

TechTrend Talk Series vol.11

L4Sで低遅延インターネットは実現できるか？

IIJ技術研究所 長 健二郎

2024-05-28

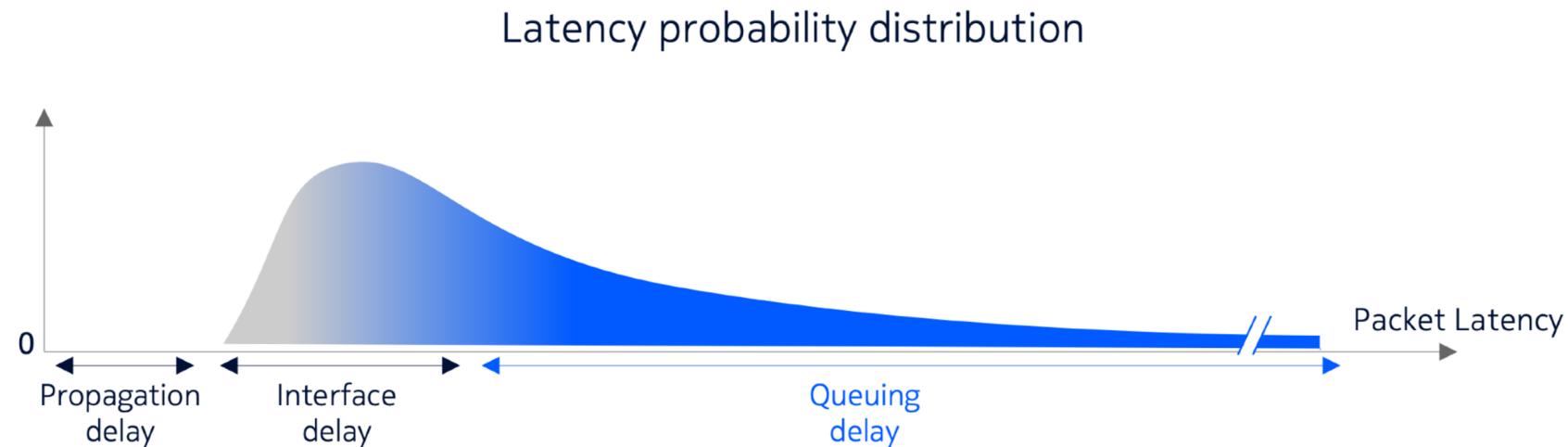
