

# TCP/IP for NMS Participant Input

## *Participant Interface Specification*

Prepared by:



Communications Engineering  
Planning and Development

**Document Number:**  
**Date:** 10/3/07  
**Revision:** 3.0

## **Copyright Notice**

Copyright © 2007 by SIAC. All Rights Reserved. Except as permitted under the United States Copyright Act of 1976, no part of this document may be reproduced or distributed in any form or by any means, or stored in a data base or retrieval system, without the prior written permission of SIAC.

## **Brand names and /or Trademarks**

Brand names or Products cited in this document may be trade names or trademarks. Where there may be proprietary claims to such trademarks or trade names, the name has been used with an initial capital. Regardless of the capitalization used, all such use has been in an editorial fashion without any intent to convey endorsement what so ever of the product or trademark claimant. SIAC expresses no judgment as to the validity or legal status of any such proprietary claims.

## **Engineering Services Disclaimer**

Information contained in this document is believed to be accurate. However SIAC does not guarantee the completeness or accuracy of any of the published information. This work is published with the understanding that SIAC is supplying information, but not attempting to render engineering or other professional services. If such services are required the assistance of the appropriate professional should be sought.

**REVISION LOG****Document Number:****Title:****TCP/IP for NMS Participant Input, Participant Interface Specification**

<b>Version</b>	<b>Date</b>	<b>Rev by</b>	<b>Pages affected</b>	<b>Comments</b>	<b>Approval</b>
0.2	7/18/97			Initial Draft	
0.3	7/21/97	ML	2,3	Material added	
0.4	7/23/97	ML	All	New sections	
1.0	8/7/97	ML	All	First release	
1.1	6/15/98	ML	All	Issues resolved	
1.2	6/20/98	ML	All	Additional chugs	
1.3	7/7/98	ML	All	Additional chugs	
1.4	12/31/98	RL	All	Change to be an official specification for distribution to Participants	
1.5	2/2/99	RL		Text added to section 6.2.1 regarding 4 byte header	
1.6	2/8/99	RL		- Expanded detail to section 6.2.1 regarding 4 byte header - TCP/IP connections chart added to Appendix A	
1.9	7/11/02	CE, RL	All	Removed references to legacy connections. Treatment of TCP/IP as the only option and not as the "new" option. Remove reference to data center locations. Remove references to non-Ethernet connectivity options	
1.91	7/22/02	RL	Appendix A	minor change to table	
2.05	10/7/03	RL, CE	All	Added ITS and compliance with SFTI	
2.06	10/16/03	RL	Appendix A	Table updated for ITS	
2.07	12/26/03	LG	Page 8	Addition of Pad character for ITS in message block	
3.0	10/3/07	ML, PH	All	- Removed references to ITS - Changed references to NMS mainframe to NMS system	

## Table of Contents

<b>1.</b>	<b>Introduction .....</b>	<b>1</b>
<b>2.</b>	<b>Overview TCP/IP .....</b>	<b>1</b>
<b>3.</b>	<b>Existing NMS IP Multicast Environment.....</b>	<b>1</b>
<b>4.</b>	<b>Application Overview.....</b>	<b>2</b>
4.1	Logical Connection Allocation and Fairness.....	4
<b>5.</b>	<b>NMS Input Network .....</b>	<b>5</b>
5.1	Inbound TCP/IP Support .....	5
5.2	Redundancy .....	5
5.2.1	Redundancy Example .....	5
5.3	Security.....	6
5.4	Data Pacing .....	7
<b>6.</b>	<b>Participant Requirements.....</b>	<b>7</b>
6.1	Communications.....	7
6.2	Applications .....	8
6.2.1	General Message Format .....	8
6.2.2	Safe-store Exposure.....	8
<b>7.</b>	<b>Network Layer Connectivity .....</b>	<b>9</b>
7.1	IP .....	9
7.2	IP Addressing .....	9
7.3	TCP/IP Framing .....	9
7.3.1	IP Header Field.....	10
7.3.2	TCP Header Field Description.....	10
7.4	Dual Access Network.....	11
7.5	IP Network Considerations.....	11
7.6	Data Security .....	11
<b>8.</b>	<b>Physical, Media Layer, and Network Connectivity via SFTI .....</b>	<b>12</b>
<b>9.</b>	<b>Appendix A: References.....</b>	<b>13</b>

