NOTE: This is a working draft document, and is subject to rapid change. Copies of this document will be distributed through the X12C TG1 list server.

ABSTRACT
This guideline provides a technical foundation for using TCP/IP communications to transfer EDI interchanges between business partners. TCP/IP is the primary transport mechanism used on the Internet. This guideline describes basic methods available for EDI communications using TCP/IP on the Internet, and outlines options for using these methods.
Table of Contents

1 Preface..................................................................................................................................... 4

2 Introduction............................................................................................................................. 5
  2.1 Layering of Protocols....................................................................................................... 6

3 MIME Encapsulation............................................................................................................... 7
  3.1 MIME Parameters............................................................................................................ 7
    3.1.1 Data Content............................................................................................................. 7
    3.1.2 Character Set............................................................................................................. 7
    3.1.3 Transparency Flag..................................................................................................... 7
    3.1.4 Physical Record Characteristics................................................................................. 7
    3.1.5 File Length................................................................................................................ 8
    3.1.6 User Defined Variables.............................................................................................. 8
  3.2 MIME Object Characteristics............................................................................................ 8
  3.3 MIME Security................................................................................................................. 8

4 SMTP Interchange Guideline....................................................................................................9
  4.1 Advantages and Disadvantages......................................................................................... 9
  4.2 Level of Service................................................................................................................ 9
  4.3 Routing Information....................................................................................................... 10
  4.4 SMTP Parameters........................................................................................................... 10
    4.4.1 Version Numbering................................................................................................. 10
    4.4.2 Application Code.................................................................................................... 10
    4.4.3 Message Class......................................................................................................... 10
    4.4.4 Sender..................................................................................................................... 10
    4.4.5 Receiver.................................................................................................................. 11
    4.4.6 Originator............................................................................................................... 11
  4.5 Security.......................................................................................................................... 11

5 HTTP Interchange Guideline.................................................................................................. 12
  5.1 Advantages and Disadvantages....................................................................................... 12
  5.2 Level of Service.............................................................................................................. 12
  5.3 Routing Information....................................................................................................... 12
  5.4 Parameters...................................................................................................................... 12
  5.5 Security.......................................................................................................................... 12

6 FTP Interchange Guideline..................................................................................................... 13
  6.1 Advantages and Disadvantages....................................................................................... 13
  6.2 Level of Service.............................................................................................................. 14
  6.3 Routing Information....................................................................................................... 14
  6.4 Other Features................................................................................................................ 14
    6.4.1 Transparency........................................................................................................... 14
    6.4.2 File Management.................................................................................................... 14
    6.4.3 Error Handling........................................................................................................ 14

ANSI ASC X12C, TG1  2  01/31/97
TCP/IP Guidelines
EDI TCP/IP GUIDELINE

1 Preface

In the past few years, the rapid growth of the Internet as a communications media has prompted many attempts to implement EDI communications using this public network. This document outlines the ASC X12C Subcommittee on Communications and Controls’ attempt to define a framework for transferring electronic data interchange (EDI) information using the Internet and TCP/IP.

The ASC X12C Subcommittee on Communications and Controls recognizes that much of the standards work for interchange of Electronic Commerce data over the Internet is being addressed by the EDIINT working group of the Internet Engineering Task Force (IETF). IETF is the body that develops standards for communications and protocols on the Internet. ASC X12C will update these guidelines to track the work done by the EDIINT group as the proposed standards become more solidified. These guidelines are intended to promote communications and interchange interoperability among disparate systems.
2 Introduction

This guideline covers EDI interchanges using the following Internet communications methods:

1. SMTP (Simple Message Transfer Protocol, or Internet email)
2. FTP (File Transfer Protocol)
3. HTTP (HyperText Transfer Protocol)
4. TCP Socket (Session based direct communications)
5. Telnet (Remote terminal session)

Each of the methods will be outlined. A list of advantages and disadvantages of each method will be presented followed by suggested guidelines for use. Security for each method will also be addressed.

For the purposes of this guideline, Electronic Data Interchange (EDI) is defined as the computer to computer transmission of common business documents by electronic means in a standard format. The standards are embodied in documents published and maintained by the American National Standards Institute Accredited Standard Committee X12 (ANSI ASC X12) and by the United Nations Electronic Data Interchange For Commerce and Trade (US/EDIFACT).

An Internet Host is any computer that supports Internet protocols. An Internet Provider is a company that provides data transport services between Internet Hosts.

