HTTP Datagrams, UDP Proxying, and Extensible Prioritization

draft-pardue-masque-dgram-priority

IETF 115 – London – 2022-11

Lucas Pardue – lucaspardue.24.7@gmail.com
Background

RFC 9000 - stream multiplexing can have a significant effect on application performance

    QUIC does not provide a mechanism for exchanging prioritization information.

RFC 9114 - HTTP/3 punts on stream prioritization.

RFC 9218 - Extensible Prioritization Scheme for HTTP/2 and HTTP/3

RFC 9221 - QUIC DATAGRAM frames. No transport multiplexing identifier.

RFC 9297 - HTTP DATAGRAMS and the Capsule Protocol

RFC 9298 - Proxying UDP in HTTP

MASQUE proxying and WebTransport definitely can exercise stream and datagram multiplexing
Venn and the art of protocol maintenance

HTTP WG
QUIC WG
MASQUE WG
WebTrans
MoQ
Extensible HTTP Priorities recap

HTTP Extensible Prioritization for streams defines signals:

- **urgency** (“u”) - between 0 and 7. Smaller the value, higher the precedence
- **incremental** (“i”) - response can be processed incrementally (data as it arrives)

And some scheduling guidance:

- Expressing priority is only a suggestion.
- RECOMMENDED to respect urgency, serve in stream ID order.
- RECOMMENDED to respect incremental, fair bandwidth sharing between incremental at same urgency
Extensible priorities stream scheduling

<table>
<thead>
<tr>
<th>u, ID</th>
<th>ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>0, 24</td>
<td></td>
</tr>
<tr>
<td>0, 28</td>
<td></td>
</tr>
<tr>
<td>0, 12</td>
<td></td>
</tr>
<tr>
<td>0, 16</td>
<td></td>
</tr>
<tr>
<td>1, 20</td>
<td></td>
</tr>
<tr>
<td>6, 0</td>
<td></td>
</tr>
<tr>
<td>0, i=1</td>
<td>12</td>
</tr>
<tr>
<td>0, i=1</td>
<td>16</td>
</tr>
</tbody>
</table>
HTTP datagram scheduling

FIFO??

| Qrtr stream ID = 0 | Qrtr stream ID = 1 | Qrtr stream ID = 2 | Qrtr stream ID = 0 |

Sticking stuff in the same bucket/queue is a bit basic
Stream and datagram scheduling?

0  u=0, ID = 24  u=0, ID = 28  u=0, i=?1, ID = 12  u=0, i=?1, ID = 16
1  u=1, ID = 20
2
3
4
5
6  u=6, ID = 0
7

Datagram:  Qrtr stream ID = 0  Qrtr stream ID = 1  Qrtr stream ID = 2  Qrtr stream ID = 0
Capsule and datagram scheduling?

Same stream ID - capsules likely to be more important than datagrams?
Bouncing around

- **HTTP Datagram Issue #46** – The spec should discuss how h3-datagram works (or does not) with priority.
  - Closed with a PR that says:

    Prioritization of HTTP/3 datagrams is not defined in this document. Future extensions MAY define how to prioritize datagrams, and MAY define signaling to allow endpoints to communicate their prioritization preferences.

- **HTTP Priorities #1550** – How are DATAGRAM frames prioritized?
  - Closed with a PR that says:

    The priority scheme defined by this document considers only the prioritization of HTTP messages and tunnels ... Where HTTP extensions change stream behavior or define new data carriage mechanisms, they MAY also define how this priority scheme can be applied.
Extend extensible priorities with a compatible parameter: `datagram-urgency` ("du").

Identical to `urgency`, except that it applies to datagrams.

Omission of `datagram-urgency` is a signal to use the default. But there is no default value. Instead the default is to use the stream’s `urgency`.

Where stream and datagrams have the same urgency, default recommendation is to share bandwidth between them when packetizing. E.g., 50/50 split between stream data and datagram data, or some other proportion
Capsule and datagram scheduling?

<table>
<thead>
<tr>
<th>Time</th>
<th>Datagram ID</th>
<th>Stream ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>u=0, ID = 24</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>u=0, ID = 28</td>
<td></td>
</tr>
<tr>
<td></td>
<td>u=0, i=?1, ID = 12</td>
<td></td>
</tr>
<tr>
<td></td>
<td>u=0, i=?1, ID = 16</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>u=1, ID = 20</td>
<td></td>
</tr>
<tr>
<td></td>
<td>du=1, Qrtr stream ID = 1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>du=1, Qrtr stream ID = 2</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>du=4, Qrtr stream ID = 0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>du=4, Qrtr stream ID = 0</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>u=6, ID = 0</td>
<td></td>
</tr>
</tbody>
</table>

Stream and datagrams of same urgency share bandwidth

Same stream ID - datagrams more important
Is the problem academic?
Liaisons venn-gèreuses
HTTP Datagrams, UDP Proxying, WebTransport, MoQ, and Extensible Prioritization

draft-pardue-masque-dgram-priority

IETF 115 – London – 2022-11