IEN 175

J. Postel
ISI
13 March 1981

Internet Meeting Notes -- 28-29-30 January 1981

I. WELCOME

The meeting was held at USC Information Sciences Institute. The host
was Jon Postel. Jon opened the meeting and gave the usual
explanations about the arrangements.

Vint Cerf extended the welcome and called special attention to the
participation of representatives from CSNET, Canada, and Germany.

II. OBJECTIVES

Vint Cerf described the main objective of the meeting to be to put
issues on the table. Problems could be resolved by subsequent
arrangements between ARPA and the contractors. In particular, the
topics of performance, addressing, and documentation are of special
interest.

III. STATUS REPORTS

A. UCLA

Charley Kline gave a report on the activities at UCLA. Bob Braden
from the UCLA Computer Center is on a one year visit to UCL.
Charley is involved in a research project in the Computer Science
Department. This report is about the work Charley is involved in.
UCLA is building a system called LOCUS. It is based on Unix and a
ring network. The ringnet hardware is the LNI developed by UCI
and modified by MIT. A UCLA developed private protocol is used on
the ring. The current work is on the distributed operating
system. The goal is for the users of the system and for remote
hosts to see the collection of hosts and the ring net as one
system. The most recent effort is on file system aspects. Very
good results have been obtained yielding essentially equal access
times for local and remote files.

B. UCL

Peter Kirstein gave most of the presentation with some help from
Ron Jones. Peter noted that with Bob Braden visiting at UCL use
of the UCLA 3033 from London has increased and the service is

quite good. However, he has observed random outages of line 77
lasting 5 to 15 minutes several times per week. Bob has installed
echo servers both at the TCP and IP levels at UCLA for use with
UCL measurements and tests. Peter described UCL's work on: (1) a
protocol converter for service between SATNET and SRCNET for
terminal access, SRCNET is an X.25 network; (2) a cambridge ring
at UCL where terminal access is now working via UMCZ80 interfaces;
(3) a local message system is up on the UCL Unix using the MMDF
system developed by Dave Crocker; (4) the NIFTP will be upgraded
to match the latest document; and (5) UCL has worked over the TIU
TCP and made some parameter changes and installed a dynamic
retransmission timeout procedure, also MOS was revised, and a
document is available on this.

C. RSRE

John Laws gave the RSRE report.

Gateway Status: The line between the UCL gateway and the RSRE
gateway, known as net 35, was upgraded from 4.8Kb/s to 9.6Kb/s in
November 1980. The gateway machine is still an LSI-11. Local
performance measurements on the gateway were instrumented by
writing a process timer for the MOS system which showed how long
and how often individual processes ran for and the total time
spent idling in the scheduler. These measurements showed that the
gateway did handle 50Kb/s lines and would give the expected
throughput on these lines for large data packets (>128 user data
bytes). The switching capability was 50 packets per second using
the RSRE line cards. The time taken to switch packets from one
interface to another was 4.8msecs. This includes use of hashing
tables to determine the route of the packets. Although
fragmentation has been implemented and tested as reported at the
last meeting no use has yet been made of this facility. Source
routing code awaits a final imprimatur of the doctrine to be
employed.

Service Availability: A comparison between ARPANET availability
from the PPSN for one month periods before the last meeting and
this meeting show that although the outages total a similar total
time the ones for the Nov-Dec 80 period are mostly due to
developments in the UCL gateway rather than to SATNET as was the
case for the Aug-Sep 80 period. Two points are worth noting: (1)
Some of the longer outages were due to debuging sessions on the
UCL gateway. These are difficult because they require personnel
at RSRE, UCL and BBN to be available simultaneously. They often
require loading software at the RSRE gateway for the tests and
reloading software after the tests. (2) The large number of short
outages were traced to resetting of the Host-SIMP interface by the

UCL gateway caused by shortage of buffers. Ginny alleviated this to a large extent in January by removing code from the gateway. However if UCL and RSRE are heavily using the gateway we can still get up to 20 resets a day.

Local Network Status: The PPSN (net 25) now has three network access protocols available to users, these are datagram type (with multi-address), virtual call (PPSN type with multi-address), and X.25 virtual call. We have developed a flexible evaluation host both for use on the PPSN and the British Post Office PSS network which we hope to be connected to in February. This evaluation host supports 2 DTEs and includes a command console for setting up multiple calls from one terminal, traffic generators, and a measurement package. The protocols employed are the Network Independent Transport Service (NITS) [formerly "Yellow Book"], and X.25 level 3 with BPO options and the RSRE X.25 level 2 line units.

TCP and IP: The CORAL implementation of IP with fragmentation and reassembly is complete. The design for the CORAL TCP including full state machine representation is complete and coding is in progress. We have obviously gained a lot of useful experience for this from using the Mathis TCP. Besides TCP measurements we now have the capability to perform measurements of IP using echoers on ISIE and EDN Unix as well as GGP echo packets. The IP measurements package has the capability to make single round trip measurements at a user settable protocol type, packet length, and packet rate of 1 to 50 packets per second.

Action Items from RSRE:

1. Dynamic timeouts for retransmission.

2. Flag bit for copied options in IP fragments.

3. Specification for strict source routing.

4. Standard addresses for Echo, Discard, and Character Generator servers.

5. Techniques for preventing silly window syndrome (name due to Dave Clark).

6. No changes to addressing for a while.

Postel

D. SRI

Holly Nelson reported that the TIU with fragmentation and reassembly is in progress, that a new version of the port expander program was delivered, that a measurement host being installed is version 7 Unix with IP separate from TCP, and that Ron Kunzelman will be leaving SRI in mid February.