## 1. Introduction

SIAC has deployed the NMS IP Network to support the dissemination of the NMS data to the recipient community. Since the Participants are also recipients with a connection to this network, it was determined that this network would also support the input from the Participants. Participants connect to the NMS IP network via the Secure Financial Transaction Infrastructure (SFTI – pronounced ‘safety’). With respect to connecting directly to SFTI, customers are urged to review the SFTI Interface Specification for Directly Connected Customers for more information, call 1-888-873-7422, or visit the SFTI website at <http://www.nysetransacttools.com/sfti/>

## 2. Overview TCP/IP

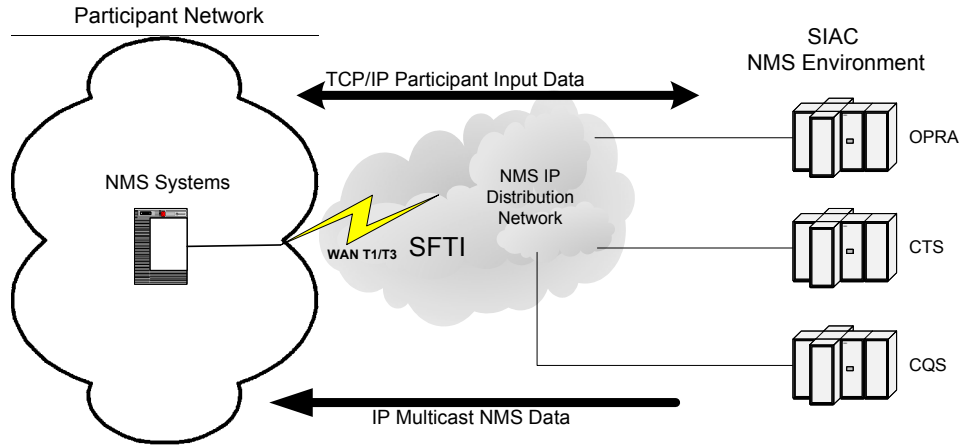
TCP/IP is a network layer connection oriented protocol that will provide a reliable network based point to point connection between the Participants input system and the NMS. Reliability in transmission is guaranteed using a handshaking mechanism in which the receiving system will send an IP acknowledgment packet back to the sending system upon reception of one or more packets. When a packet is received with errors, the acknowledgment packet can also request a retransmission from the source. Conversely, if the source system does not receive an acknowledgment for packets sent, it will assume they were never received by the destination system and will re-transmit them.

The TCP handshake mechanism also performs implicit data pacing between source and destination guaranteeing that the source system will not overload the input of the receiving system. This occurs because the acknowledgment messages have the ability to signal the source of a busy/clear condition, which effectively paces the packets across the network. The guaranteed delivery and pacing occur transparently to the applications running on both the sending and receiving systems.

## 3. Existing NMS IP Multicast Environment

Figure 1 provides a high level view of the NMS IP environment. For NMS market data dissemination the NMS hosts (OPRA, CTS and CQS) output their messages encapsulated or wrapped in a series of IP/UDP packets. These packets are forwarded into the IP distribution network with an IP multicast destination address. The multicast address selected per packet is based on the unique data line and system that the message data supports.

For a recipient system to receive a particular data line, it would send a multicast subscription message containing the mapped multicast address to their recipient router, which is connected to SIAC’s distribution network. On reception of the subscription message, the network will replicate the requested IP multicast data stream and forward it on to the recipient’s network. The Figure also illustrates how the participating market centers can connect via TCP/IP to the NMS system.



