BBR Congestion Control Draft

draft-ietf-ccwg-bbr-05

Internet Draft Editors:

Neal Cardwell (Google), Ian Swett (Google), Joseph Beshay (Meta)

Speaker: Ian Swett

Outline

Overview

- Outline recent BBR Internet Draft updates
- Summarize open issues (3) and open pull requests (2)

Goals for this talk:

- Provide a road map for...
  - Readers of the draft
  - Implementers of BBR reading the draft
  - Members of the CCWG/ICCRG community who would like to contribute
- Inviting the community to…
  - Read the draft
  - Contribute to the draft

Overview of draft-ietf-ccwg-bbr

- BBR was adopted as a CCWG WG item in October 2024
- Intended status: experimental RFC
- IETF CCWG members are collaborating on github:
  - https://github.com/ietf-wg-ccwg/draft-ietf-ccwg-bbr
- Latest published revision is at:
  - https://datatracker.ietf.org/doc/draft-ietf-ccwg-bbr/
- Latest editor's copy (with hot-off-the-press changes not in published revisions above):
  - https://ietf-wg-ccwg.github.io/draft-ietf-ccwg-bbr/draft-ietf-ccwg-bbr.html
- Draft editors:
  - Neal Cardwell (Google)
  - Ian Swett (Google)
  - Joseph Beshay (Meta)

Changes in draft-ietf-ccwg-bbr-05: summary

Changes between draft-ietf-ccwg-bbr-04 and draft-ietf-ccwg-bbr-05 [text diff]:

Logic updates

- Fix pacing spec in "application-limited" section by @nealcardwell in #47
- Specify behavior to handle spurious loss recovery episodes by @nealcardwell in #98
- Clarify what to do when looking at packet loss if SACK is not available by @nealcardwell in #97
- Update Drain logic to add exiting after 3 RTTs by @jbeshay in #101

Changes in draft-ietf-ccwg-bbr-05: summary

Editorial changes

- Generalize Transport-Specific language:
  - Remove reference to TCP-specific _seq variables by @jbeshay in #99
  - Generalize Offload Budget for QUIC by @ianswett in #88
- Pseudocode fixes:
  - Fix arguments to IsNewestPacket() by @ianswett in #89
  - Replace uses of "?:" operator with then clause omitted with the equivalent "max()" expression by @antoniovicente in #91
  - Fix typo, clarify "bubble" and dropping short-term model in ProbeBw_REFILL by @jbeshay in #96
  - Use MarkConnectionAppLimited() function to remove a duplicate statement by @nealcardwell in #93
  - Minor editorial fixes by @jbeshay in #100

Open Issue 1 of 3: test cases section in the draft

Open editorial issue:

- Consider adding test cases to an appendix #71

Interest in a list of test cases for implementers to use for verification

- Qlog traces in an external repository
- Use of a common tool that can test and trace arbitrary binaries with network activity under certain conditions

Open Issue 2 of 3: generalizing to non-TCP transports

Open editorial issue:

- Section 5.5.9 sounds very TCP-specific #69

The intent is to make the draft as transport agnostic as possible.
We are making progress but not done.

Goal: Ensure implementation of BBR across as many transports as possible

Non-Goal: Create universal approach for mapping any congestion control to any transport

Open issue 3 of 3: organizing definitions of variables

Open editorial issue:

- Per-Packet and Connection State variables are spread across multiple Sections #68

PR: ack_phase: pending experiments

An open PR for an algorithm change waiting for performance data from experiments:

- #5 Remove BBR.ack_phase from pseudocode
  - A minor algorithm simplification
  - Has one implementation (mvfst QUIC BBR2)
  - We'd like
    - A second implementation
    - Internet performance data to verify there is no performance regression

PR: Delivery rate sampling and restarting from idle

An open PR for an algorithmic question for generalizing from TCP to other transports:

- Clarify Unacknowledged rather than inflight for rate sample #56
  - In Sec `4.1.2.2. Transmitting a data packet` of -03
  - **How to check for a connection "restarting from idle", in a transport/implementation-agnostic way?**
    - e.g.: connection is certain that no data packets are in the network
  - TCP-centric version was: `if (SND.NXT == SND.UNA)  /* no packets in flight yet? */`
  - `if (C.inflight == 0)` is not sufficient, for subtle reasons
    - packets spuriously marked lost and later delivered would have incorrect timestamps that could cause significant bandwidth overestimation
  - PR proposal is: `if (UnacknowledgedBytes() == 0)`
  - Open for discussion; ideas are welcome

Conclusion

- Inviting the community to...
  - Read the draft: draft-ietf-ccwg-bbr
  - Offer contributions/comments/edits, in whatever manner you prefer
- Thanks!

Backups...

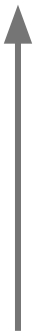
Changes in draft-ietf-ccwg-bbr: how to view

- To see recent changes, you can use several approaches, depending on your preference:
  - From the command line:
    - git clone https://github.com/ietf-wg-ccwg/draft-ietf-ccwg-bbr.git
    - cd draft-ietf-ccwg-bbr/
    - git log -p
  - From github:
    - Commits
    - Merged pull requests

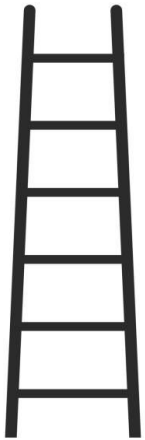
Thoughts about ways to contribute

- Contributions at any "rung of the ladders" below are welcome!
- The higher on the "ladders" (the more concrete/specific/tested the contribution is)...
  - The more useful to the BBR draft effort
  - Given editor time constraints, the more likely the eventual inclusion in the draft
- To finalize significant algorithm changes, we'd like to ultimately reach the top rung of the ladder
- Collaboration encouraged: e.g., idea from person A, implemented by person B, tested by sites B/C

Editorial changes:



- Github pull request with draft text
- Github issue describing the idea
- CCWG email/meeting suggestion



Technical algorithm changes:

- Multiple at-scale Internet deployments
- At-scale Internet deployment data
- Lab/simulation experiment results
- Patch to an open-source BBR code base
- Github pull request with draft text
- Pseudocode
- Github issue describing the idea
- CCWG email/meeting suggestion

Goals of evolving the BBR draft text

Goals as we evolve the BBR draft text:

- Clarification
- Simplification
- Better coexistence with Reno/CUBIC
- Better performance
- Avoiding performance regressions in the real world

Proposed bar for publication (keep in mind the target is an experimental RFC):

- Multiple deployments at scale in QUIC and TCP
- Text both TCP and QUIC implementations can follow
- Fair sharing with other BBR flows, coexistence with Reno and Cubic

Thesis: It's better to publish a good draft with deployment experience in a reasonable timeframe than evolve BBR indefinitely without shipping an RFC.