The Internet is made up of computers operated by cooperating organizations linked by dedicated telecommunications transmission equipment; lines and routing devices. There is no governing body to monitor and enforce levels of service. Levels of service are provided on a voluntary basis by organizations that choose to follow the RFC’s. Transmission Control Protocol/Internet Protocol (TCP/IP) are the transport and networking protocols defining electronic transmission over the Internet. The Internet Protocols (IP) define packet switching protocols which allow for the reliable transmission of data packets from point to point. The Transmission Control Protocol (TCP) is an end to end protocol which divides messages into packets to be transmission at the sender by IP andreassembles those messages at the recipient.

Data transmissions traverse the Internet in packets which travel unpredictably from Internet host to Internet host. There are no provisions in TCP/IP protocols for routing, timely delivery or other levels of service provided by traditional EDI Value Added Networks (VANS). Portions of the Internet wholly under the control of a primary backbone supplier may be more closely controlled for routing, delivery time, etc. than the entire “public” Internet.

Due to the nature of the Internet and the lack of enforcement of standards by a governing body, it is very difficult to guarantee the levels of service provided by traditional EDI Value Added Networks (VAN). Some Internet providers do provide enhanced service by controlling the end to
end path of each message. Messages do not travel on the “public” Internet, but instead travel
within these “private” Internets if possible.

2.1 Layering of Protocols

The layering of data transport protocols in the context of providing EDI has been an issue in the
EDI community for many years. There is a clear delineation as to the responsibilities of the
various software processes that manipulate EDI data. EDI data is clearly in the application layer
(refer to the Open Systems Interconnect model). SMTP lies with the TCP layer which is
analogous to the transport layer of OSI.

Information giving semantic meaning to EDI data should remain in the application layer. Pieces
of information, such as sequence numbers are already contained in EDI envelope structures and
with EDI data elements. Applications using Internet protocols to transport EDI data should not
rely on Internet headers to determine this type of information.
3 MIME Encapsulation

The IETF working group standards are developed around the use of MIME encapsulation on a variety of transports. For the purposes of this document, MIME encapsulation will be recommended for SMTP, HTTP and FTP transports.

3.1 MIME Parameters

The following tags and parameters are allowed in SMTP messages.

Need more information on MIME, tags, etc. here.

3.1.1 Data Content

Data content can be specified and ascertained by the “Content Type: Application” header. The following are supported by RFC 1767:

    Content-Type: Application/EDIFACT
    Content-Type: Application/EDI-X12
    Content-Type: EDI-Consent

The last item, “Content-Type: EDI-consent” can be used to specify industry standards, or formats agreed by trading partners by mutual consent. Other Content Types can be included in the RFC if there is enough usage to warrant inclusion as an Internet standard.

3.1.2 Character Set

RFC 822 specifies the use of the ASCII character set. EDI objects need to be encapsulated into MIME object for many reasons including this one. A MIME object will not be interpreted by intervening hosts as they pass the message through the Internet and therefore do not need to interpret the EDI message. The character set is only the concern of the sending and receiving applications and are not the concern of any intervening ‘MTA’s.

The character set can be specified in the MIME envelope. Use the “Content-type: charset=” parameter.

3.1.3 Transparency Flag

There is no data transparency indication in Internet SMTP RFC’s at this time. A data transparency flag can be added by amending RFC 1767.

3.1.4 Physical Record Characteristics

Traditional Internet hosts were usually using the UNIX operating system. Modern Internet Hosts can be UNIX or PC based using Microsoft Windows NT, Novell Netware or other operating systems. These systems usually treated all data as stream data, that is as a continuous stream of bytes and are not interpreted as records.
Legacy and mainframe systems are not stream oriented, but need information to determine record and field characteristics. Extensions need to be made to RFC 1767 to add physical record characteristics.

### 3.1.5 File Length

Specifying file length in the context of Internet messages could lead to some inconsistencies. Internet messages can be encoded such that the length of the resulting message is different from the original message. The message length would then have to change as the messages passes from host to host or the message length would have to be specified as the original message length.

File length is not specified in X12 or EDIFACT standards.

Therefore, File Length might not be required.

### 3.1.6 User Defined Variables

User defined fields can be defined by trading partners as needed. Additional header fields can be defined as long as they do not conflict with existing published header fields. These fields will be ignored by intervening hosts.

Care must be taken, however, with user defined fields. It is possible that specific header fields could be pre-empted by subsequent Internet publications. It is incumbent on the EDI community to register these fields and be aware of the evolving Internet standards.

### 3.2 MIME Object Characteristics

The MIME standard does not specify if the object contains wrapped or unwrapped data. It is possible that the nature of Internet data transport will make this specification unnecessary.

