

TechTrend Talk Series vol.1

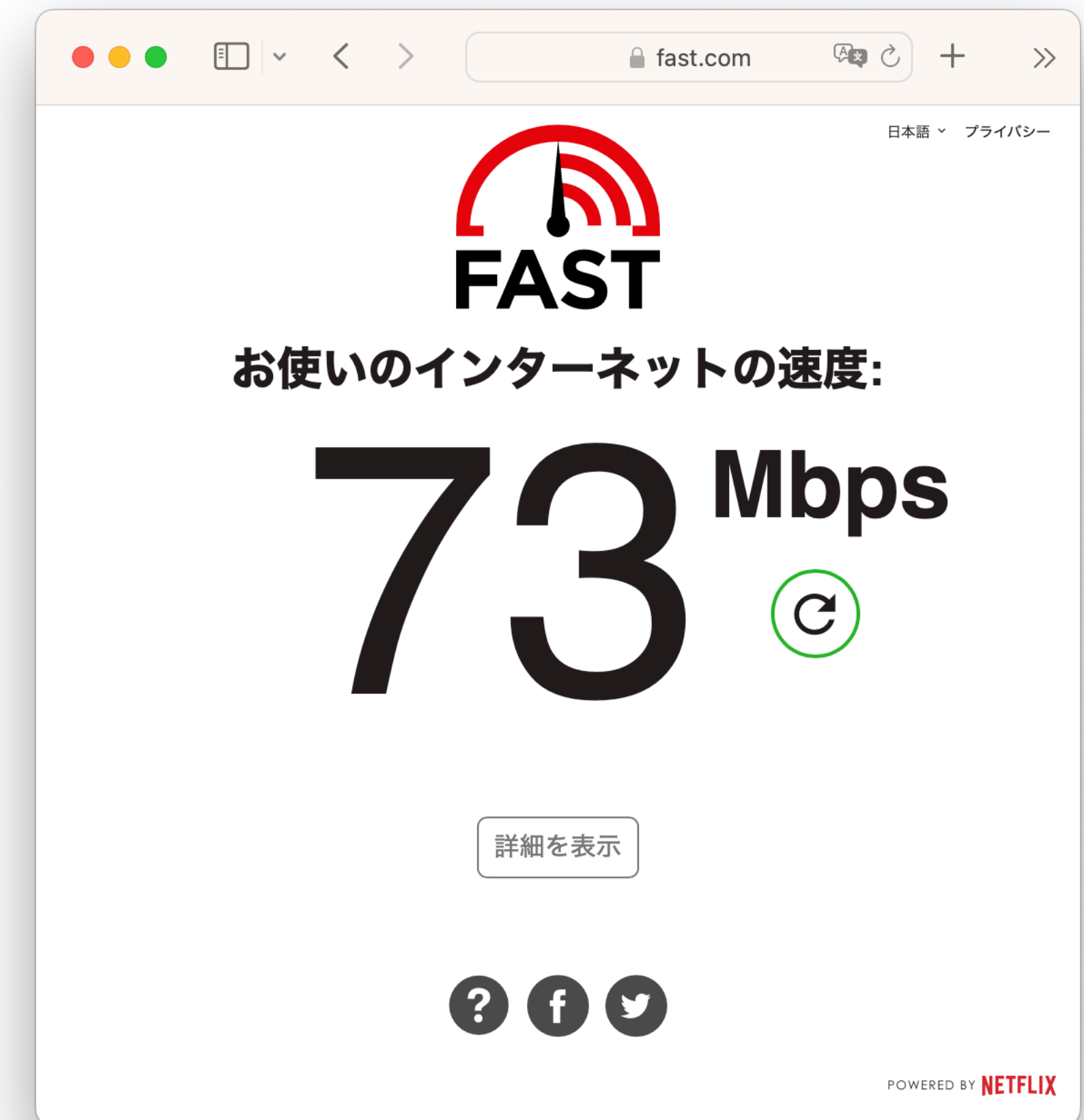
ブロードバンドサービスの品質計測について

IIJ技術研究所 長 健二郎
2023-04-18



インターネットのスピードテスト

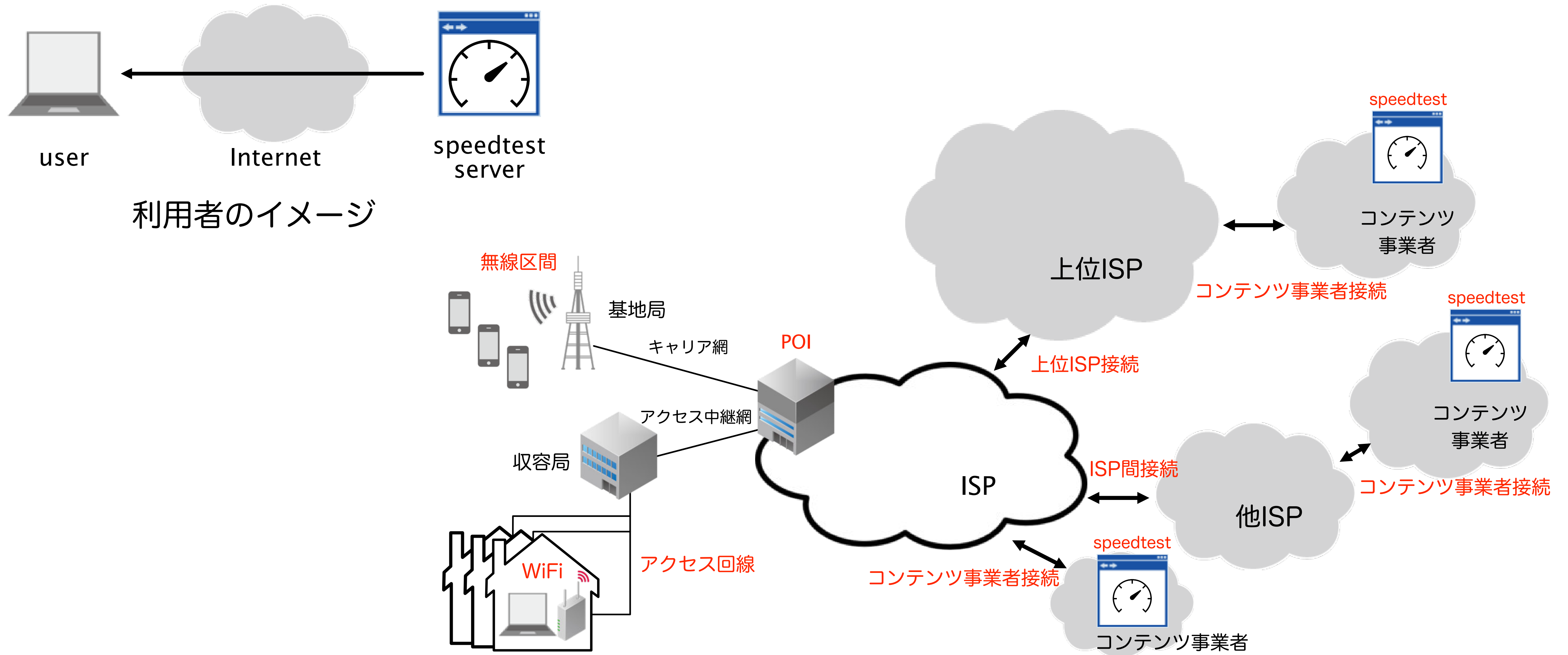
- スピードテストはみんな大好き
 - 使い方が簡単
 - わかり易い（大きい方がい！比較も簡単）
- 正確な計測や結果の正しい理解は意外に難しい
 - WiFiや他のトラフィックの影響などを受ける
 - 統計多重の仕組みとボトルネック箇所の理解が必要
 - 計測方法、まわりの影響と時間変動など、ばらつきの解釈が必要