Overview of connecting to NMS hosts via SFTI and the NMS network  
Figure 1

## 4. Application Overview

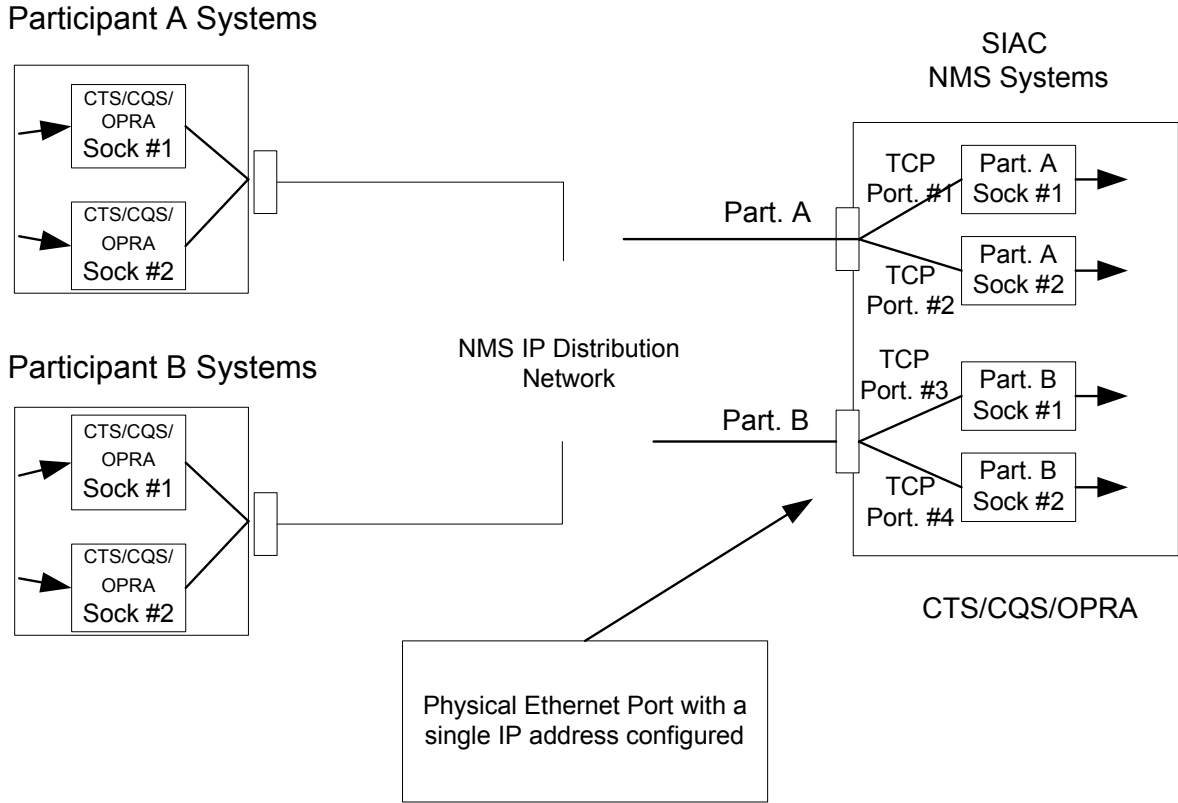
For the applications to support the TCP/IP protocol the following must be considered:

- The message format for CTS/CQS/OPRA is outlined in section 6.2
- The number of logical connections to each system by each Participant will be based on a Participant's projected maximum input rate (in messages/second) and maximum data rate that can be supported on a logical connection.
- The Participant will establish the TCP/IP connections. SIAC will always be the listener.
- A set of IP addresses will be assigned for each Participant's access to each system (CTS, CQS and OPRA), with one or more TCP/IP ports associated with each of these IP addresses.

The Participants' application systems will communicate with SIAC's NMS (CTS, CQS and OPRA) systems via TCP/IP socket connections. Sockets are established by the Participants' application system connecting to a pre-determined unique application port number in the NMS system. Socket addresses (IP address and port number) for all logical connections will be pre-assigned by SIAC for each Participant.

For CTS/CQS Participants, each will be assigned a unique physical connection on each NMS application system. Thus, each Participant has a unique IP address it will use to access the CTS/CQS systems. The IP addresses for backup and foldover connection will also be unique for each CTS/CQS Participant. Figure 2 depicts how the TCP/IP socket connections are established by the Participants' application systems for CTS/CQS.

