A Proposal for Simple Measurement Support for Users

IEN 161

R.G.Jones

November 1980

University College London

Contents

1. Introduction

The majority of this note was originally written to
address a wider problem than measurements solely within
the ARPA catenet and the orientation of the discussion
is therefore in those terms. However, it is felt that
the proposed technique would also be appropriate for
user measurements confined to the catenet.

The intention in distributing this note is to gather
initial reactions to the proposed technique and
therefore no attempt has been made to fill in all the
details.

2. Motivation

Increasingly users will wish to evaluate
communication performance across a concatenation of
varied networks. The purpose of this note is to propose
a minimal set of easily implemented facilities which
would support measurements of network performance at
the user level, in particular at the network and
transport levels. It is based on ideas originally
circulated in an INDRA Group Working Paper 1 .

Given the rapid proliferation of networks and the
corresponding increase in the variety of their type and
interconnection, it is felt that the facilities should
not be tightly bound to a particular protocol but
should be supportable across as many networks as
possible. In particular, for the purposes of examples,
the environment in which the measurements would be
performed is assumed to extend beyond the ARPA catenet
system to include, for example, VANs or PTT X25-based
networks.

The note deliberately concentrates on a subset of the
requirements for a comprehensive measurement program in
the hope of ensuring rapid implementation of this
essential subset.

3. Measurements

The performance of a packet-switched network is
typically characterized by delay as a function of
throughput and the facilities proposed are intended to
support this type of measurement. In the case of
networks with a virtual call interface, the call set-up
time is obviously an additional important parameter,

but this can be measured without requiring co-operation
from a remote site.

Throughput and delay measurements are most easily
undertaken by employing timestamps. This technique
requires at least the provision of an echo server at a
remote site. Useful results, such as round-trip time,
can be obtained with even the simplest echo server. If
global clock sychronization can be achieved (for
example, within a satellite-based net or by the use of
broadcast time standards) then timestamping at
intermediate points is particularly valuable in
pinpointing the origins of delay and throughput
constraints. However, the insertion of timestamps by
remote sites requires an agreed format and location
within the packet for the timestamps and an agreed
triggering mechanism. In a network with adaptive
routing, it is important that the intermediate
timestamps should have some identification of their
origin associated with them.

The minimum requirements for a measurement program
are therefore seen to be a widely available set of echo
servers with triggerable timestamping facilities. A
more thorough program of work would require, for
example, participating sites to make available traffic
generators which were remotely controllable.

4. Current Measurement Support

Within the domain of nets supporting ARPA protocols,
timestamping is an option in the internet header.
Currently, however, there is no associated
identification of the origin of the timestamp.
Furthermore, the number of timestamps is limited by the
maximum size of the internet header and this is even
more restrictive if, for example, source routing is to
be employed. In measurements beyond the ARPA catenet
the internet header may well be lost in traversing
other networks where gateways may, for example, do
protocol conversion at the transport level.

Echo servers are available at gateways for packets
utilizing the Gateway-Gateway Protocol but since such
packets are treated differently from normal traffic
these are useless for throughput tests and provide
optimistic delay measurements.

R.G.Jones

Other available facilities suffer from comparable
restrictions. Current support is therefore felt to be
inadequate not only in a wider context, but even within
the ARPA catenet system.

The situation with regard to other nets, for example
PTT X25-based nets, is even worse and it is unrealistic
to suppose that PTTs will be forward in providing
facilities. Users will therefore be forced to be co-
operative if they wish to pursue a measurement program
of reasonable extent.

5. Proposed Support

The basic requirement is for a mechanism which will
allow measurements to be made across very different
nets. With regard to interworking between nets based
on different protocols, it is unreasonable to expect
groups to implement protocols specific to another
network. Central to the proposed mechanism is a
standard format for the measurement data area which is
independent of the underlying network protocol. Such a
data area can be employed as the user data of the
selected protocol. Measurement data is distinguished at
the underlying protocol level from normal traffic by
the use of a reserved protocol number or port or
whatever is appropriate to the protocol.

