RateLimit Headers

Communicate service status

IETF 106 Singapore

draft-rpolli-ratelimit-headers

[see the specifications]

RateLimit HTTP Header Fields

- communicate service limits, so clients can stop before being throttled out
- align all the *already existing* ratelimit headers and end headers proliferation
- express multiple RateLimit policies

STOP header fields proliferation

X-RateLimit-UserLimit: 1231513

X-RateLimit-UserRemaining

X-Rate-Limit-Limit: name=rate-limit-1,1000

x-custom-retry-after-ms

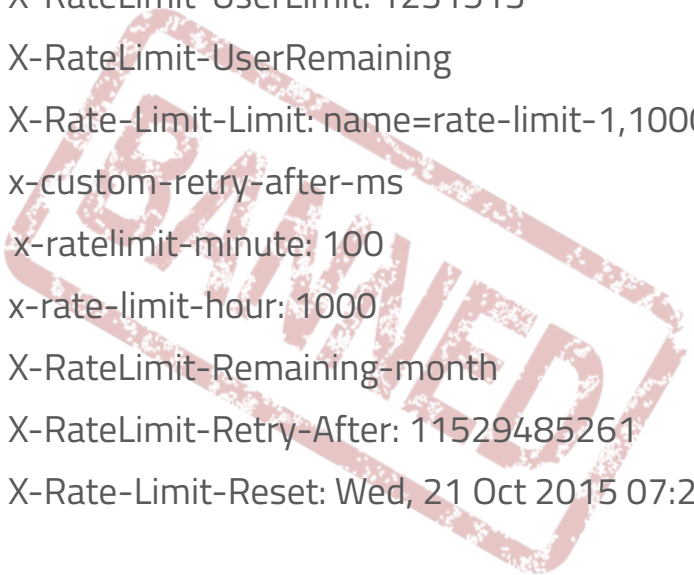
x-ratelimit-minute: 100

x-rate-limit-hour: 1000

X-RateLimit-Remaining-month

X-RateLimit-Retry-After: 11529485261

X-Rate-Limit-Reset: Wed, 21 Oct 2015 07:28:00 GMT



RateLimit-Limit: #quota-units

RateLimit-Remaining: #quota-units

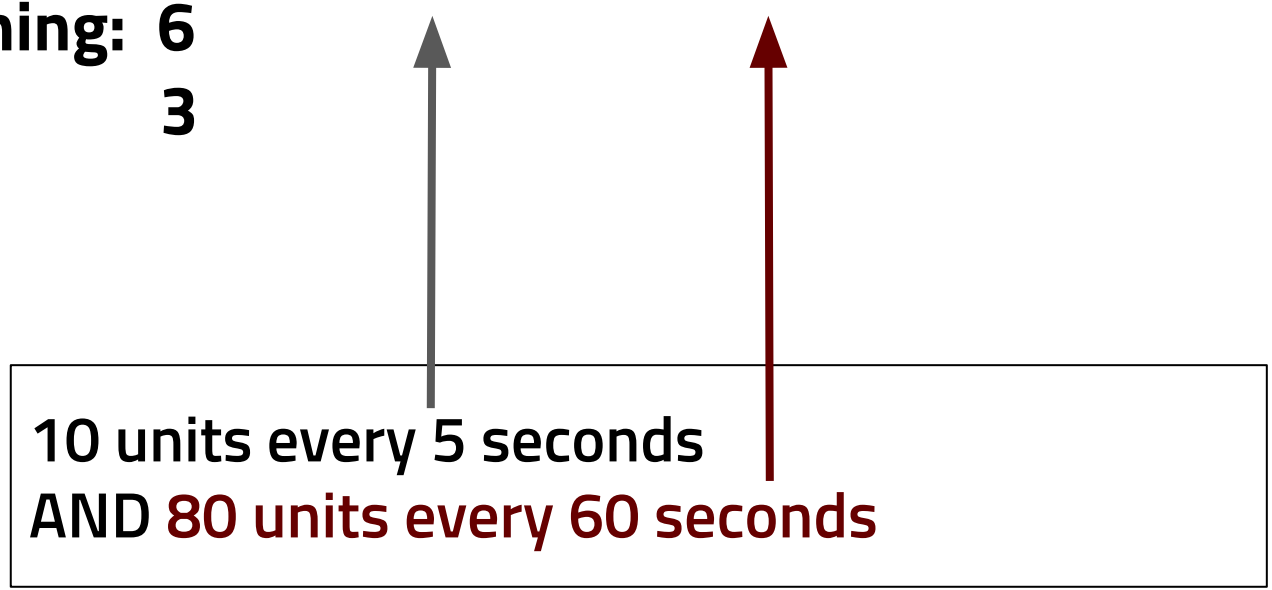
RateLimit-Reset: #delta-seconds

... and many more!

