HTTP Priorities Design Team Update

October 2019

Deployment Data from Chrome/Google

Compared

- SPDY (current deployment)
- Chrome's use of H2 priorities
- Round Robin (H2 default)
- FIFO
- LIFO

Deployment Data from Chrome/Google

Compared

SPDY (current gQUIC deployment) - Quite good
Chrome's use of H2 priorities - Slightly better than SPDY!
Round Robin (H2 default) - Worse than SPDY or H2
FIFO -
LIFO - QoE regressions for VOD video streaming

What we believe we know

We have experimental evidence that N priorities + 1 bit works well for most use cases

We have application performance metrics confirming this

We have experimental and application metrics indicating some options (particularly LIFO and RR) cause regressions

Areas of Consensus

Small number of priorities(~8) + 1 bit for incremental delivery

Defining semantics of the priority levels is worthwhile

Re-prioritization is a requirement

Need for a negotiation mechanism

Focus on the web page load and VOD streaming use cases

Emerging Consensus

Use header and embed the same information in a frame

- Headers are how to accesses the semantics of HTTP
- Allows further extensions and improvements
- Allows the header to be used for a wider set of use cases
- Allows the header to potentially be used even with 1.1

Base on draft-kazuho-httpbis-priority

Proposed Header

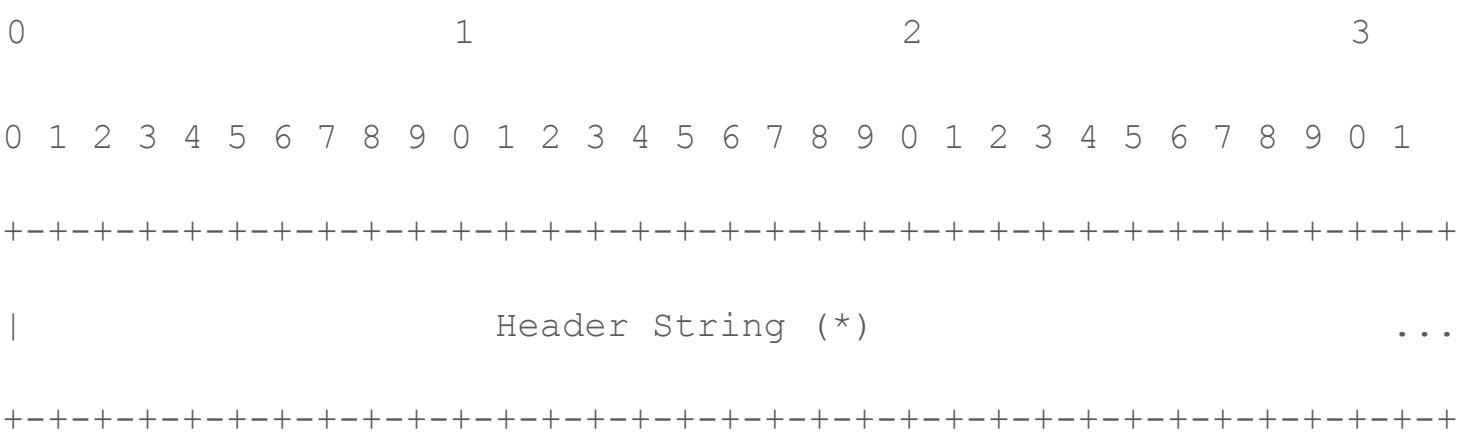
Conveys:

- The SDPY style priority
- A single bit to indicate whether to round robin at a level

```
priority = urgency=3, progressive=?1
```

Proposed Frame

Sent on the control stream



Open issue: Exactly how to indicate support?

ie: draft-lassey-priority-setting-00

May want to indicate support for multiple schemes/features

Header Concerns

Minimal deployment experience with a header

Few existing headers that are similar

Concern that header may ‘leak’ through proxies

Unknown unknowns

Header: Areas of Further Discussion

Is a header Hop-by-hop or End-to-end

And what does End-to-end really mean?

Is a priority header cacheable?

Open issue: Does ‘background’ change transport?

Issue #48 'background' priority should suggest changing congestion control to be less aggressive

```
The background urgency (value 6) is used for responses of
which the delivery can be postponed without having an impact
on using other responses.
```

- Known use cases (ie: software updates, sync)
- But is it too cross-layer?

H2 compatibility

Simple to convert this scheme to HTTP/2 if a server only supported HTTP/2 priorities

Can we ship something in time for HTTP/3?

If it's a new header, do we need to include anything in H/3?

What does the QUIC/HTTP WG want soon?

- A new header in HTTP?
- A frame in HTTP/3 to convey a provably useful system?
- Both/something else?

Thanks!

To all design team members, especially Robin Marx for taking notes and Kazuho and Lucas for their draft