HTTP/2 ORIGIN Frame
(and related challenges)

Erik Nygren <erik+ietf@nygren.org>
April 4, 2016
The problem...

- When to (re)use a connection for a given Origin?
- Server identity validation (eg, TLS cert) is a MUST for HTTPS
- Challenges / under-specified:
  - Require IP handed out in the DNS? (for Origin name or Alt-Svc server)
  - Allow pushed server certs?
  - DNS TTL Expiry? When to re-resolve and re-connect?
  - Allow reusing conns for multiple Origins for improved perf?
- Reuse *sometimes* desirable for perf
- Reuse *often* has operational challenges
What RFC 7540 has to say

10.1. Server Authority

HTTP/2 relies on the HTTP/1.1 definition of authority for determining whether a server is authoritative in providing a given response (see [RFC7230], Section 9.1). This relies on local name resolution for the "http" URI scheme and the authenticated server identity for the "https" scheme (see [RFC2818], Section 3).
Some issues… (partial list)

● Servers may have certs covering names they aren’t yet prepared for:
  ○ Not yet/still “live” (eg, transitioning hosts to/from CDN or hosting provider)
  ○ Multi-CDN/Hoster Load Balancer
  ○ Using features not yet H/2 tested and ready
  ○ Different levels of production/staging
  ○ Wildcard, SAN, etc certs make this worse

● Different origins may prefer different connections for various reasons:
  ○ Different preferred cipher suites or TLS config; client certs
  ○ Different load balancing / QoS / mapping
  ○ Desire H/2 for some and HTTP/1.1 for others
  ○ DNS / Mapping TTL expired
  ○ …
Potential mechanisms

- “421 Not Authoritative”
- Alt-Svc / ALTSVC
- GOAWAY
- ORIGIN frame ←
- Perhaps Others:
  - DNS record push (requiring DNSSEC signatures?)
  - Server cert push
- Caveat: many of these don’t help unprepared/misconfigured Origins if clients overly optimistic about reuse
Security challenges

- Every new mechanism we add for connection reuse when Origin is not from DNS-resolved-IP increases exposure to compromised server identities (or injected MitM CAs)
  - Expands from requiring local/inline attack or DNS poisoning to many new vectors
  - Opportunities for attackers to combine vulnerabilities in new and “exciting” ways
The big open question(s)...

- Using the ORIGIN frame to constrain reuse seems safe and valuable

- When is it safe (and a good trade-off) to increase scope?
- What are good defaults for conn usage and reuse?
- What operational guidance should we give for clients?
- How do servers know how clients will behave?