Example with multiple quotas

mandatory part | **optional** parts with policy details and comments

RateLimit-Limit: 10 , 10;w=5 , **80;w=60**;comment="bar"
RateLimit-Remaining: 6
RateLimit-Reset: 3



Technical choices

- #60 support only delta-seconds (no ntp skew & adjustment issues) like Retry-After
- #49 quota expressed in units, may or may not be requests support multiple quota policies and comments
- flexible semantics to express dynamic policies, sliding windows and concurrency limits
- don't mention infrastructural concepts like connections

https://github.com/ioggstream/draft-polli-ratelimit-headers/pulls?q=is%3Apr+is%3Aclosed

Open Issues Needing Input

- http-core#99 Define a throttling scope, related to Retry-After
- #42 Define header dependencies
- #35 Use Structured-Headers
- #41 Upper bound for RateLimit-Reset
- #34 Header names

https://github.com/ioggstream/draft-polli-ratelimit-headers/issues

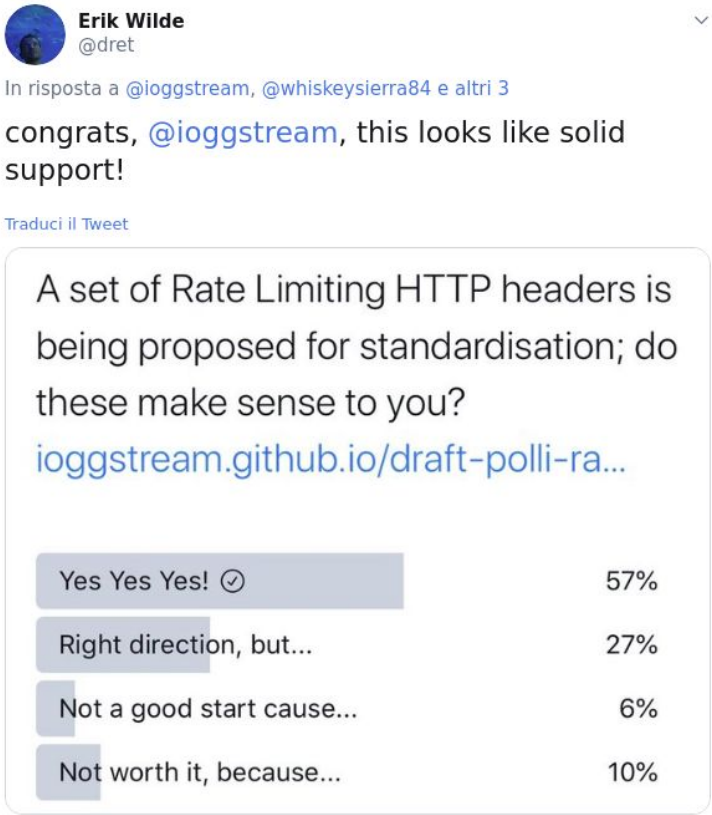
Thanks!

Roberto Polli - robipolli@gmail.com

Alex Martinez - amr@redhat.com

Backup slides

Who wants it?


Erik Wilde
@dret
In risposta a @ioggstream, @whiskeysierra84 e altri 3
congrats, @ioggstream, this looks like solid support!
Traduci il Tweet
A set of Rate Limiting HTTP headers is being proposed for standardisation; do these make sense to you?
ioggstream.github.io/draft-polli-ra...
Yes Yes Yes! 57%
Right direction, but... 27%
Not a good start cause... 6%
Not worth it, because... 10%

Example...after 40 seconds

mandatory part | **optional** comment parts with policy details

RateLimit-Limit: 80 , 10;w=5 , 80;w=60;foo="bar"
RateLimit-Remaining: 0
RateLimit-Reset: 20

^--- now use this

After 40 seconds, client consumed 80 units.
The enforced quota is the second one.

Why proliferation is bad?

Currently every API gateway implements custom ratelimit headers

Clients consuming APIs behind different gateways have to support different ratelimit headers.

The reality is that they ignore them