スピードテストの仕組み

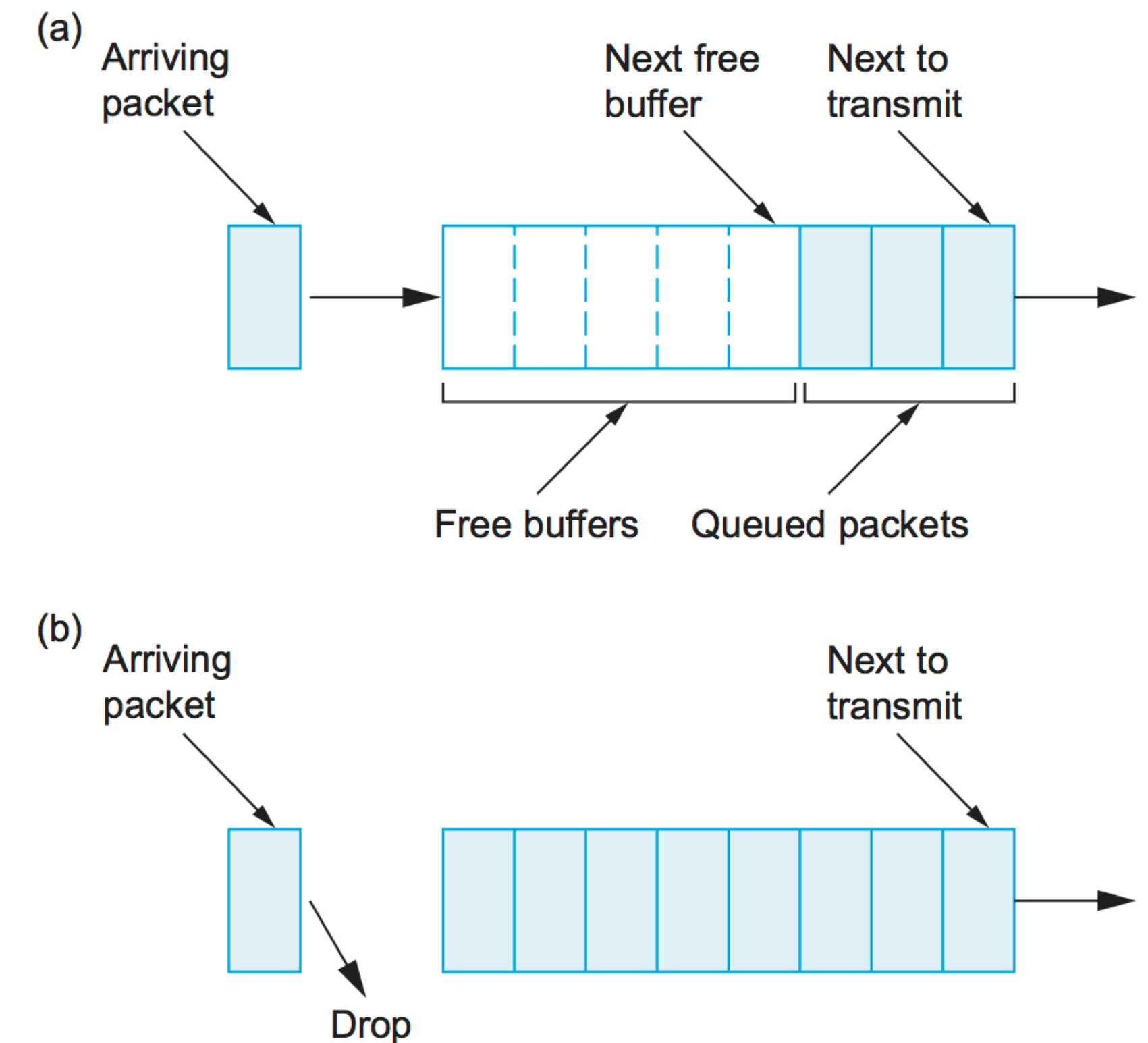
- デバイスと測定サーバーの間でスループット（転送性能）を計測
 - $\text{スループット} = \text{転送量} / \text{転送時間}$
 - 測るのは伝送路の太さ：時間あたりどれだけ沢山のデータ量を送れるか
- スループットは回線の混み方によってばらつく
 - 空いていれば最大性能に近い値が出るが、混んでくると低下する
 - 時間や場所その他環境条件
 - スピードテスト/TCPの違い、サーバ側の処理能力、測定端末、宅内Wi-Fi

