                        TFTP Blocksize Option

Status of this Memo

   This document specifies an Internet standards track protocol for the
   Internet community, and requests discussion and suggestions for
   improvements. Please refer to the current edition of the "Internet
   Official Protocol Standards" (STD 1) for the standardization state
   and status of this protocol. Distribution of this memo is unlimited.

Abstract

   The Trivial File Transfer Protocol [1] is a simple, lock-step, file
   transfer protocol which allows a client to get or put a file onto a
   remote host. One of its primary uses is the booting of diskless
   nodes on a Local Area Network. TFTP is used because it is very
   simple to implement in a small node's limited ROM space. However,
   the choice of a 512-octet blocksize is not the most efficient for use
   on a LAN whose MTU may 1500 octets or greater.

   This document describes a TFTP option which allows the client and
   server to negotiate a blocksize more applicable to the network
   medium. The TFTP Option Extension mechanism is described in [2].

Blocksize Option Specification

   The TFTP Read Request or Write Request packet is modified to include
   the blocksize option as follows. Note that all fields except "opc"
   are NULL-terminated.

```
        +-------+---~~---+---+---~~---+---+---~~---+---+---~~---+---+
        |  opc  |filename| 0 |  mode  | 0 | blksize| 0 | #octets| 0 |
        +-------+---~~---+---+---~~---+---+---~~---+---+---~~---+---+
```

   opc
       The opcode field contains either a 1, for Read Requests, or 2,
       for Write Requests, as defined in [1].

   filename
      The name of the file to be read or written, as defined in [1].

   mode
      The mode of the file transfer: "netascii", "octet", or "mail",
      as defined in [1].

   blksize
      The Blocksize option, "blksize" (case in-sensitive).

   #octets
      The number of octets in a block, specified in ASCII. Valid
      values range between "8" and "65464" octets, inclusive. The
      blocksize refers to the number of data octets; it does not
      include the four octets of TFTP header.

For example:

```
+-------+--------+---+--------+---+--------+---+--------+---+
|   1   | foobar | 0 | octet  | 0 | blksize| 0 |  1428  | 0 |
+-------+--------+---+--------+---+--------+---+--------+---+
```
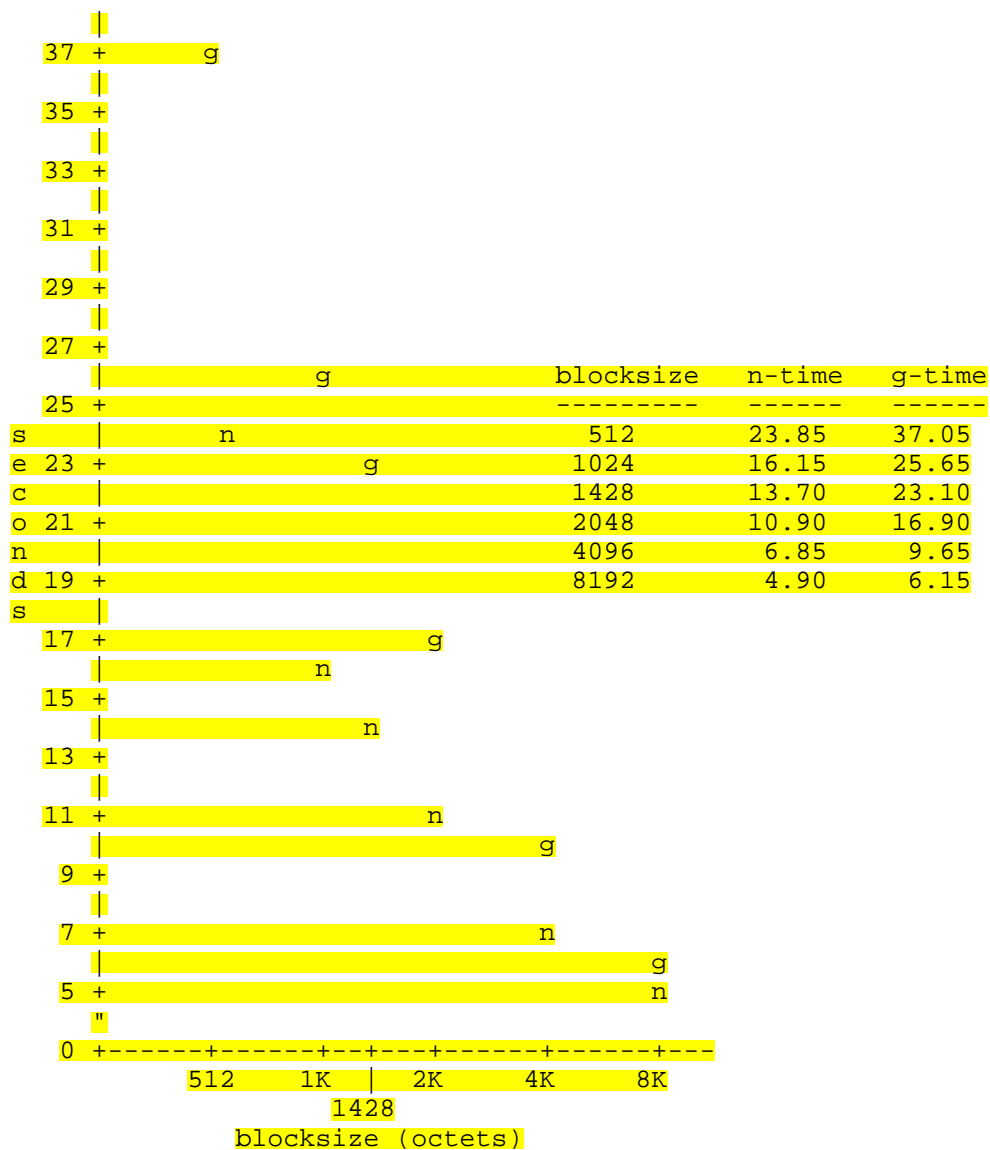
is a Read Request, for the file named "foobar", in octet (binary)
transfer mode, with a block size of 1428 octets (Ethernet MTU, less
the TFTP, UDP and IP header lengths).