OPRA is designed differently, such that each Participant will be given access to multiple physical interfaces that are shared among multiple Participants. Participants are assigned a unique TCP port number to which only their applications will be permitted to connect. **Error! Reference source not found.** 3 depicts how the TCP/IP socket connections are established by the Participants' application systems for OPRA.



**Figure 2**  
**CTS/CQS/OPRA TCP/IP Architecture**

#### **4.1 Logical Connection Allocation and Fairness**

For each CTS/CQS/OPRA host system, SIAC will be defining a Maximum Message Rate (MMR) allowed for a given logical connection (TCP/IP socket connection). This will be implemented in the form of an application throttle, which will keep the message rate at or below a preset rate. The MMR will be based on actual testing and known CPU capacity restrictions. It will also be system dependent (i.e., could be different for each NMS system). The application throttles will be set for each logical connection so that the total aggregate rate (across all connections) meets the Participants projected maximum message rate while not exceeding the systems capacity. The throttles do not need to be set uniformly across all of a Participant's logical connections as long as the total message rate is at or below the Participants' projections and no throttle is set higher than the allowed MMR. Several examples illustrate this approach in the next section.

Fairness between Participants is defined as each Participant always being able to input their projected maximum message rate. The NMS systems have been sized to support the projected message rate from all Participants concurrently.

## 5. NMS Input Network

### 5.1 Inbound TCP/IP Support

The NMS IP Distribution Network is used to support TCP/IP input connectivity to the NMS hosts in addition to multicast distribution as shown in Figure 1. The network, regardless of whether the packets are IP multicast framed or TCP/IP framed, properly routes packets across the network. In the case of multicast packets, the packets are forwarded via a single path to all networks that have subscribed to the multicast group listed in the IP header of each multicast packet. TCP/IP packets follow standard unicast routing algorithms as dictated by the destination IP address of each TCP/IP packet.

When routing TCP/IP packets, the Participants systems would encapsulate their application data within a TCP/IP packet and provide a destination IP address and destination TCP port number in order to direct the packet to a particular NMS host address and application port.

### 5.2 Redundancy

Though not shown in Figures 1 and 2, all recipients and Participants will connect to the NMS network and its hosts via the Secure Financial Transaction Infrastructure (SFTI). There is a separate network interface specification available that describes in more detail how to connect to the edge routers of SFTI that are currently available for direct connectivity via several access centers in the United States. By connecting via the SFTI Access Centers, customers are provided routed access into the NMS distribution network two data centers. Once the packet is in the network, it has equal access to all of the NMS hosts regardless of their physical site location. The packets will be properly routed based on destination address.

Participant application hosts will be responsible for initiating the connections to the NMS host systems.

With regard to the NMS hosts themselves, multiple network-based connections on each system will be maintained to support redundant connectivity to the Participants in the event of a controller or other type of connection failure.

#### 5.2.1 Redundancy Example

To illustrate the resiliency of the TCP/IP input environment Figure 3 shows a single Participant input system redundantly connected into the NMS networking environment. It can be assumed in this example that the customer has WAN connectivity between its network and at two SFTI Access Centers. The Participant system is dual homed to customer networks  $y$  and  $z$  having source addresses  $y.I$  and  $z.I$  respectively. The NMS host connections supporting this Participant are shown on dedicated subnets,  $a, b, c, d, e, f, g, h$ .

Using OPRA as an example, the Participant could make a TCP/IP input connection from  $y.I$  to  $a.I$  on OPRA A. If  $a.I$  or the switch supporting  $a.I$  were to fail, the connection would time out, and the Participant would re-establish the connection from  $y.I$  to the backup interface  $b.I$ . For added redundancy the switch supporting  $b.I$  is different than the primary,  $a.I$ . If OPRA Primary needed to fold into OPRA Back-up, the Participant would re-establish a connection from  $y.I$  to  $g.I$ . Note that the Participant could source his connection from either  $y.I$  or  $z.I$ . The NMS network provides routable access to all system connections regardless of its ingress point into the NMS network via SIAC's two data centers. Note that this example holds true for CTS and CQS as well.