E. NDRE

Yngvar Lundh reported that progress at NDRE continues to be hampered by the lack of people. In spite of this, progress is being made and speech now runs in the local net. Paal Spilling has returned to NDRE after a one year visit to SRI. Paal will be trying speech experiments with Lincoln.

The agreement with the Norwegian Telecommunications Authority is being renegotiated for 2-3 years. NDRE has had discussions with the German group.

Action Items from NDRE:

1. A HDLC port on the SATNET gateway is needed.

2. ARPANET availability must be improved.

F. MITRE

Anita Skelton discussed the development by MITRE of a Z8000 based TCP. The program is written in C and cross compiled on Unix. The system is a modified C MOS. The TCP is about 600 lines of C code. This TCP has been tested talking to itself, but not against any other TCP. The data rate measured is 350 Kb/sec. Later in the meeting Steve Holmgren gave a presentation on this.

G. MIT

Dave Clark described developments at MIT. Planning for a computer network architecture at MIT is being formulated in terms of a system involving 52 buildings and 10,000 hosts. As Dave noted, he reported again that the version 2 ring net (10 Mb/s) is almost working. The development is a joint project with PROTEON. Some interfaces are prototyped in multiwire. The use of IP and TFTP based services is spreading. A Unix on the CHAOS net uses IP to communicate with a Unix on the Ring net. Fragmentation and Reassembly is used widely in the MIT environment. MIT did its own reworking of MOS -- this time to make it easy to bind with programs from other source languages.

Postel   [Page 4]

On the multics front Dave reports that IP/TCP was used over a HASP connection on a 9.6 Kb/s line between Cambridge and Phoenix with a round trip time of 8 seconds. Various investigations suggest hosts should be careful about also being gateways, for example, the gateway echo traffic could be more than the host could handle. In one experiment a routing loop was detected and datagrams would have looped forever except for Multics decrementing the time to live. Multics is being connected to TYMNET and TELENET, initially as a server for remote terminal access. A NIFTP server is being developed, and a MTP server is up on both NCP and TCP for RFC 772 style mail.

The TOPS20 system, MIT-XX, still has trouble with the IP implementation crashing the system, and the performance is poor otherwise.

Dave put together a minimal IP/TCP/Telnet in BCPL for an Alto. The total package is about 3000 words of code.

The Swallow Distributed Computing Project is basing its communication primitives on IP/UDP. A C-30 IMP is scheduled for delivery to MIT in early February.

Action Items from MIT:

1. A maximum segment size TCP option is needed.

H. Lincoln Laboratory

Jim Forgie described the work Lincoln has done on the Voice Terminal (VT) and the ST Gateway. Several VTs now exist and are functioning on the LEXNET. Two modes of speech data are supported 64Kb/s PCM and 2.4Kb/s LPC. A ST gateway is partially implemented in an 11/45 running EPOS and experiments are being done via the ARPANET with ISI. The gateway currently supports only the point-to-point ports of the ST protocol (no conference support yet). The measured gateway throughput is about 100 packets/second. The next few milestones on the Lincoln schedule are:

1. Move the gateway from the 11/45 to the 11/44.

2. Demo speech over the WBCNET.

3. Deliver a LEXNET and VTs to ISI.

4. Add source routing to ST and the gateway.

5. Add conferencing features.

I. Linkabit

Estil Hoversten discussed Linkabit's role in the support for
SATNET and WBCNET. Linkabit provides clever modems to make the
most of the channel. Estil pointed out that when Germany and
Italy join the SATNET the channel will be more shared and less
capacity will be available to existing communication partners.
For the WBCNET Estil reviewed the state of installation. All the
equipment is in place at ISI and LL, but the RF and power
equipment needs tuning up. The DCEC and SRI equipment should be
fully in place before the next IN MTG. Something called "Remote
Control and Alarming" is needed to make the operation legal. The
FCC requires that an operator (FCC licensed) be able to control
the transmitter (shut it off) if something goes wrong. Western
Union is installing the remote control mechanism so their operator
in New Jersey can satisfy this requirement.

Estil also mentioned that Linkabit has some communication between
two sites in San Diego several miles apart to support terminal
access for terminals at one site to TOPS20s at the other site.
They use dumb multiplexors and a 45 Mb/s microwave link. Also
Linkabit is part of M/A-COM which is planning to link its four
major sites via a satellite channel in the 800 Kb/s to Mb/s range.

J. ISI

Jon Postel reported on the activities at ISI. The Internet
Project has three areas of interest: Protocol Verification,
Protocol Design, and Protocol Applications. In the verification
area Carl Sunshine made a presentation later in the meeting. In
the design area Danny Cohen has led our efforts primarily in
discussion of addressing issues (see IEN 165). In the application
area we are developing two computer mail applications: the MPM
and the MTP. The MPM is a multimedia mail delivery system
following the specification of RFC 759. This is now working on
TOPS20, and was demonstrated at the end of the meeting. The MTP
is a text only mail server for multinetwork operation and follows
the specification of RFC 772. This is partially implemented on
TOPS20. It can now accept mail via TCP connections and leave it
for distribution via mailer and NCP connections. MTP was
demonstrated at the end of the meeting. Another program was
developed to better understand using TCP and can be used as a user
Telnet. This program is called TT. A description was distributed
at the meeting and TT was demonstrated at the end of the meeting.

Other work at ISI included the development and installation of

code in the TOPS20s to guard against sending the 9th outstanding
message to the same destination in the ARPANET and thus be blocked
by the IMP. This has been installed in the five TOPS20s at ISI.
The PDP-11 which is the front end for the Penguin printer has been
modified to also route internet datagrams to other hosts on the
ARPANET or the local Ethernet. The WBC project has participated
with Lincoln in the Speech experiments reported on by Lincoln.
Also the WBC project has interfaced a radio clock to the speech
host so that the operating system (EPOS) gets the time from the
radio clock.