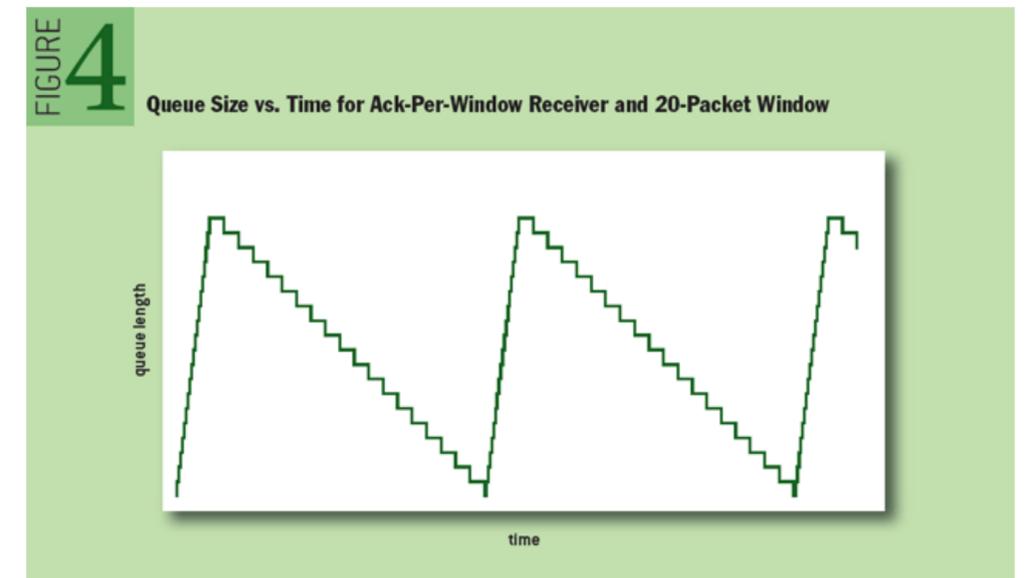
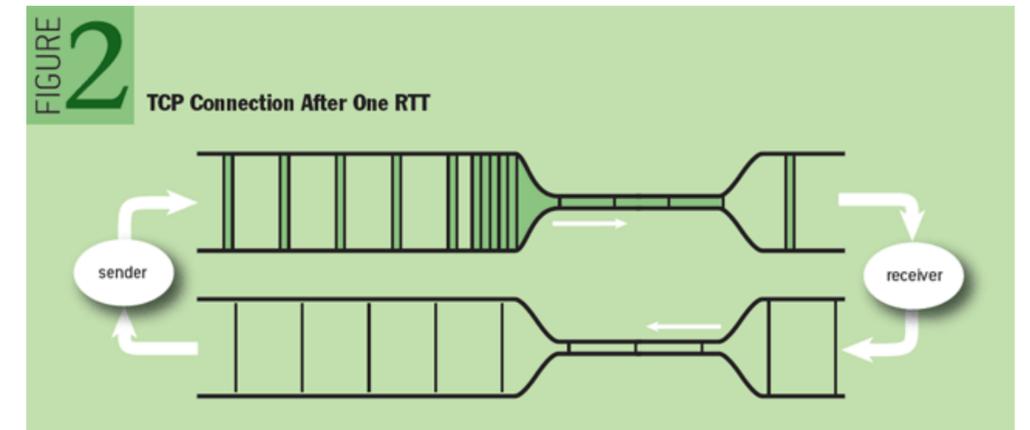
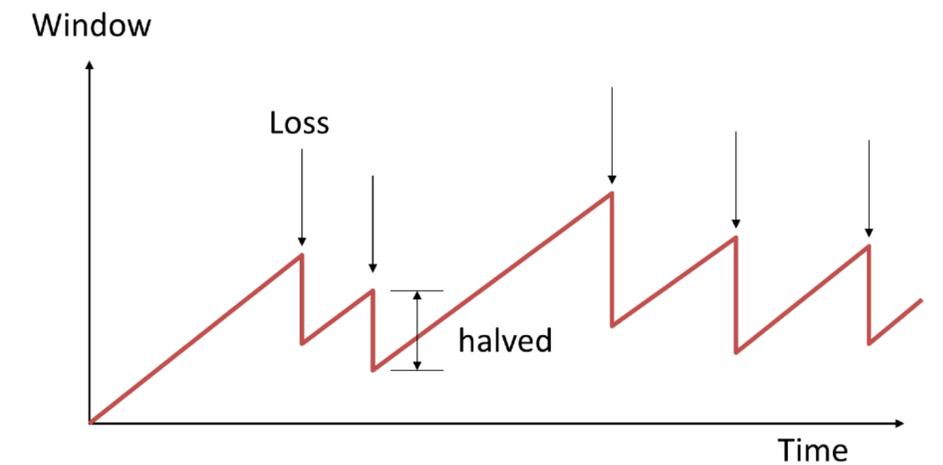