### 3.3 MIME Security

Need to discuss encryption and signatures of MIME envelopes here.
4 SMTP Interchange Guideline

SMTP (Simple Message Transfer Protocol) is an Internet standard developed to facilitate the transport of electronic mail. It is in widespread use in many implementation and on many computer platforms. The SMTP protocols are defined in Request For Comment (RFC) documents number 821 and 822. There is a lot of software written to these standards and many are bundled with versions of UNIX, Windows NT and other operating systems. There are also many implementation of software implementing SMTP that is commercially available.

SMTP is a store and forward, text message based protocol. SMPT mail messages can travel unpredictably from Internet host to Internet host. Each intermediate Internet host will receive the message and attempt to forward it to another host closer to the intended recipient until it is delivered to that recipient.

SMTP was originally written to transport human readable text in 72 character column format in the clear (with no security). The Multipurpose Internet Multimedia Extensions (MIME) protocol (RFC 1896), was written to facilitate the transport of non-textual binary object using SMTP protocols. Security is provided by adding features to the messages and sending them using SMTP and MIME.

4.1 Advantages and Disadvantages

add in notes from June meetings here.

4.2 Level of Service

SMTP providers can only guarantee levels of service from end to end when the entire transmission path is owned by the Internet Service Provider. The levels of service available to SMTP users is as follows:

1. “Time Of Delivery “ - Time of delivery cannot be guaranteed by Internet Service Providers using normally available services. Therefore message expiry should be specified.

2. “Message Expiry” - Message expiry is defined as the time at which a message is no longer valid; the information in the message should not be acted upon. For example, if a Purchase Order was delivered after the specified expiry date, the goods or services should not be delivered.

3. “Expiry Date Indication” - Expiry Date Indication is not directly supported by RFC 822 which specifies that no automatic action should be expected. Therefore Message
Expiry should be implemented by the receiving Message Transfer Agent (MTA). Care should be taken that expired documents are not opened when not authorized.

The enforcement of Message Expiry is dependent on the receiving MTA and therefore is out of the control of the sending MTA. The sender then has no guarantee that expired messages will not be viewed.

In light of the limitations on Message Expiry, documents sent with Message Expiry dates should be such that can be viewed after the expiration date. The receiving firm must understand that such messages are no longer valid and should not be acted upon. The sending firm must understand that expired messages might be viewed.

### 4.3 Routing Information

Messages cannot be routed on the “public” Internet due to the way in which the Internet was designed and evolved. The Internet is made up of cooperating organizations whose computers and routers are linked in a mesh topology. Internet protocols dictate that messages and packets “find” the least cost path from sender to recipient. The protocols are written so that if a network host fails, other hosts substitute and forward data. When failed hosts come back on-line, the routing of messages and packets can change again.

Messages requiring specific routes, for security or other reasons, should not be sent on the public Internet. Such messages should be sent on private networks or by “VAN”s that can guarantee routing.

A receiving MTA can determine the route that a message took either by examining the forward-path list in the SMTP MAIL command or by examining the “Return Path” header.

### 4.4 SMTP Parameters

The following tags and parameters are allowed in SMTP messages.

**4.4.1 Version Numbering**

The version number specified in the “Message-ID” SMTP Reference Field is used to indicate the version of the SMTP handler, and cannot be used to indicate the version of the EDI data being sent.

**4.4.2 Application Code**

Not yet included in Internet RFC’s.

**4.4.3 Message Class**

Not yet included in Internet RFC’s

**4.4.4 Sender**

The sender can be determined by examining the “FROM” SMTP command or the “From:” field in the message header.
4.4.5 Receiver
The receiver can be determined by examining the “RCPT” SMTP command or the “To:” field in the message header.

4.4.6 Originator
The originator can be determined by examining any of the following SMTP commands: “MAIL FROM”, “SEND FROM”, “SAML FROM”. Or the “Reply To:” field in the message header.

4.5 Security
There are security facilities available to Internet users:

Any MIME encapsulated object can be encrypted.
Originator authenticity can be ensured by the use of digital signatures.
Internet and SMTP headers must remain in the clear in order for the message to be routed through the Internet. Hosts must be able to decipher the headers in order to forward messages.

If messages must be routed along secure paths, trading partners must arrange for such services to be provided by VAN’s or other network providers that do not use the public Internet.
5 HTTP Interchange Guideline

HTTP (HyperText Transfer Protocol) is an Internet standard developed to facilitate the transport of electronic mail. It is in widespread use in many implementation and on many computer platforms. The SMTP protocols are defined in Request For Comment (RFC) documents number 821 and 822. There is a lot of software written to these standards and many are bundled with versions of UNIX, Windows NT and other operating systems. There are also many implementation of software implementing SMTP that is commercially available.

5.1 Advantages and Disadvantages

add in notes from June meetings here.