K. DoD

Ray McFarland mentioned that questions are being raised about the
Internet Addressing and its implications for Network
Architectures, and on the relationship between Internet
Architectures and Distributed Processing Systems. Later in the
meeting Ken Shotting discussed his work on specification of IP.

L. DCEC

Ed Cain gave the report on DCEC activities. The hosts in the EDN
now do reassembly. There are two gateway programs. The old
gateway does fragmentation, talks GGP, and sends in CMCC reports.
The new gateway performs well. There was some confusion about GGP
routing messages.

It seems that a type 11 message was used in the BBN-built gateways
to avoid confusing a new routing message format with the one
documented in IEN 109 (How to Build a Gateway). The DCEC gateway
used the type 11 message in order to exchange routing information
with its neighbors. However, since the agreed-upon resolution is
to modify IEN 109 to show the new format, the BBN-built gateways
have been gradually reverting back to the new-format type 1
message. During the transition, the DCEC gateway has kept track
of which gateways use the two routing types, and serves as a
routing "bridge" between the two communities. The right type code
for routing messages is Type 1, and no Type 11 messages should be
used anymore.

DCEC also implements the security and precedence features of IP.
A TIU for AUTODIN II is being developed by DCEC, and an
implementation of MTP is in progress. DCEC coordinated a seminar
on IP and TCP at NBS in November which 405 people attended.

Action Items from DCEC:

1. How do we provide testing facilities to companies that develop TCP products?

2. How do we control transit traffic?

M.   COMSAT

Dave Mills gave a brief review of the configuration at COMSAT. The most important new development is the idea of virtual hosts which are processes with internet address and move from one real host to another within the DCNET. Another idea being developed is the maintenance of synchronized clocks in the hosts of the DCNET. This is done with time stamps in routing update messages between the hosts. Dave reported some interesting results about clocks and time variations that came to light in this work (see IEN 173). Also at COMSAT a MTP mail server is working. The FAX program now handles multipage documents. NIFTP works between COMSAT and ISIE -- 5 million bytes of RFCs and IENs were transferred. The INTELPOST project conversion to IP/TCP continues. A copy of the DCNET software has been installed at FORD research center at Dearborn with a 1200 baud dial up connection between Dearborn and Washington.

N.   BBN

Jack Haverty reported on the development of three new TCPs: for the TAC, the HP3000, and the VAX Unix. In the TAC, the TCP connection opening and closing has been completed, and work continues on the data handling code. In the HP3000 version the TCP is done and work is in progress on Telnet while waiting for a hardware interface. The VAX Unix (and C70) version is passing packets but work on TCP continues. The work on the VAN gateway (ARPANET/TELENET) continues; testing of X.25 level 3 is being done with an Applied Data Communications Pro/Tester. The TIP Login design is just about done; it is being designed as a special case of resource allocation. BBN is modifying CMOS for use in the C50 environment. The CMCC data has shown some anomalous behavior. For example short periods (5-10 minutes) with a high percentage of dropped datagrams and at the same time NCC data showing very busy IMP's. Hard to track this down -- better fault isolation tools are needed.

Ginny Strazisar reported on the developments in the BBN-built gateways. A new version of the gateway code was distributed to all sites (except Ft. Bragg) which will do fragmentation (but not with options). The gateway at UCL was modified to only use small

(SATNET size) buffers and remove local operator support to improve performance. The fragmentation routine will send on fragments as large as the network can handle. Work is in progress to change to a routing procedure which detects partitioned networks.

Dale McNeill reported on SATNET. After several months of poor performance, on November 6 there was a significant decrease in packets received with checksum errors on the SATNET channel. Performance is consistent with a change in the BER from about $10^{**}-4$ to about $10^{**}-6$. No information on the potential cause has been given by INTELSAT. In any case, SATNET channel protocol stability and the ARPANET direct connection via SATNET performance (ARPANET line 77 to the London TIP) are much improved. The US-Norway ARPANET satellite circuit (ARPANET line 41 between the SDAC IMP and the NORSAR TIP), which was changed last spring from commercially-controlled to military-controlled, continues to provide poor service. This line is composed of a military satellite hop and a number of commercial circuits. Hence, NDRE gateway to UCL gateway traffic is still disallowed in the Tanum Satellite IMP, thereby breaking internetwork traffic between the two gateway sites. A bug in the Host-SATNET protocol was found, and a fix was installed in the Satellite IMPs. The same fix needs to be inserted in the gateway to provide complete protection in the protocol. A major milestone was passed in MATNET, the secure shipboard copy of SATNET being built for the Department of the Navy; namely, a Satellite IMP successfully transmitted packets through a Navy FLTSAT satellite for the first time. The test setup was at the plant facility of E-Systems, ECI Division, where preliminary integration is being done between the network controllers and the radio equipment.

Action Items from BBN:

   1. Rubber EOL and Buffer Size has implementation impact in TAC.

O. ARPA

Vint Cerf reported that the ARPANET switched over to 96-bit headers on January 1st. Bill Carlson is leaving ARPA to join Western Digital and plans to build ADA machines. Vint mentioned a project at CCNY which is encoding video using adaptive delta modulation. ARPA plans to make a pair of these devices available to ISI and DCEC in July 1981. ARPA plans to replace all the 316 TIPS by TACs plus C30 IMPs. TIP Login will be used for access control. Germany and Italy are likely to join SATNET soon, Canada is in the discussion stage.