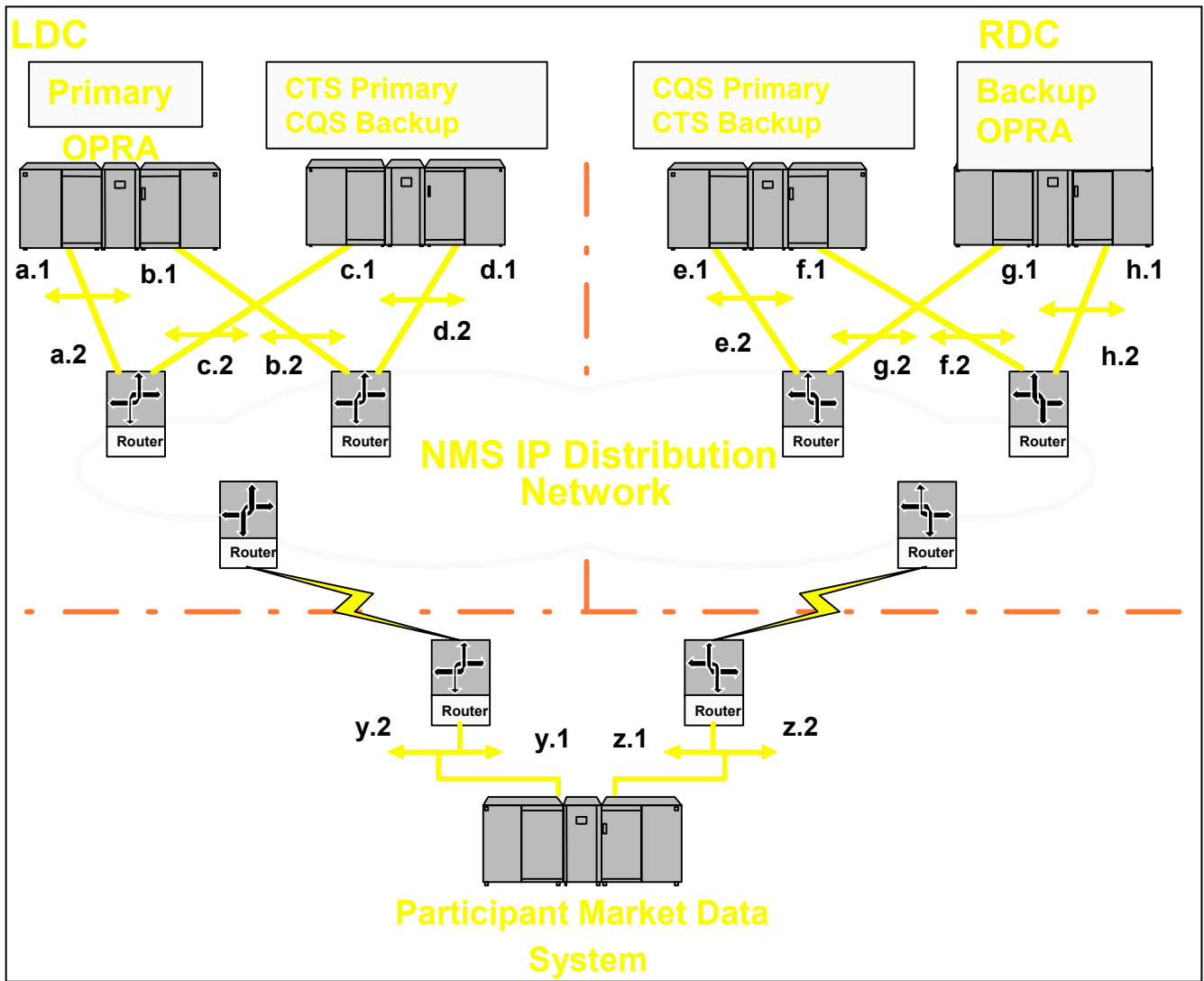


Figure 3

### 5.3 Security

Network security within the NMS IP Distribution Network is provided with the use of router filters installed in the NMS and SFTI routers at appropriate points in the network path. These packet filters create an access policy of sorts between each Participant and their logical connections to the NMS hosts. In this configuration, only valid TCP/IP communications will be allowed. Anti-spoof filters will also be implemented for preventing mis-configured or hostile systems from impersonating a valid system and/or connection.