5.2 Level of Service

5.3 Routing Information

5.4 Parameters

5.5 Security
6 FTP Interchange Guideline

File Transfer Protocol (FTP) is a real-time protocol that requires both the local and remote systems to be active at the same time; an Internet Service Provider (ISP) or VAN may be used as an intermediate party if direct logins are not permitted or if the systems are not operational at the same time. FTP can be used to send or to retrieve data.

FTP is widely available and there are numerous user applications available. FTP may be used in conjunction with other protocols, such as HTTP/WWW, or Telnet.

FTP is a connection-oriented protocol. Two virtual circuits (channels) are established for FTP: one for control, and one for the data transfer.

The FTP protocols are defined in Request For Comment (RFC) documents ??? (need research here). FTP is implemented and supported under a wide variety of operating systems.

ANONYMOUS FTP

One login convention is referred to as “anonymous FTP”, where a login name of “anonymous” is used, and no particular password value is required.

Anonymous FTP is useful for distribution of publicly-available EDI documents, where the files are set with read-only file permissions. One example is a Request for Quote (RFQ, 840) transaction set (TS) that is made available for any bidder.

Anonymous FTP could also be used if the EDI documents were write-protected and the contents were encrypted, so that only authorized users would be able to decrypt and process the EDI documents.

A drop-box may be used, where a user logs on anonymously and deposits files in a directory on the remote system. With a drop-box, the user is not able to list the contents of the directory or to delete existing files. Some file name conventions may be needed to ensure that a file being deposited does not clobber an already existing file in that drop-box.

6.1 Advantages and Disadvantages

FTP requires the originating system to know the login and password of the remote system, and to actively log onto that system. A different login ID should be allocated for each user that is to connect to a system. The user permissions should be restrictive, to only allow that login ID to do the minimum that is necessary. Possible FTP activities include:

- upload file
- download file
One system may have to poll the remote system on a regular basis to see if any EDI data has been posted in the directory waiting for pickup. Alternatively, each system can initiate the FTP connection to push data to the other system. This approach may be impractical if there are many systems to be polled or to connect with.

6.2 Level of Service

6.3 Routing Information

User-specified routing is not supported. The remote system’s IP address or name is used in establishing the connection, and the connection is established using TCP/IP.

6.4 Other Features

6.4.1 Transparency

FTP supports two modes of transfer: binary and text data. Binary mode should be used if there is any doubt about the type of data contained in the file; FTP transfers in text mode will corrupt any control characters, other than a few common characters such as carriage returns and tabs. Binary mode may also be needed if EBCDIC data is being transferred (???).

6.4.2 File Management

The trading partner agreements should include which party will delete files once they have been retrieved.

6.4.3 Error Handling

The FTP protocol should recover from transient network errors. During the data transfer phase, an FTP system will repeatedly resend a packet until it receives an acknowledgment. If after many attempts (e.g., over 6) the packet has still not been acknowledged, FTP may drop the data channel. The application should check the file size of the transferred file to verify that the entire file was copied successfully.

The TA3 can be used for acknowledgment of transferred data and to report detected errors. To ensure that undetected communications errors do not result in data loss, the absence of receipt of a TA3 should result in retransmission of the interchange.

A TCP/IP Ping may be used to determine the error rate that is currently experienced on the network. If the error rate is high, corrective action or an operator alert is necessary, or the FTP session may be deferred until later.

Procedures should be established for handling of duplicate copies of EDI transactions. (This is a general statement.)
6.5 Security

Several options are available for security (to be elaborated later):
- Secure Sockets Layer (for data confidentiality during the transfer) (also authentication of the connection??).
- File encryption (for data confidentiality).
- X12.58 (for data confidentiality, data integrity, non-repudiation of origin).

Also, Firewalls.

The password is sent in the clear. One solution is to make passwords useless after they have been used once, such as the generation of one-time passwords (using a smart card, for example). Kerberos for protection of passwords within an organization or intranet???
7 TCP Sockets Interchange Guideline

7.1 Advantages and Disadvantages
add in notes from June meetings here.

7.2 Level of Service

7.3 Routing Information

7.4 Parameters

7.5 Security
8 Telnet Interchange Guideline

8.1 Advantages and Disadvantages
add in notes from June meetings here.

8.2 Level of Service

8.3 Routing Information

8.4 Parameters

8.5 Security
References

Simple Mail Transfer Protocol
RFC822: Standard For The Format Of ARPA Internet Text Messages
RFC1082: Post Office Protocol
RFC1767: MIME Encapsulation Of EDI Objects
RFC1865: EDI Meets the Internet, Frequently Asked Questions About Electronic Data Interchange (EDI) on the Internet

Appendix A - SMTP Examples
Appendix B - HTTP Examples
Appendix C - FTP Examples