P. CSNET

Dave Farber gave a brief overview of CSNET. The idea is to
provide network services to a group of university computer science
departments. It is a 5 year NSF grant. The contractors are
University of Wisconsin, University of Utah, Purdue University,
The RAND Corporation, and 9 others. There is a steering committee
(Chair Bill Kerns, NSF) and a technical committee (Chair Dave
Farber, UDEL). CSNET will use Telenet, ARPANET, and telephones to
connect the universities. The services to be provided are:
Computer Mail, File Transfer, Terminal Access, and Distributed
Processing Applications. More information is available from Dave
(FARBER@UDEL).

Q. DTI

Gary Grossman gave an overview of DTI's work on a front end for
the WWMCCS H6000 hosts. The INFE (interim network front end) has
been up for some time, and has had some testing. It implements
the January 80 IP and TCP and has been tested with the EDN
systems. DTI is now working on the COS/NFE which is to be
multi-level secure and will be based on DTI's HUB (TM) operating
system. The COS/NFE is to be verified by SDC.

R. NAVELEX

Frank Deckelman described the interests and activities of NAVELEX
Code 330. There are four areas: DBMS, Distributed Processing,
Decision Aids, and Networking. In Networking the focus is on:
Global networks (e.g., MATNET), local networks (e.g., CCN - uses
Ungerman-Bass hardware), and C2 work stations (multimedia and AI).
One current program that involves all four areas of responsibility
is the ACCAT at NOSC and the Remote Site Modules at NPS, FNOC,
CINCPACFLT, NRL, and other sites.

S. XEROX

John Shoch gave a brief status report on networking at XEROX.
There are between 30 and 40 networks in operation, connected with
20 to 30 gateways. The new 10 Mb Ethernet is up. The XEROX
series 8000 products are based on the Ethernet. There is a Print
Server, a File Server, and a Communication Server (gateway), as
well as work stations. John also mentioned that a paper on
"Measured Performance of the Ethernet Local Network" appears in
the December 1980 issue of the Communications of the ACM.

IV. ACTION ITEMS

A. The problem reinitializing UCL gateway was resolved.

B. The design notes on the VAX, HP3000, and TAC TCPs were done [see IENs 166,167,168].

C. The memo on IP Errors and GGP style messages was distributed in DRAFT form.

D. The demonstration of the SRI Name Server was postponed indefinitely.

E. The MTP was demonstrated at the meeting.

V . PROTOCOL SPECIFICATION AND VERIFICATION

A. Overview

Carl Sunshine gave a well-received presentation on formal models for protocol analysis. Carl presented a basic model of a protocol and identified the aspects that need to be specified and verified. Then he described the methods for specifying protocols and identified particular programs or techniques that implement each method. Finally, Carl reviewed the verification methods and again identified particular systems which implement each method.

The specification methods are:

    Programs
        inductive assertions
    State Machines
        FSA
        Abstract Machines
        Petri Nets
        Sequence Expressions

The verification methods are:

    Program verification
    State Exploration
    Structural Induction
    Symbolic Execution
    Design Rules

A DRAFT version of a report on "Formal Modeling of Communication Protocols" was distributed. The final version is ISI RR-81-89.

B.   Three Way Handshake

Danny Schwabe presented an analysis of the three way handshake
made in the SPEX language. A description of a protocol in SPEX is
a "spexification." Spexifications can be translated into an
AFFIRM abstract data type. Proofs of properties of this abstract
data type can be done using AFFIRM. The analysis revealed a
possible problem in the three way handshake. As currently
described in the TCP specification (IEN 129), a connection may
become established on one side and potentially may accept data
from a previous incarnation of the connection. This is a very low
probability case and would be quickly detected, but the analysis
showed it was possible in principle.

The problem is that the document is in error in allowing an ACK to
be accepted before a SYN has been received.

A DRAFT report "Formal Specification and Verification of a
Connection Establishment Protocol was distributed.

C.   IP Specification

Ken Shotting described his specification of IP using SRI's HDM and
SPECIAL. Ken broke down IP into several mini-layers to make use
of the methodology. There was some difficulty making assertions
about global behavior since several fields are changed between the
source and destination, e.g., Time-to-Live. This also causes
problems with the checksum field. The order of processing assumed
for various features has an impact on the formal specification.
Ken's evaluation of HDM for protocol analysis is that it provides
good support for developing a hierarchy of abstract functions and
for state machine transitions, but is weak in data representation,
exception handling, and concurrency. Ken's work is documented in
a University of Maryland Technical Report "On the Formal
Specification of Computer Communication Protocols," Computer
Science TR-973, December 1980.

D.   NBS Work

Jack Haverty gave an overview of the work BBN is doing for the
NBS. The work is being done by Tom Blumer (TPB@BBN-Unix) and
Richard Tenney (RTenney@BBN-Unix). This protocol specification
system is based on a finite state machine model, augmented so that
there are variables, and arbitrary program segments may be
associated with each transition. These program segments are
written in a Pascal-like language. There is a syntax checker for
specifications, a specification editor, a FSM analyzer, a compiler
that produces C and a test generator.

For further information contact Tom Blumer or see BBN Report 4581.

E. DCA Work

Dave Kaufman reported on the work SDC is doing for DCA. The main
emphasis is on developing complete specifications of existing
protocols based upon the original design intent and upon a set of
specification guidelines being prepared in parallel by SDC. These
guidelines include the following sections: Introduction, Service
Specification, Lower Level Service Specification, Interface
Specifications, Peer Level Mechanism Specification, and Execution
Environment. These sections correspond with the basic protocol
model presented by Sunshine. A distinction is made between the
service specification which is the global view of the service and
the interface specification which is the user/service local
interface. The execution environment section describes protocol
requirements of the operating system (or more correctly of the
protocols runtime environment) such as timers and interprocess
communication. IP is currently being specified according to these
guidelines and during this effort several areas have been
uncovered which require further definition.