If the server is willing to accept the blocksize option, it sends an
Option Acknowledgment (OACK) to the client. The specified value must
be less than or equal to the value specified by the client. The
client must then either use the size specified in the OACK, or send
an ERROR packet, with error code 8, to terminate the transfer.

The rules for determining the final packet are unchanged from [1].
The reception of a data packet with a data length less than the
negotiated blocksize is the final packet. If the blocksize is
greater than the amount of data to be transfered, the first packet is
the final packet. If the amount of data to be transfered is an
integral multiple of the blocksize, an extra data packet containing
no data is sent to end the transfer.

Proof of Concept

   Performance tests were run on the prototype implementation using a
   variety of block sizes. The tests were run on a lightly loaded
   Ethernet, between two HP-UX 9000, in "octet" mode, on 2.25MB files.
   The average (5x) transfer times for paths with (g-time) and without
   (n-time) a intermediate gateway are graphed as follows:

```
      ||
   37 +      g
      ||
   35 +
      ||
   33 +
      ||
   31 +
      ||
   29 +
      ||
   27 +
      |             g          blocksize  n-time  g-time
   25 +                        ---------  ------  ------
 s    |      n                    512      23.85   37.05
 e 23 +          g               1024      16.15   25.65
 c    |                          1428      13.70   23.10
 o 21 +                          2048      10.90   16.90
 n    |                          4096       6.85    9.65
 d 19 +                          8192       4.90    6.15
 s    |
   17 +             g
      |            n
   15 +
      |           n
   13 +
      ||
   11 +               n
      |                g
    9 +
      ||
    7 +                   n
      |                     g
    5 +                     n
      "
    0 +------+------+--+---+------+------+---
           512    1K  |  2K     4K     8K
                    1428
                blocksize (octets)
```

The comparisons between transfer times (without a gateway) between
the standard 512-octet blocksize and the negotiated blocksizes are:

```
    1024     2x    -32%
    1428    2.8x   -42%
    2048     4x    -54%
    4096     8x    -71%
    8192    16x    -80%
```

As was anticipated, the transfer time decreases with an increase in
blocksize. The reason for the reduction in time is the reduction in
the number of packets sent. For example, by increasing the blocksize
from 512 octets to 1024 octets, not only are the number of data
packets halved, but the number of acknowledgement packets is also
halved (along with the number of times the data transmitter must wait
for an ACK). A secondary effect is the efficiency gained by reducing
the per-packet framing and processing overhead.

Of course, if the blocksize exceeds the path MTU, IP fragmentation
and reassembly will begin to add more overhead. This will be more
noticable the greater the number of gateways in the path.

Security Considerations

   The basic TFTP protocol has no security mechanism. This is why it
   has no rename, delete, or file overwrite capabilities. This document
   does not add any security to TFTP; however, the specified extensions
   do not add any additional security risks.

References

   [1] Sollins, K., "The TFTP Protocol (Revision 2)", STD 33, RFC 1350,
       October 1992.

   [2] Malkin, G., and A. Harkin, "TFTP Option Extension", RFC 2347,
       May 1998.

Authors' Addresses

   Gary Scott Malkin
   Bay Networks
   8 Federal Street
   Billerica, MA 10821

   Phone: (978) 916-4237
   EMail: gmalkin@baynetworks.com

   Art Harkin
   Networked Computing Division
   Hewlett-Packard Company
   19420 Homestead Road MS 43LN
   Cupertino, CA 95014

   Phone: (408) 447-3755
   EMail: ash@cup.hp.com

Full Copyright Statement

   Copyright (C) The Internet Society (1998). All Rights Reserved.

   This document and translations of it may be copied and furnished to
   others, and derivative works that comment on or otherwise explain it
   or assist in its implementation may be prepared, copied, published
   and distributed, in whole or in part, without restriction of any
   kind, provided that the above copyright notice and this paragraph are
   included on all such copies and derivative works. However, this
   document itself may not be modified in any way, such as by removing
   the copyright notice or references to the Internet Society or other
   Internet organizations, except as needed for the purpose of
   developing Internet standards in which case the procedures for
   copyrights defined in the Internet Standards process must be
   followed, or as required to translate it into languages other than
   English.

   The limited permissions granted above are perpetual and will not be
   revoked by the Internet Society or its successors or assigns.

   This document and the information contained herein is provided on an
   "AS IS" basis and THE INTERNET SOCIETY AND THE INTERNET ENGINEERING
   TASK FORCE DISCLAIMS ALL WARRANTIES, EXPRESS OR IMPLIED, INCLUDING
   BUT NOT LIMITED TO ANY WARRANTY THAT THE USE OF THE INFORMATION
   HEREIN WILL NOT INFRINGE ANY RIGHTS OR ANY IMPLIED WARRANTIES OF
   MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.