTP multicast address and to see if Announce messages are coming from more than one source, or from a source that should not have become the master. One way to avoid this situation is to configure all of the ordinary devices as slave-only so that they do not try to take over as master even if announcement messages do not arrive.

Another area where problems have been encountered in laboratory tests is in mixed IGMPv2 and IGMPv3 networks. IGMPv3 has a valuable added feature called Source Specific Multicasting, which can be useful in large networks with many different sources. However, many current generation devices will only support IGMPv2, which can cause headaches on network installations. In particular, the default behavior for IGMPv3 devices is to revert to operating in IGMPv2 mode whenever messages are received from IGMPv2 devices. The best advice is to choose either version 2 or version 3 and ensure that every device on the network fully supports that version and is configured accordingly.

Troubleshooting Questions

Here is a list of some questions that may be useful to answer when troubleshooting a PTP network:

1. Is the slave device in a network segment that is included in the master's multicast?
2. Are both the slave and the master using the same PTP domain?
3. Are both the slave and the master set for either one-step or two-step?
4. Are there any network asymmetries (100BaseT <-> GigE for example) along the path between the slave and the master?
5. Can delay request messages reach the master from the slave?
6. Are delay response messages being sent multicast or unicast?
7. Has a multicast querier been set up for the PTP multicast, so that snooping data inside switches becomes properly populated?
8. Does the Grandmaster have a lower (better) priority value than other devices on network?

About the Author

Rafael Fonseca, Vice President Product Management

Rafael brings to Artel significant experience in executive management, product management and marketing, sales, and sales operations at established and startup companies. Previously, Rafael was Vice President of Product Management, Marketing, and Sales Engineering and member of the founding team at Cedar Point Communications, acquired by GENBAND specializing in technology, telecommunications, and multimedia IP solutions. Early in his career, Rafael was a member of Bell Laboratories technical team responsible for research and development in the areas of networking synchronization and equalization for data communication and optical networking systems. He holds a M.S.E.E. from Cornell University, a B.S. in Electrical Engineering from the University of Puerto Rico at Mayaguez, and a MBA Certification from Wharton School of Business.

Rafael has authored several articles and spoken at industry venues on the operational costs of delivering next generation services, media transport, audio over IP, and other topics.

Related White Paper

It's About (Precision) Time

A primer on the importance of PTP, using the SMPTE Epoch for synchronization; benefits, system components, and applications for media.

Related Products

Artel Quarra 10G PTP Ethernet Switch

Artel Quarra 1G PTP Ethernet Switch

Quarra PTP Ethernet Switches

Artel's Quarra family supports the SMPTE ST 211, standard for System Timing and Definition and ST 2059-2 permitting interoperable use of IP-based media equipment with conventional genlocked SDI equipment. The Artel Quarra switches are designed for audio/video broadcast, defense and security, finance, utilities, telecom, and enterprise IT applications in which accurate timing and control are required.

About Artel

Artel Video Systems is a world-class provider of innovative, real-time, multimedia delivery solutions serving global markets. Today the majority of the live events in the US traverse Artel products to support their mission critical work-flows. Artel's expertise in IP- and fiber-based technologies spans more than 30 years and has established Artel as a trusted partner in the development of reliable, standards-based, IP infrastructures. Artel's integrated solutions include fiber and IP based multimedia delivery, precision timing, OTT, and data networking. An employee-owned business since 2014. More information is available at www.artel.com.

World Headquarters
5B Lyberty Way
Westford, MA 01886
USA
Tel: (978) 263-5775
Fax: (978) 263-9755
Artel.com | info@artel.com

Contents subject to change without notice.
AR207-900003-00_C