The proposed format is outlined in Figure 1.

The type field specifies, for example, whether the
data is to be echoed, is an echo, or whether it should
simply be dropped by the receiving process.
The flag field can be used to specify whether the
station address should be inserted by each timestamping
station and to specify the categories of sites which
should timestamp. For example, for measurements made at
the ARPANET internet datagram level, the flags could
specify that intermediate gateways should timestamp as
well as the destination.
The length field is used at the transport level as a
record delimiter.
The sequence number field is simply a field in which
the originating process can insert a data area sequence
number if so desired.
The offset field points to the location at which the
next timestamp (and possibly station identification)
should be inserted. This is updated by each
timestamping station to point beyond the timestamp it
has inserted.

```
+----------------+----------------+
|      type              flags    |
+----------------+----------------+
|             length              |
+----------------+----------------+
|         sequence number         |
+----------------+----------------+
|             offset              |
+----------------+----------------+
|        origin of ts#1           |
|          (optional)             |
+----------------+----------------+
|          timestamp              |
|          number 1               |
+----------------+----------------+
|                                 |
|          timestamp              |
|          number n               |
+----------------+----------------+
```

FIGURE 1: Format of Measurement Data Area

A sequence of timestamps (possibly with associated
station identification) then follows the offset field.
For experiments with user data of varying sizes, the
measurement data area would be followed by padding to
the appropriate length.
For throughput tests with minimal size user data it
would, of course, be possible to dispense with all
fields except for length and type and flags, with the
latter set to indicate no timestamping.

It is intended that the traffic generator would
supply a data area large enough to contain the
anticipated number of timestamps. If it were possible
for there to be a large variation in the number of
networks traversed for a given destination that might
prove difficult to initially estimate. However, a
station unable to timestamp because of lack of space
would set a flag to indicate the fact.

Station identification in measurements across nets
with an homogeneous addressing scheme (such as the ARPA
catenet) presents no problems. However, identification
in other systems rules out a standard field size. The
PTT networks, for example, specify an international

address of up to 14 digits and local nets will have a
variety of further schemes. The use of timestamp
identification in the most general case is therefore a
matter for further consideration. In many situations
the problem will not arise since, for example, across
PTT nets timestamps will be available only from the
destination echo server. It may be desirable to add an
additional field either for each origin or each
origin/timestamp pair to indicate length and format.

The suggested format allows servers to be exceedingly
simple with much common code between servers at
different protocol levels.

In the ARPANET environment, echo servers for such
data should be provided above the internet datagram
level and above TCP. Thus, an internet protocol number
would be reserved for measurement packets and all such
packets processed by an internet layer in a host or
gateway would be passed to an internet datagram echo
server. It should be possible to specify timestamping
in a gateway functioning as a transit gateway rather
than the packet destination. A "well-known" TCP port
would be reserved for a server for TCP tests.

For X25-based networks, at the virtual circuit level,
a protocol number could be reserved for the "protocol
field" of the call user data area (bytes 1 through 4).
In the absence of such a universally recognized
identification, substitutes can be found. For example,
on PSS "well-known" (to co-operating sites) sub-
address digits could specify the echo server in the
DTE.

6. Summary

Attention has been drawn to the inadequacies of
current measurement support for the network user. An
outline has been given of the minimum requirements for
a user measurement program and a method of meeting the
requirements has been proposed. The central mechanism
of the proposed support is a standardized format for
measurement data, which lends itself to easy
implementation above various network protocols
accessible by the user. This scheme means that the
measurement data is independent of the protocol level
in use and may be more easily preserved across
networks.

   The note has deliberately defined a minimal subset of
the requirements for a comprehensive measurement
program, not least because it is felt that these
facilities could be quickly provided.

   It does not attempt to provide the necessary
facilities for detailed investigation of network
behaviour, but does (quickly and easily) provide
sufficient resources for the user to be able to
identify problems for further investigation.

R.G.Jones