Example areas noted by SDC were:

1. Who supplies the "protocol" field for the IP header, IP or
the transport protocol?

2. What is the nature of the interaction between IP and GGP?

3. Is source routing loose or strict?

Jack Haverty raised an additional related issue:

4. How does IP interact with the local net, on errors, on flow
control,etc.?

F. NDRE Work

Yngvar Lundh noted that a graphical technique had been used to
produce an augmented state machine description of TCP for use in
an NDRE implementation effort.

VI. FRONT ENDS

A. Overview

Vint Cerf introduced the subject of Front Ends for protocols and
TCP in particular. The idea seems to be that in some cases IP/TCP
etc. may be too complex to put into the existing operating system,
so the protocol modules are put into a small front end machine.
Then there must be a communication procedure between the host and
the front end. This communication procedure is called Host-Front
End protocol (HFP). The argument is that HFP is much simpler for
the host than IP/TCP would be. Note that not everyone shares this
view.

B. WWMCCS HFP

Gary Grossman gave a detailed presentation on the HFP developed
for the WWMCCS H6000 machines. DTI has developed a protocol for
use between the host and the front end based on the notions of
"services" and "channels." The lowest level of the HFP is the
link level which provides a single channel with flow and error
control. The function of the link level corresponds to the
functions of an HDLC connection. The middle level is the channel
level. This is the level where separate logical streams appear
and are multiplexed based on logical channel number. At the top
level are services (i.e. applications). Typical services would be
Telnet and FTP. Gary gave quite a bit of detail on the protocol
formats and control procedures.

Gary also talked about the specification technique used to
describe the HFP. For each message type the following points are
described: function, when sent, sending state (and next state),
receiving state (and action to take), subsequent actions by sender
and receiver. For the service protocols, a specification includes
a "specification" and an "adaptation." The specification gives
the semantics of the fields used and the procedure for
communicating via the channels. The adaptation gives the format
details and any restrictions due to the particular machine
environment.

DTI and MITRE have done a fair amount of performance testing. The
comparison of an old implementation of NCP in the H6000 versus an
NCP implementation in the Front End show the Front End about 2 to
1 better in throughput. There are many differences in the two
configurations. Some testing recently showed that from the H6000
through the FE and over the ARPANET to DTI a data rate of 18Kb/s
was achieved. In local tests, 64Kb/s was measured over the
H6000-->FE-->I path; and with a source and sink in the FE, 84Kb/s

was measured for source-->I-->sink. Finally using a "mirror"
program in the FE instead of sending to the IMP, 125Kb/s was
measured for source-->mirror-->sink. An analysis of the CPU time
in the TCP code in the FE revealed that about 20% of the time was
spent on TCP functions and 80% on interprocess communication and
other "system" functions in the monitor.

Gary noted that details of the HFP are described in DTI report
DTI-80043.C-INFE.18, and DTI-78012.C-INFE.14, and the measurements
are described in the papers "Control Structure Overhead in TCP,"
NBS Computer Networking Symposium, May 1980, and "A Performance
Study of a Network Front End," Sixth Data Communication Symposium,
November 1979.

C. Z8000

Steve Holmgren of MITRE described a Z8000 based TCP. The goal is
to provide network communication for very small hosts which might
otherwise be peripheral devices of a computer. The Z8000 actually
supports a Network Access Protocol (NAP). The NAP is a layered
protocol with an option approach. The layers are: link,
transport, service, and user. The option approach allows the
users of the protocol to select the features they need and to
ignore others.

Steve gave some details of the protocol functions and formats.
The access to the protocol is via a system call style mechanism,
and the user is allowed to pass parameters and select options.

D. Discussion

John Shoch said he questioned the desirability of Front Ends, and
wondered, if aside from communication efficiency and pragmatics,
there was a good technical principle for front ends?

It was noted that we often excuse the current work of computing
checksums by saying we will someday have a checksum instruction in
our machines. Steve Holmgren asked "Why not a TCP instruction?"

When does a "front end" become an "instruction?"

VII. PERFORMANCE

A. Multics

Dave Clark discussed his measurements of TCP performance in
Multics. Dave described some detailed timing studies made of the
different TCP functions. The IP implementation is 5.3 K
instructions, and the TCP implementation is 9.0 K instructions.

Dave also discussed throughput in TCP when there is a lot of data
to send (e.g., in file transfers). A problem, previously pointed
out by Dave Mills, can arise where many small segments are sent.
This is due to the receiver reporting additional window each time
a small amount of the data has been processed, and the sender
immediately sending a segment to fit that additional window. Dave
Clark proposes to call this phenomena "Silly Window Syndrome"
(SWS). One way to prevent SWS is for the receiver not to report
new window unless the increase is a reasonable size. The receiver
can acknowledge incoming segments at any time, but limit window
updates to points when a reasonable increase can be made. It is
up to the receiver to decide what is reasonable, perhaps something
like 20% of the total buffer space for this connection. The
sender can also try to stay out of SWS by only sending big
segments and waiting until the window is large enough to allow it.
If the sending user indicates an end of letter then a smaller
segment may be sent.

Dave suggests that sending a segment with one octet of new data
into a zero window as a probe may lead to SWS. It may be better
to send an octet of old data.

Dave suggested that a performance measure might be the size of
bursts of segments the net, gateway, or host could handle. For
example, can a gateway handle a burst of 10 datagrams of 576
octets at the network input speed? Could measures be devised and
feedback control be exercised in terms of bursts? Wild idea
number one: the receiver can tell the sender the maximum burst
size it should send. Wild idea number two: have gateways turn on
a bit in the IP header if they are busy, and have the receiving
hosts integrate this information for use in determining a burst
size to feedback to the source.