背景：低遅延サービス

- スループットより低遅延（ジッタ）
 - いまや遅延の差が品質を左右（ゲームだけでない）
- 低遅延の実現に特別なプロトコルはいらない？
 - 既存のアプローチはエンドで頑張る方向
 - 問題の本質はTCP/IPの輻輳フィードバック（の欠如）
 - バッファリング遅延を200msから1msに出来る！？

現状TCPは必要以上の遅延

- バッファリング遅延（待ち行列遅延）
 - 普段5msの遅延でも輻輳時は100ms超に！
 - TCPのcwndは帯域幅×RTT
 - 多くがボトルネックのバッファに溜まる
- バッファブロート問題
 - 負荷時の遅延計測の必要性が認識されてきている



L4Sとは

- TCPの遅延を劇的に減らす技術（TCPに限らない）
 - ECN改良によるスケーラブルな輻輳制御（CC）
- 技術的には、DCTCPを広域インターネットで利用可能に
 - レガシーTCPとの安全な共存
 - インクリメンタルなディプロイメント

Low Latency, Low Loss, and Scalable Throughput (L4S)

- DCTCP [SIGCOMM2010]のアイデア
 - スケーラブルな輻輳制御： Scalable TCP[Kelly2003]
 - ECNを即時信号に変更し送信側で解釈
 - 2015年からIETFで標準化に向けた活動
- 2023年1月にRFC 3本
 - RFC9330: L4S architecture
 - RFC9331: ECN protocol for L4S
 - RFC9332: DualQ coupled AQM for L4S
- 実装
 - TCP Prague in Linux, Apple next MacOS/iOS, Google QUIC, DOCSIS 3.1
 - Comcastが実証実験

L4Sの簡単な歴史

- いろんな話が関連、情報が散らばっていて分かりにくい
 - IETFでも複数のWGで標準化が進行：tsvwg, tcpm, iccrg, etc
- 90年代からのQoSやIPngのアイデアの発展の歴史
 - L4Sは誰かの提案というよりコミュニティの知見の集大成
- RED (CCR1991), ECN (CCR1994, RFC3168 2001)
 - diffserv (RFC2475, 1998), IPv6 (RFC2460, 1998)
- Scalable TCP (2003), high-speed TCP, Fast TCP, etc.
- DCTCP (SIGCOMM2010)
 - accurate ECN (ID2011-)
- IETF: TCP Prague requirements for L4S (2015) ここからすでに9年

scalable congestion control

- recovery timeがスループットに拘らず一定

Renoではスループットが10倍になると、リカバリタイムも10倍に

100Mbps x 30msRTT
reno cnwd: 250 pkts
recv time: 125 RTTs

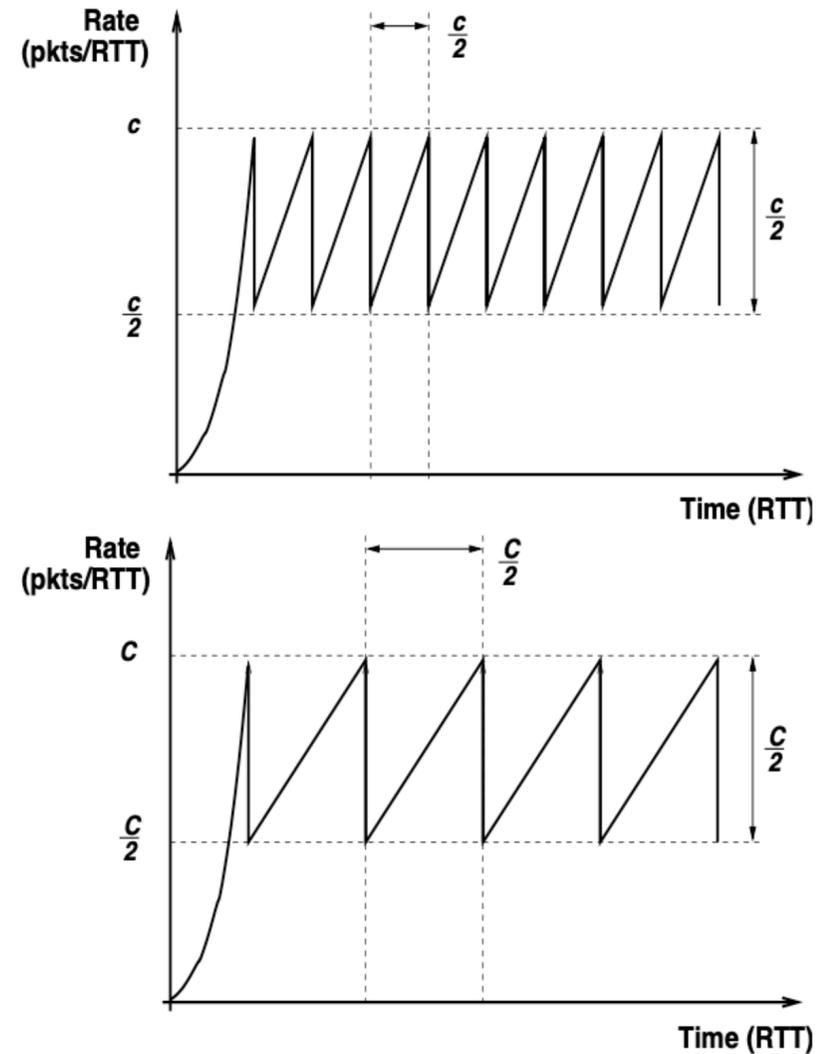


Figure 1: Traditional TCP scaling properties.