## **5.4 Data Pacing**

The OPRA, CTS and CQS applications are set to read data received from Participants (via the TCP/IP stack) at a preset maximum rate to assure that each Participant is given appropriate and fair access to NMS system resources.

The TCP/IP protocol itself performs implicit data pacing between the Participants input system and the NMS hosts via the use of the TCP/IP window size advertisement and back-off timers. TCP/IP's window size pacing mechanism becomes active if a Participant is sending messages at a faster rate than the NMS application can process them and thus causing the NMS host's TCP/IP input buffer to fill up (the same is true in the other direction, i.e., NMS host to Participant host). The window size advertisement lets each end of the connection know how many bytes of data can be sent without losing data. Depending on a combination of factors including network WAN speeds, system capacities and network congestion, the TCP/IP window size adjustments can result in data pacing.

The protocol also uses packet acknowledgements for identifying when retransmissions are required. However if retransmissions occur frequently, the protocol's algorithms will automatically "back-off" from sending data at the existing rate, i.e., moving to a slower rate, until successful TCP packet acknowledgements resume and the back-off timers have readjusted to normal processing.

## **6. Participant Requirements**

### **6.1 Communications**

A Participant is expected to connect to the NMS IP Distribution Network via the SFTI access centers. Participants should consult the SFTI Interface Specification for Directly Connected Customers for more information, call 1-888-873-7422, or visit the SFTI website at <http://www.nysetransacttools.com/sfti/>.

## 6.2 Applications

The Participant input systems will need to support TCP/IP. In addition, support for a 4 byte block length header will be required (see below).

### 6.2.1 General Message Format

The following general message block formats are necessary to support TCP/IP protocol for NMS Participants.

#### CTS/CQS

<b>Block Length (4 Bytes)</b>	STX	Block Header	US	msg #1	US	msg #2	US	.....	msg #n	ETX	*PAD
-------------------------------	-----	--------------	----	--------	----	--------	----	-------	--------	-----	------

#### OPRA

<b>Block Length (4 Bytes)</b>	SOH	msg #1	US	msg #2	US	.....	msg #n	ETX	*PAD
-------------------------------	-----	--------	----	--------	----	-------	--------	-----	------

- The 4-byte **Block Length** Header containing the size of the block (in bytes) is required at the beginning of every block. The header contains two binary (not ASCII) fields. The first binary field is two bytes (16 bits) in length and indicates the length of the block. The byte ordering is left-to-right, that is, the left byte is the Most Significant byte and the right byte is the Least Significant byte (Big-Endian format). The second binary field is also two bytes (16 bits) in length and is reserved. The second field will be set to binary zero (NULL).
- In the case of OPRA, CTS, and CQS, the **Block Length** calculation includes the sum of the byte counts associated with the SOH and ETX characters, the PAD character, if present, and the 4-byte Block Length header itself. Thus, the maximum length for a block from a Participant to the NMS hosts is increased from 1,000 to 1,004 bytes and the maximum length for a block from the NMS hosts to a Participant is increased from 300 to 304 bytes.
- The Sequence Number field in the message header, **for OPRA only**, is equal to 7 digits.
- For OPRA, CTS and CQS, it is necessary for a block to consist of an even number of bytes. In the event that the block does not consist of an even number of bytes, a PAD character (Hex 'FF') will follow the ETX character at the end of the block.

### 6.2.2 Safe-store Exposure

With the TCP/IP protocol, the Safe-Store before Acknowledgment cannot be guaranteed.

When using TCP/IP protocol, the application system has no control of the TCP/IP ACK messages. If the receiving system, in this case SIAC's system, experiences a problem processing input data, the messages that have been acknowledged by the TCP/IP stack but have not been passed to the application system would be lost and the Participant's system would not have this information.

The number of messages that could be potentially lost is determined by the buffer space allocated for the specific TCP/IP socket on the SIAC system. To deal with this unlikely event, the NMS input applications already perform sequence number checking which will notify the Participant when a sequence gap is detected. For OPRA, CTS and CQS, this is an automated process.