B. UCL

Ron Jones reported on measurements of datagram traffic between UCL
and ISIE. The source of the traffic was an LSI-11 host. This was
connected to a port expander. The PE was in turn connected to the
UCL Gateway. The UCL Gateway communicates via the Goonhilly SIMP,

the SATNET, and the ETAM SIMP, to the BBN Gateway. The BBN
Gateway sends the datagrams through the ARPANET to ISIE.

Ron had a number of measurements which he described, but interest
focused on the discovery that in some tests from 25% to 90% of the
datagrams were lost between the ETAM SIMP and the BBN Gateway.
Some of the tests showed a significant step in the throughput
graph at the single-packet/multi-packet threshold.

C.   BBN

BBN believes that the problem originated because the ARPANET
IMP 40 would not accept packets from the BBN gateway fast enough.
The packet loss is manifest as multiple packet retransmissions
from the Etam SIMP to the BBN gateway, until the packet eventually
times out and is discarded in the SIMP. This behavior underscores
two inadequacies in the system. First, the BBN gateway should
discard the packet and send the appropriate message to the Etam
SIMP; the machinery is already present in Host-SATNET protocol to
accommodate this. Second, the Etam SIMP should notify the
Goonhilly SIMP which should notify the UCL gateway that problems
are arising. Unfortunately, a flow control mechanism like source
quenching was never developed within SATNET, although its need has
been known for some time; hence, when packets are dropped in
SATNET, gateways are never informed and therefore cannot generate
internet source quenching messages.

D.   RSRE

John Laws presented some findings of measurements performed by
Andrew Bates. These are datagram measurements of a single round
trip (at earlier meetings RSRE has reported on TCP measurements of
double round trip times). The source/sink is a TIU connected to
the RSRE gateway. One set of measurements seems to show that the
round trip time to the ETAM SIMP is about 2 seconds with very
little spread, but the round trip time to the BBN Gateway is also
about 2 seconds but with about 25% of the datagrams delayed (a
bimodal Histogram). The measurements were made at different times
of day, and may be due to a difference in the load on SATNET.

E.   CMCC

David Flood Page told of tracking down the cause of some looping
datagrams. In December the CMCC data showed that in the course of
a day one of the PRNET/ARPANET gateways had processed several
million datagrams. When the people who normally use that gateway
were asked what experiment they were conducting, they replied
"What experiment? We aren't doing anything." Several instances

of this incredibly high traffic level in at least two different
PRNET/ARPANET gateways were detected over the next several weeks.
David began investigating and eventually found the cause.

The loader server that loads the port expanders attached to the
PRNET gateways uses XNET 2.5, with Internet 2.5 headers, to do the
loading. The loader server sends the packets to logical host 15
(decimal) on the port expanders ARPANET address, which is the
address recognized by the bootstrap (obviously). The last packet
sent by the loader server is the one that says "go" to the
bootstrap; this causes the bootstrap to send out an acknowledgment
and then start up the port expander software. One of the first
things the PE does is to flip the ready line to the IMP. Now if
the acknowledgment from the bootstrap is still in the IMP, and the
IMP sees the ready line down, it will drop the ack (and any other
outstanding packets from that host). So the loader server gets no
ack, and retransmits the "go" command.

The "go" command arrives at the port expander addressed as before,
i.e. to logical host 15 on the PE; but the PE doesn't have a
logical host 15. It is set to hand off any packets which it
doesn't know how to route to its logical host 0, which is the
gateway; the gateway sees that it is a packet destined for
somewhere on the ARPANET that is not itself, so sends it back out
its ARPANET interface, i.e. to the PE. The PE has the same
trouble as before, so we get a loop. Because the packet is an
Internet 2.5 packet, it does not have a time to live field, so the
packet loops indefinitely until either the PE goes down, the
gateway goes down, or a flood of real packets causes enough
congestion to result in the dropping of the looping packet.

This didn't happen every time the PE was loaded, because some of
the time, the acknowledgment got out of the IMP fast enough not to
be dropped when the ready line was flipped. It was largely a
matter of luck.

The problem will disappear either when the new gateway version,
which drops all Internet 2.5 packets, goes into the PRNet
gateways; or when the V4 bootstrap goes into the PEs. In fact I
think the new gateway version may already be in place. Holly did
put a fix in the PE as well, but I'm not sure what it was. One
important point to note is that the port expander is on the
ARPANET side of the gateway, not the PRNet side.

David says the lesson that should be learned from this is that
more remote access to debugging tools is needed. The key piece of
evidence in solving this problem was provided by a "packet

printer" trace program that could only be run locally at the
gateway/port expander site - which is normally unmanned.

F. Radio Clocks

Steve Casner described (actually showed) one of the Radio Clocks.
Since we are about to start doing coordinated measurements of one
way times we must have an agreement on what size and precision of
time stamps to carry around in datagrams and programs. The "no
thought" proposal is 32 bits of milliseconds since 1 January 1980.
Others suggested that we might want more precision later so maybe
we should have micro seconds and more bits. This was not
resolved.

VIII. WORM PROGRAMS

John Shoch briefly described an experiment in distributed computing.
Programs were constructed to operate in segments with each segment
allocated to a personal computer. The total program is called a
worm. The basic version of the program simply maintains its
existence. If a segment dies the remaining segments cooperate to
find a free machine, load it with the segment code, and start it
running as part of the worm. John described some of the interesting
things which can go wrong, and ways to prevent them. The experiment
is discussed in more detail in IEN 159. One comment to the internet
group is that this experiment showed the usefulness of
multidestination or group addresses. The conclusion is that with
datagram style communication, high speed local networks, and new
control procedures, the tools are in hand to begin some very
interesting multi-machine applications.