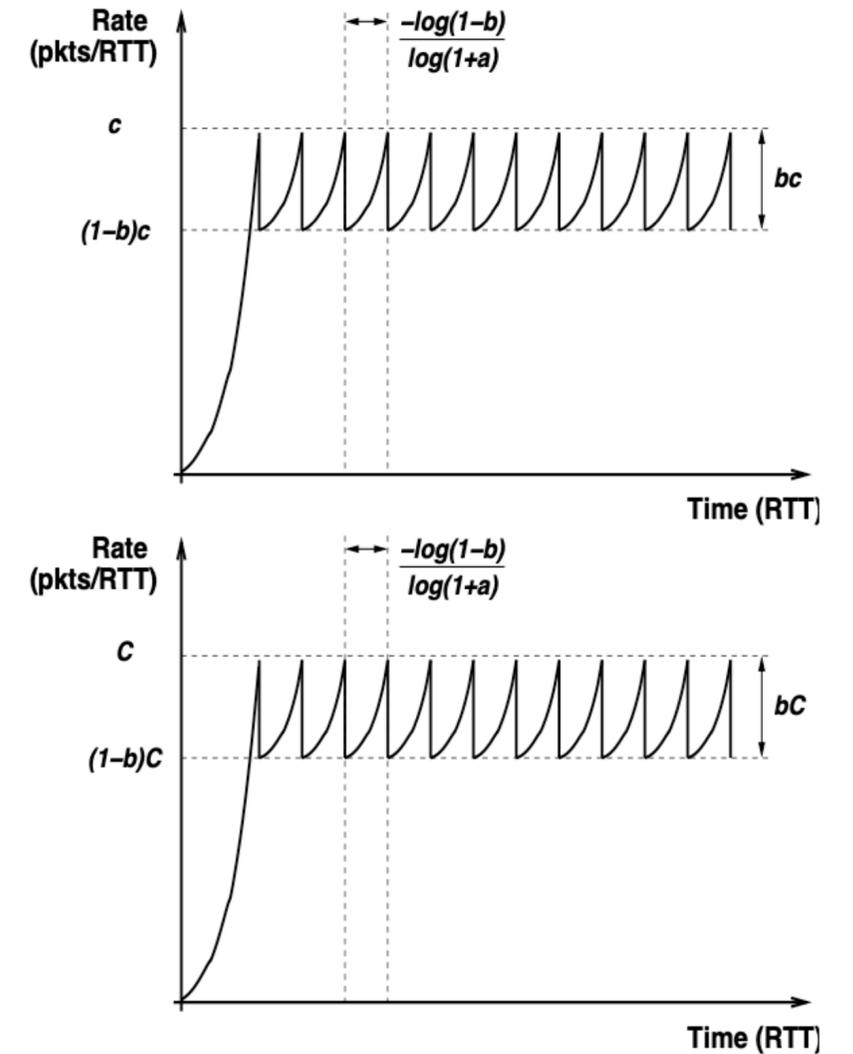
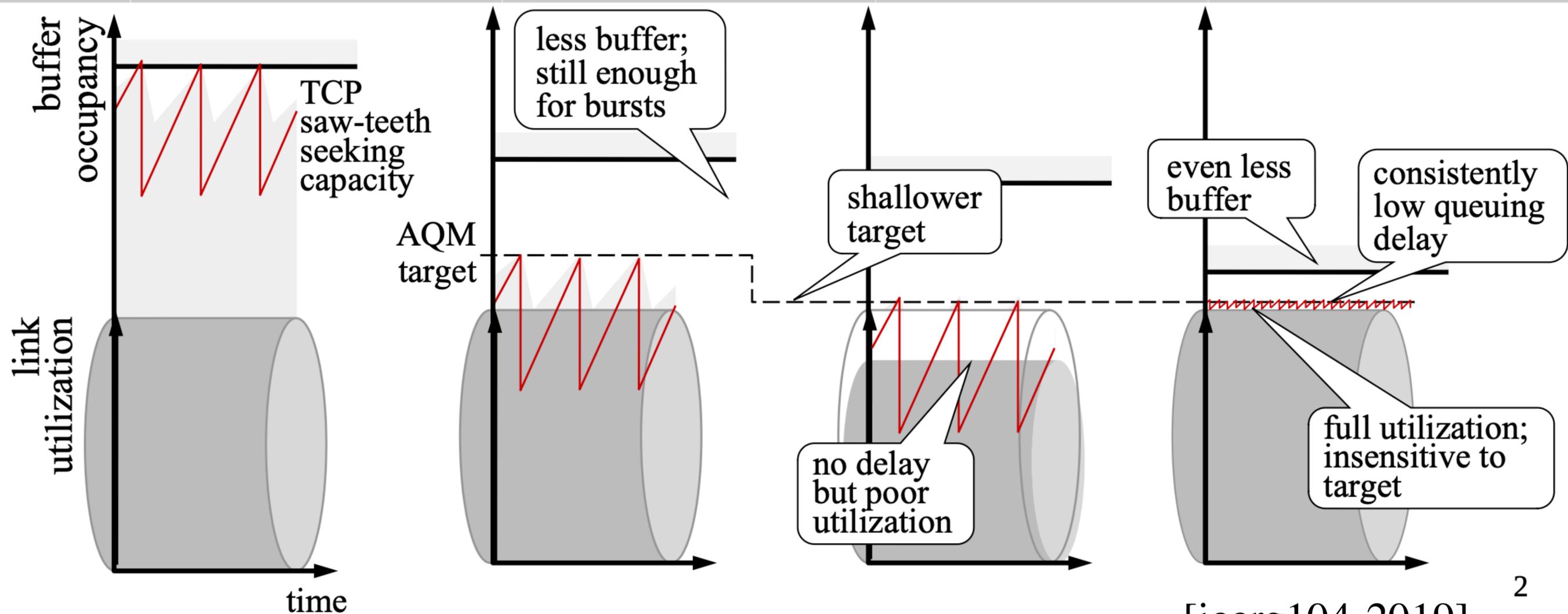