## 7. Network Layer Connectivity

### 7.1 IP

The Internet Protocol suite, referred to as IP, defines a data encapsulation method that allows data to traverse multiple networks through intermediate network devices known as routers. Addressing between end stations is based on a source and destination IP address. Routing packets (data frames) between networks is a point-to-point operation; involving one source, one destination, and one or more intermediate routers.

### 7.2 IP Addressing

End stations participating in the NMS input network must use unicast IP addresses within the Class A, B, or C address ranges which includes 1.0.0.0 through 223.255.255.255 (using standard IP address notation). When an IP router encounters a packet with a destination unicast Class A, B, or C address, the packet is routed using the standard IP routing table when determining the next hop interface or destination network to send the packet.

### 7.3 TCP/IP Framing

The application data is encapsulated in a TCP/IP frame as shown in Figure 7.3-1. The IP datagram includes the IP and TCP headers plus the application data block as described in section 2 above. The datagram fields can be read left to right starting at the top and working your way down through the datagram. The size of each field (excluding the TCP data field) is represented in bits across the top and bytes going down. Bits are transmitted across the link starting with bit 0, 1, 2 and so forth. This is called the “big endian” representation where the most significant bits are transmitted first.

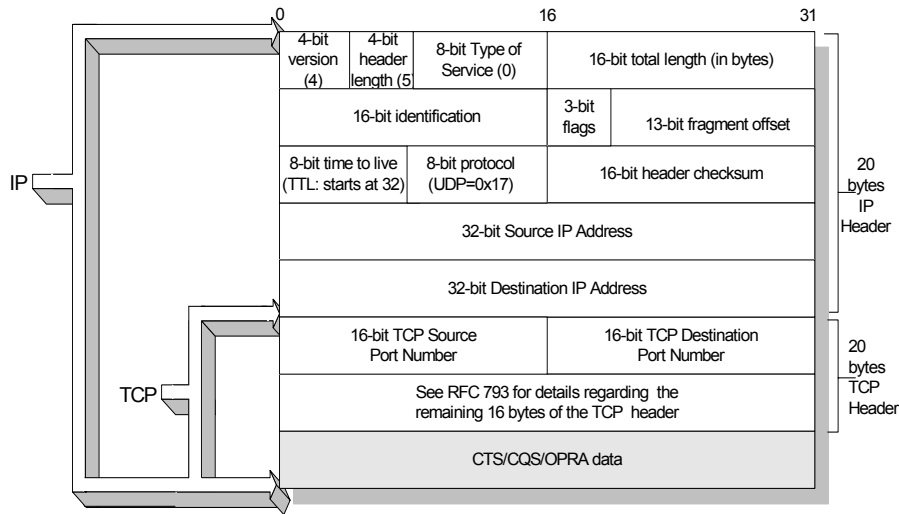


Figure 7.3-1 TCP/IP Datagram Format

### 7.3.1 IP Header Field

- **Version** - This is a 4 bit field which defines the current version of the IP protocol. It is currently set to 4.
- **Header Length** - This 4 bit field contains the number of 32 bit words in the IP header portion of the datagram. For all multicast packets being generated by this network the IP header will be 20 bytes long, which means this field will contain the value 5.
- **Type of Service** - The first 3 bits are the precedence sub field and are ignored by most network equipment. The next four bits are flags that define minimize delay, maximize throughput, maximize reliability, and minimize monetary cost respectfully. They are set to zero (0) for this application. The last bit is always set to zero. Based on this description this field will always have the value of zero (0) for all multicast packets.
- **Total Length Field** - This 16 bit field contains the length in bytes of the entire IP datagram. This includes the IP and TCP header plus the application data (TCP data).
- **Identification Field** - This 16 bit field contains a value that is incremented by one for each packet sent by the source system. It only has relevance on the receiving system when packets are either fragmented and/or TCP is used as the transport protocol.
- **Flags and Fragment Offset** - The combined 16 bit field is only used when an IP datagram is fragmented.
- **Time to Live (TTL)** - This 8 bit field contains a value that determines the number of routers that this datagram can pass through. Each router that forwards this datagram will decrement this value by one; when it reaches zero the next router throws it away.
- **Protocol** - This 8 bit field contains a value representing the next level encapsulated protocol. In this case it is TCP, which has a value of 6.
- **Header Checksum** - This 16 bit field contains a checksum made up of the IP header fields only. The calculation is based on the ones complement sum of the header broken into 16 bit words.
- **IP Source Address** - This 32 bit field contains the IP address of the source system.
- **IP Destination Address** - This 32 bit field contains the IP address of the destination system.