IX. IP ERRORS

Jon Postel distributed a draft memo on error reports in IP. The memo
is titled "What every host should know about GGP". The intention is
to use the GGP messages on routing advice and destination dead as a
base and to add a few other messages to cover the host-to-host error
conditions. The discussion focused on whether or not this should
remain part of GGP or be in a separate protocol. From an
implementation point of view there seems to be little difference,
from an administrative point of view it seems to be cleaner for it to
be separate. There were some strong difference of opinions. The
matter was not resolved.

X.   ADDRESSING DISCUSSION

A.   Overview

Vint Cerf opened the topic with a call for a very general
discussion of the issues and spent some time developing a list of
goals for the the addressing and routing procedures. The goals
were:

1. Keep routing simple.

2. Keep routing tables small.

3. Keep computation costs small.

4. Most datagrams should be delivered.

5. Use any available connectivity.

6. Multidestination delivery.

7. Handle mobile hosts.

8. Efficiently use constituent networks.

9. Support multi-homing.

10. Dumb hosts should be included.

11. Pantheism should be dealt with.

12. Degradation should be automatically fixable.

13. The system should scale up.

14. The system should be measurable.

15. Ability to use hierarchy.

There was some discussion of these points and it was noted that
some are service specifications while others are "good job"
criteria.

B.   Presentation

Noel Chiappa described the MIT complex which includes 9 networks
and three major protocol families (IP, PUP, and CHAOS). All these
networks share one "Internetworkly known" network number. Noel

listed a number of problems and made some observations. One
possible long term development may be that all hosts are on local
networks and only gateways are attached to long haul networks.
The thrust of Noel's remarks seems to be that things are going to
grow in a complex, but generally hierarchical, way; and any
workable system will have to be prepared to grow, probably by
taking advantage of the structure.

C. Discussion

Vint Cerf led a further discussion on addressing. The main focus
was on the tradeoff between a flat address space and a
hierarchical one. In a flat address system there is no relation
between the address and the location, and routing has to be by
central knowledge or broadcast. In a hierarchical address system
the address is composed of fields which identify the location in a
routing tree.

Vint suggests that we have both in one! Let an address be
composed of two parts: a hierarchical address (called an address)
and a flat address (called an identifier). The address can be
used as a routing guide or hint, but the identifier must match for
a message to be delivered. The identifiers are globally unique
names for hosts and do not change when hosts move.

There are many questions to be answered in this scheme. For
example, How does source routing work? Or Multi-homing?

XI. FILE TRANSFER IMPLEMENTATION STATUS

Vint Cerf took a survey of the implementers present to find out the
status of FTP implementations. There are four different FTP systems
being used: ARPA FTP (RFC 765), AUTODIN II FTP (IEN 101), NIFTP (IEN
99), and TFTP (IEN 133). The first three use TCP, and the last
(TFTP) uses UDP.

ARPA FTP

    DCN (COMSAT)   now
    TOPS20 (BBN)   April 81
    EDN-Unix (DCEC)   June 81
    Multics (MIT)   after June 81

AUTODIN II FTP

    BBN-Unix (BBN)   now
    EDN-Unix (BBN)   now

Internet Meeting Notes

NIFTP

TOPS20 (UCL)	now
DCN (COMSAT)	now
Multics (MIT)	June 81
UCL-Unix (UCL)	now

TFTP

Multics (MIT)	now
TOPS20 (MIT)	now
ALTO (MIT)	now
Unix (MIT)	now
TOPS20 (ISI)	March 81
EPOS (ISI)	April 81

XII. COMPUTER MAIL

A. Mail Facilities

Peter Kirstein discussed the problems of providing mail services between users in the UK and the US. Within the UK, computer mail will almost certainly be based on NIFTP and will likely be as proposed in IEN 169. One major concern for the US/UK transmission is cost. It is quite important that a minimum number of transmissions be made and especially that only one copy of a message destined for multiple users be sent across the Atlantic.

Peter raised the issue of an interface between NIFTP mail and MTP mail, since it seems that the Internet will use the MTP mail.

B. Mail Transfer Protocol

Jon Postel discussed the MTP mail plan and indicated that a revised document would be issued soon which will describe the MTP in a network independent style.

Jon agreed with Peter that a NIFTP-mail/MTP interface data structure should be defined soon, and promised to do so.

There are several working (or partly working) MTPs already: Multics, COMSAT, DCEC, and ISI.

C. TELEX/TWX Gateway

Geoff Goodfellow described a system that is under development at SRI to connect Telex and Computer mail. A user at a computer terminal runs a program like SNDMSG or MM and simply adds some

additional fields to the header of the message to give the
information needed for sending a telex. The message is routed to
a special host which checks the telex information and if it is
acceptable sends a telex. When a telex arrives at SRI it is
processed by the special host and if the required keywords and
information are found it is sent as a RFC 733 style message to the
recipients computer mail mail box. A small percentage of the
incoming telexes must be handled by a human operator because they
do not contain the required information in a computer parsable
form.

D. Discussion

There was some discussion of mailbox addresses and the problem of
sending (and answering) mail to (from) mailboxes in other systems
(e.g., Internet, TELEMAIL, ONTYME).

XIII. NEXT MEETING

The next meeting will be held 1-2-3 June 1981 at COMSAT in Washington
D.C. Dave Mills (Mills@ISIE) is the host. Please send agenda items
to Jon Postel (Postel@ISIF). If you plan to attend please register
with Linda (Linda@ISIF). Local arrangements information will be sent
only to those registering.