Figure 2: Scalable TCP scaling properties.

Scalable TCP[Kelly2003]

The trick: scalable congestion control

	① Today (typical)	② Today (at best)	③ Unacceptable	④ L4S
Bottleneck	Bloated drop-tail buffer	AQM	Shallower AQM	Immediate AQM
Sender CC	Classic	Classic	Classic	Scalable (tiny saw-teeth)



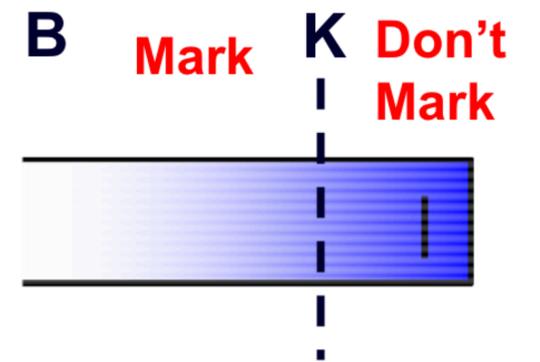
DCTCP

- データセンター利用を前提にTCPとECNを改造 (SIGCOMM2010, RFC8257 2017)
 - ECN: 輻輳信号を即時信号に変え、その平滑化を送信側で実施
 - TCP: 輻輳信号割合だけcwndを縮小
 - 大幅に性能改善

Data Center TCP Algorithm

Switch side:

- Mark packets when **Queue Length > K**



Sender side:

- Maintain **moving average** of **fraction** of marked packets (α)