### 7.3.2 TCP Header Field Description

SIAC will be expecting Participant applications to connect to pre-defined IP destination addresses and IP destination port numbers. The destination IP addresses will correspond to physical NMS system ports. The destination port numbers will be assigned to each Participant by SIAC. For detail regarding the TCP protocol and header information, the reader should refer to RFC 793.

- **TCP Source Port Number** - This 16 bit field identifies the sending process within the Participant source system. It is set by the source system.
- **TCP Destination Port Number** - This 16 bit field identifies the TCP process waiting for a connection request on the SIAC NMS system. Each Participant will be given a pre-defined set of TCP destination port numbers specific to that Participant. Participant systems must use these TCP destination port numbers when opening connections to the NMS system. Connections not matching the pre-defined values will be blocked by SIAC.

#### **7.4 Dual Access Network**

In order to provide a resilient/redundant networking environment the Participant is provided with the ability to connect to SIAC's data centers via multiple SFTI Access Centers.

#### **7.5 IP Network Considerations**

The IP networks making up the connection between the Participant's location and the SFTI edge router ports located at SFTI Access Centers are the responsibility of the Participant and/or their third party access providers per agreements between them. With respect to the connection to the SFTI edge port, the Participant is required to provide a publicly registered IP address for both ends of this connection, which includes the router port on the SFTI Edge router. If the Participant doesn't have a publicly registered address to use, SIAC will provide a range of private addresses to use. This and related IP address requirements are explained in more detail in the SFTI Interface Specification for Directly Connected Customers.

#### **7.6 Data Security**

SIAC is protecting its network and hosts using packet level filtering on the SFTI and NMS routers. Routing policies are implemented on SIAC managed routers to insure strict control over the IP routing table. These security measures maintain the integrity of network routing tables and also protect SIAC's network and hosts from intentional or accidental access by a Participant network. These measures are in no way intended to provide the same level of security to the Participants themselves. If a Participant believes that additional security is required to protect their network then it is left to them to implement it.

## **8. Physical, Media Layer, and Network Connectivity via SFTI**

Customers must connect to the SFTI network via one of the SFTI Access Centers in order to access NMS systems. At each of the Access Centers will be SFTI edge routers that provide customers with an Ethernet port to which they can connect (100Base-T or 1000Base-SX).

The device connected to the SFTI Edge Router must be a router or Layer 2 switch capable of tagging packets with VLAN membership tags as per the 802.1Q standard. Customers must configure their switches to perform VLAN tagging such that the SFTI Edge Routers can recognize the tag and forward the incoming packets appropriately.

For additional information on connectivity to SFTI, please refer to the Secure Financial Transaction Infrastructure (SFTI) Network Interface Specification for Directly Connected Customers.

## 9. Appendix A: References

### IP related standards documentation:

RFC 768	User Datagram Protocol (UDP)
RFC 791	Internet Protocol (IP)
RFC 792	Internet Control Message Protocol (ICMP)
RFC 793	Transmission Control Protocol (TCP)
RFC 826	Address Resolution Protocol (ARP)
RFC 903	Reverse Address Resolution Protocol (RARP)
RFC 950	Internet Subnetting Procedures
RFC 1042	IP over IEEE 802 Networks
RFC 1075	Distance Vector Multicast Routing Protocol (DVMRP)
RFC 1112	Host Extensions for IP Multicasting (includes the Internet Group Management Protocol (IGMP))
RFC 1122	Requirements for Internet hosts - communication layers
RFC 1123	Requirements for Internet hosts - application and support

### Ethernet standards documentation:

#### Physical Layer:

##### IEEE

- 802.3i for the 10Base-T specification
- 802.3u for the 100Base-T specification

#### Media-Access Layer:

- Ethernet Version 2.0 frame formats as specified by Digital Equipment Corporation, Intel Corporation, and Xerox Corporation.