The fall meeting will be held at UCL on 21-22-23 September 1981. The
winter meeting will be held at SRI in mid January 1982.

DOCUMENTS DISTRIBUTED

Author	Subject	Number
Cerf	INFOCOM82 Call for Papers	---
Postel	AGENDA	---
Postel	Recent Document List	---
Postel	IEN INDEX	---
Postel	Assigned Numbers	776
Farber	CSNET Description	---
Postel	TT	---
Chase	TTLUSR	---
Postel	What Every Host Should Know About GGP	---
Cohen	On IP-Addressing	170
Cohen	Addressing in the ARPANET, another visit	171
Sunshine	Formal Modeling of Communication Protocols (DRAFT)	---
Schwabe	Formal Specification and Verification of a Connection Establishment Protocol (DRAFT)	---
Bennett	A Simple NIFTP-Based Mail System	169

RECENT DOCUMENTS

Author	Subject	Number

IENs

Author	Subject	Number
Shoch	Notes on the "Worm" Programs - Some Early Experience with a Distributed Computation	159
Postel	Internet Meeting Notes - 7-8-9 October 1980	160
Jones	A Proposal for Simple Measurement Support for Users Through Slow Nets	161
Pershing	Transport, Addressing, and Routing in the Wideband Net	162
Jones	Echo Delay Measurements with GGP Packets	163
Stern	CMOS System Overview	164
Cohen	About Addressing in the WBnet	165
Hinden	Design of TCP/IP for the TAC	166
Sax	HP3000 TCP Design Document	167
Gurwitz	VAX-UNIX Networking Support Project Implementation Description	168
Bennett	A Simple NIFTP-Based Mail System	169
Cohen	On IP-Addressing	170
Cohen	Addressing in the ARPAnet, Another Visit	171
Mills	Time Synchronization in DCNET Hosts	173

RFCs

Author	Subject	Number
Postel	Internet Protocol Handbook Table of Contents	774
Mankins	Directory Oriented FTP Commands	775
Postel	Assigned Numbers	776

ATTENDEES

Name	Affiliation	Nationality	Mailbox
Vint Cerf	ARPA	USA	Cerf@ISIA
Robert Kahn	ARPA	USA	Kahn@ISIA
David Flood Page	BBN	British	DFloodPage@BBNE
Jack Haverty	BBN	USA	JHaverty@BBND
Robert Hinden	BBN	USA	Hinden@BBNE
Charles Lynn	BBN	USA	CLYNN@BBNA
Dale McNeill	BBN	USA	DMcNeill@BBNE
Paul Santos	BBN	USA	Santos@BBNE
Virginia Strazisar	BBN	USA	Strazisar@BBNA
Gerry Amey	Canada-DND	Canadian	Amey@ISIA
Hoi Chong	COMSAT	Rep. China	Chong@ISIE
David Mills	COMSAT	USA	Mills@ISIE
Marvin Solomon	CSNET	USA	UWVAX!Solomon@BERKELEY
Ed Cain	DCEC	USA	Cain@EDN-UNIX
Wayne Shiveley	DCEC	USA	Shiveley@BBNB
Hans Dodel	DFVLR	Germany	---
Helmuth Lang	DFVLR	Germany	---
Gary Grossman	DTI	USA	grg@DTI
Horst Clausen	IABG	Austria	Clausen.IABG@Mit-Multics
Ray McFarland	DoD	USA	McFarland@ISIA
Ken Shotting	DoD	USA	Shotting@SRI-k1
W. Bradbury	I.P. Sharp	Canadian	---
Stephen Casner	ISI	USA	Casner@ISIB
Danny Cohen	ISI	USA	Cohen@ISIB
Randy Cole	ISI	USA	Cole@ISIB
Dan Lynch	ISI	USA	Lynch@ISIB
Bill Overman	ISI	USA	Overman@ISIF
Jon Postel	ISI	USA	Postel@ISIF
Danny Schwabe	ISI	Brazil	Schwabe@ISIF
Carl Sunshine	ISI	USA	Sunshine@ISIF
Estil Hoversten	Linkabit	USA	Hoversten@ISIA
Jim Forgie	LL	USA	Forgie@BBNC
Noel Chiappa	MIT	British	JNC@XX
Dave Clark	MIT	USA	Clark@MIT-Multics
Steve Holmgren	MITRE	USA	Steve@MITRE
Anita Skelton	MITRE	USA	Anita@MITRE
Frank Deckelman	NAVELEX	USA	Deckelman@ISIA
Oyvind Hvinden	NDRE	Norway	Oyvind@DARCOM-KA
Yngvar Lundh	NDRE	Norway	Yngvar@DARCOM-KA
Merle Neer	NOSC	USA	Neer@ISIA
Mike Wahrman	RAND	USA	Mike@RAND-UNIX
Derek Barnes	RSRE	British	T45(ATTN.Barnes)@ISIE
John Laws	RSRE	British	T45@ISIE
Dave Kaufman	SDC	USA	Kaufman@ISIE

Geoff Goodfellow	SRI	USA	Geoff@SRI-ka
Ron Kunzelman	SRI	USA	Kunzelman@SRI-KL
Jim Mathis	SRI	USA	Mathis@SRI-KL
Holly Nelson	SRI	USA	HNelson@SRI-KL
Zaw Sing Su	SRI	Canada	ZSu@SRI-KL
Ken Biba	SYTEK	USA	Biba@SRI-KL
Ron Jones	UCL	British	UKSAT@ISIE
Peter Kirstein	UCL	British	PKirstein@ISIA
Charles Kline	UCLA	USA	CSK@UCLA-Security
Dave Farber	UD/CSNET	USA	Farber@UDEL
John Shoch	Xerox	USA	Shoch@PARC