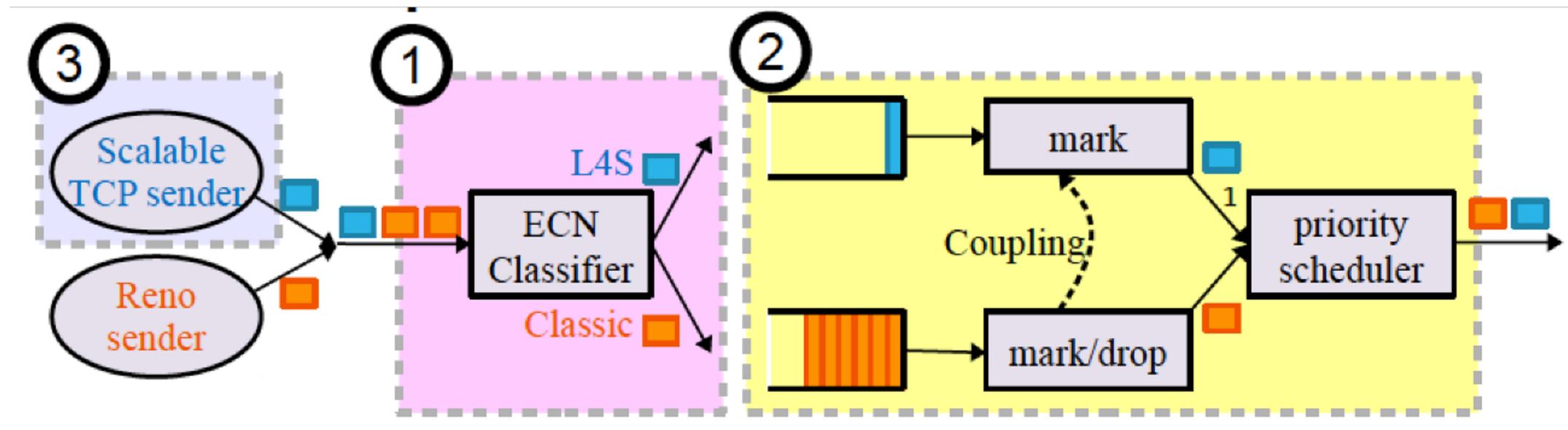
$$\text{each RTT: } F = \frac{\# \text{ of marked ACKs}}{\text{Total \# of ACKs}} \Rightarrow \alpha \leftarrow (1-g)\alpha + gF$$

- Adaptive congestion window decrease: $W \leftarrow (1 - \frac{\alpha}{2})W$

[tsv84-2012]

L4Sアーキテクチャ

- 低遅延: スケーラブル輻輳制御で実現
 - ボトルネックのキューにパケットを貯めない (<1ms)
- レガシートラフィックとはキューを分離して共存
 - ECNビットで区別



Host

Protocol

Network

[iccr104-2019]

L4Sを利用可能にするには

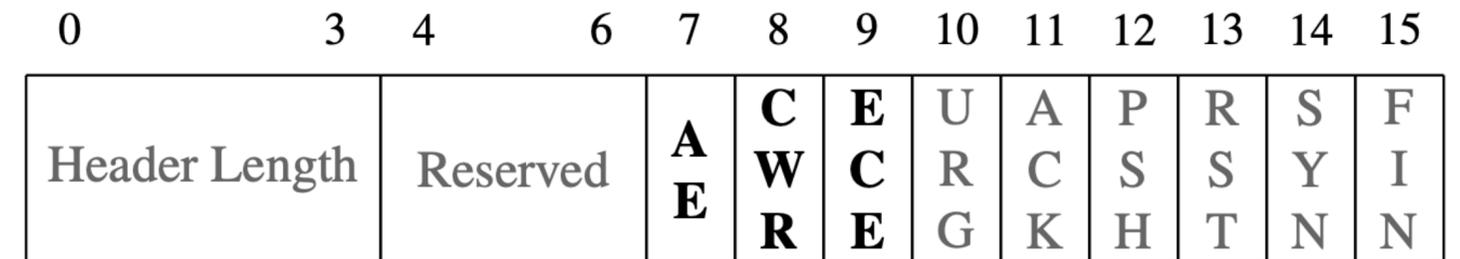
- 送信側：accECNを使ったスケラブル輻輳制御
- 受信側：accECNフィードバック
- ボトルネックルータ：L4Sサポート
 - L4Sのトラフィック分離、L4SにはaccECNマーキング
- インクリメンタルなディプロイメントが可能

More Accurate ECN feedback

- Accurate ECN: 3ビットのカウント値をフィードバック
 - 既存のTCPフラグの3ビットを読み替える
 - コネクション確立時にaccECNを使うかをネゴ
- ECN++: ECN for control packets (e.g. SYN)
- (ちなみにQUICはECNカウント値を返す仕様となっている)

Codepoint	Binary	Meaning
Not-ECT	00	Not ECN-Capable Transport
ECT(0)	10	ECN-Capable Transport (Classic)
ECT(1)	01	ECN-Capable Transport (L4S)
CE	11	Congestion Experienced

2 bit ECN field in IP header



3 bits in TCP header flag field

Dual Queue Coupled AQM

- L4Sを実現する最小限の機能を定義
 - RFC9332付録に具体的なアルゴリズム：DualQ coupled PI2
- カプリング: Reno/L4Sの定常状態のcwndサイズが同等になるようにマーキング確率を調整

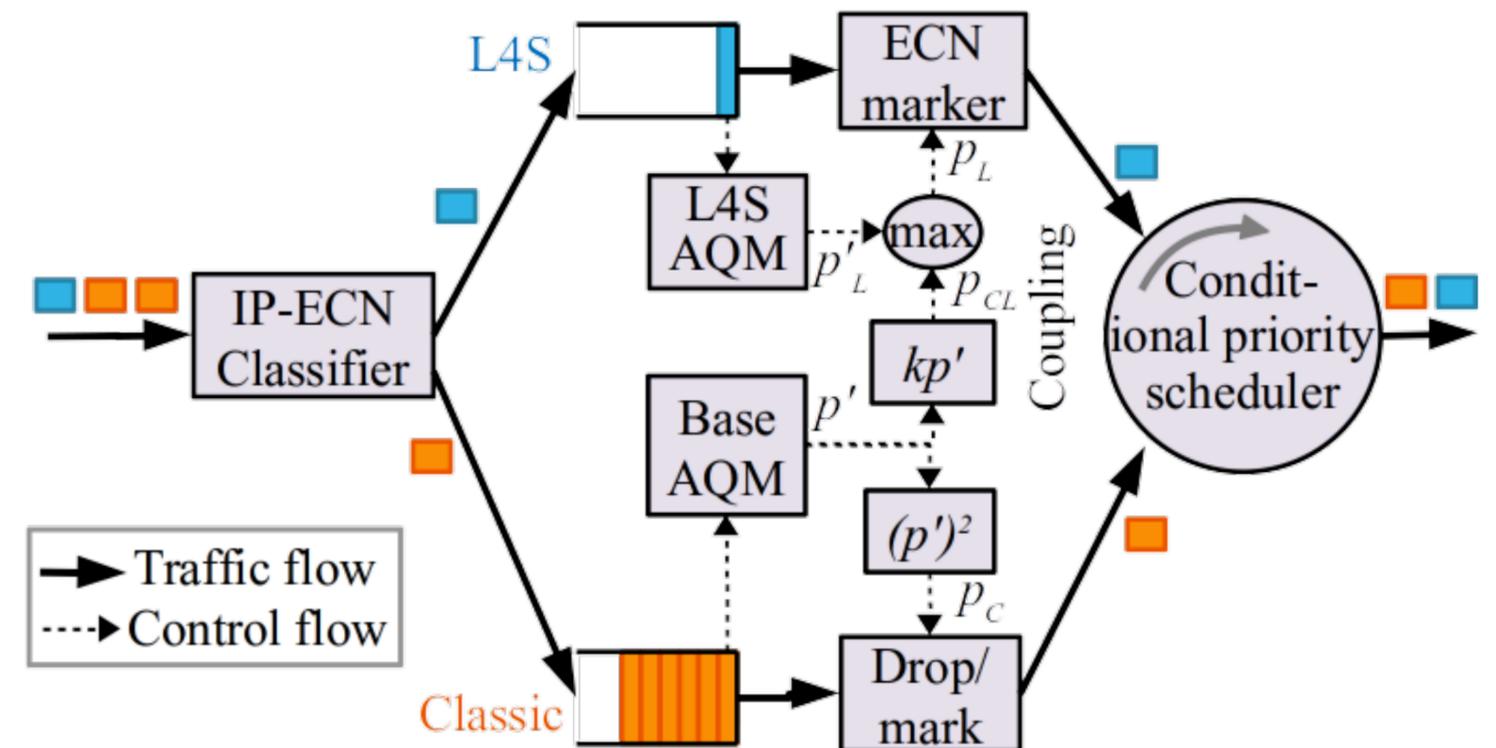


Fig. 3: Dual Queue Coupled AQM: Structure

[Schepper2022]

標準化ステータス

- RFC9330 (2023-01 Informational): Low Latency, Low Loss, and Scalable Throughput (L4S) Internet Service: Architecture
 - B. Briscoe (editor, Independent), K.D. Schepper (Nokia Bell Labs), M. Bagnulo (Universidad Carlos III de Madrid), G. White (CableLabs)
- RFC9331 (2023-01 Experimental): The Explicit Congestion Notification (ECN) Protocol for Low Latency, Low Loss, and Scalable Throughput (L4S)
 - K.D. Schepper (Nokia Bell Labs), B. Briscoe (editor, Independent)
- RFC9332 (2023-01 Experimental): Dual-Queue Coupled Active Queue Management (AQM) for Low Latency, Low Loss, and Scalable Throughput (L4S)
 - K.D. Schepper (Nokia Bell Labs), B. Briscoe (editor, Independent), G. White (CableLabs)
- Internet Drafts:
 - More Accurate ECN Feedback in TCP
 - ECN++: Adding Explicit Congestion Notification (ECN) to TCP Control Packets
 - Non-Queue-Building Per-Hop Behavior (NQB PHB) for Differentiated Services
 - Operational Guidance on Coexistence with Classic ECN during L4S Deployment
 - ISP Dual Queue Networking Deployment Recommendations

実装状況

- Linux Prague TCP/QUIC: L4S development hub
 - <https://github.com/L4STeam>
- Apple
 - WWDC2023: Reduce network delays with L4S
 - developer: `testing_and_debugging_l4s_in_your_app`
 - L4S on iOS,iPadOS,MacOS can be turned on in Developer settings
 - apple's QUIC supports L4S, TCP supports only receiver-side
- Google
 - BBRv2 TCP/QUIC supports L4S
- Comcast: Low Latency DOCSIS field tests since 2023, reported at [tsvwg118-2023][nanog90-2024]
- Nokia and Vodafone: first trial of L4S technology over an end-to-end PON (April 2024)

まとめ

- L4S
 - スケーラブルな輻輳制御、同時に、遅延を劇的に減らす技術！
 - ECNを即時信号へ改良して再定義
 - レガシーTCPとの安全な共存
- 技術は20年以上議論されてきた（L4Sだけでもほぼ10年）
 - インターネットに必要な技術、かつ、よく考えられている
 - もちろん今後改善が必要な部分は出てくる
 - 輻輳がなくなるわけではない：伝搬速度とフィードバック制御の限界
 - 標準化、研究、実装、実証実験、ベンダーとISPのサポート（課題は認知度？）
 - 輻輳フィードバックを普及させる最後のチャンス？
- 懐疑派
 - うまくいかなかった技術の蒸し返しでは？
 - ECN試したけど期待外れ
 - linuxパッチの完成度

展望：L4Sは普及するか？

- 低遅延のメリットの認知度向上
 - 1Gbps回線より1msジッタのサービスの価値
- スケーラブルなトランスポート: 1Gbps超えると必要
- L4Sはメリットしかない、導入のハードルも比較的低い
 - アクセス収容ルータでのサポートが課題
 - とはいえECNもなかなか普及しなかったが、L4Sは違いをアピールしやすい？
 - メリットを視覚化する分かりやすいツールが欲しい
- ケーブルインターネット：DOCSISとComcastが先導（WiFiとの連携も）
 - L4Sの成否を占う試金石
- モバイル：5Gの低遅延を活かすには有線区間も
- ファイバ：他と比較してメリットが見えにくい
 - 日本ではFLET'S網が対応しないと、、、