潜在的なボトルネックはいろいろ



統計多重とキューイング

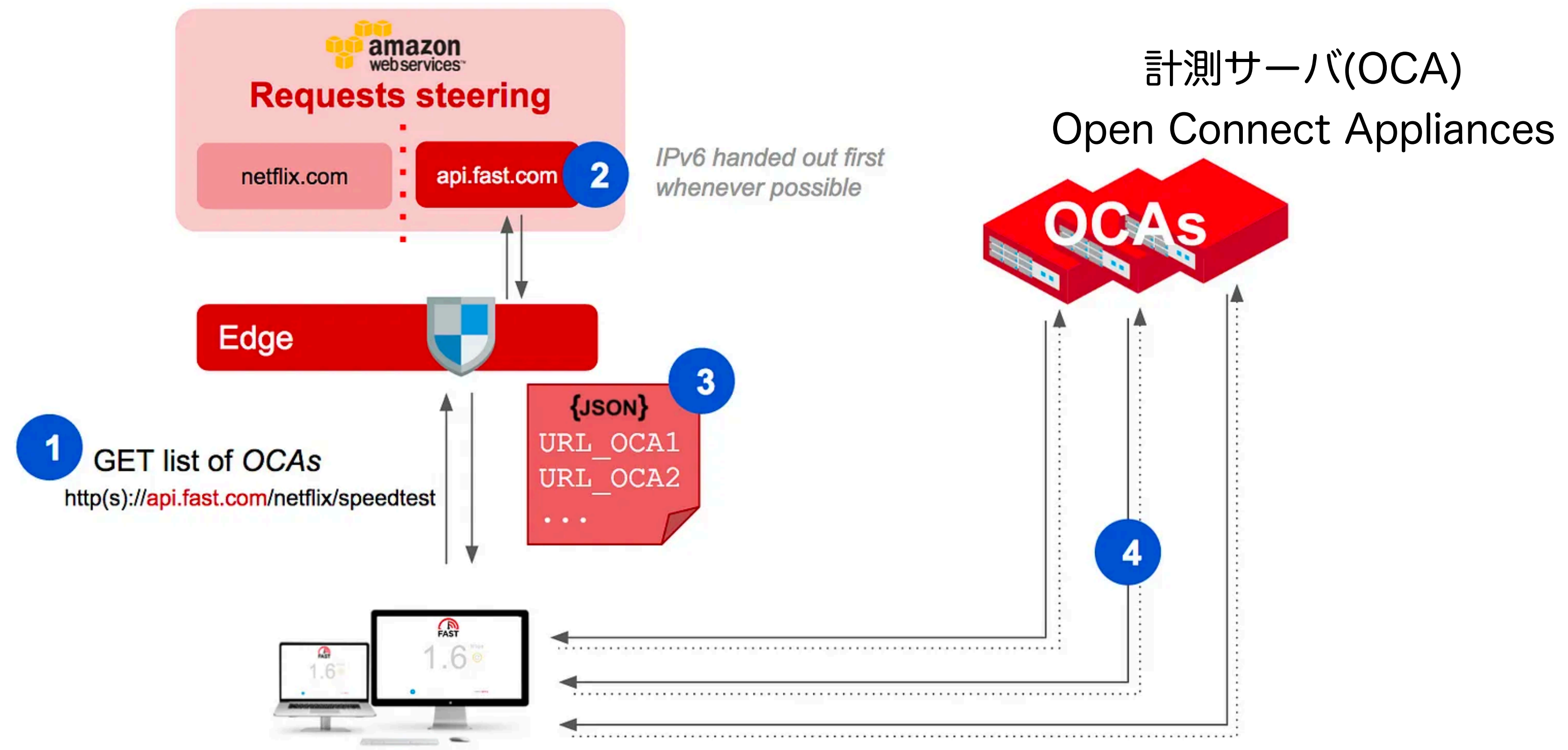
- インターネット通信：資源共有なので安価
- パケット通信と統計多重
 - 間欠的なコンピュータ通信の資源共有
 - 他の人が使っていない時は占有できる
 - 利用が重なると資源を分け合う
- キューイング（バッファリング）
 - 利用の重なり（バースト）を吸収
 - 混んでくると待ち時間が長くなる
 - キューが溢れてパケットロスするとTCPが再送



source: <https://book.systemsapproach.org/congestion/queuing.html>

Fast.comの例：計測の構成

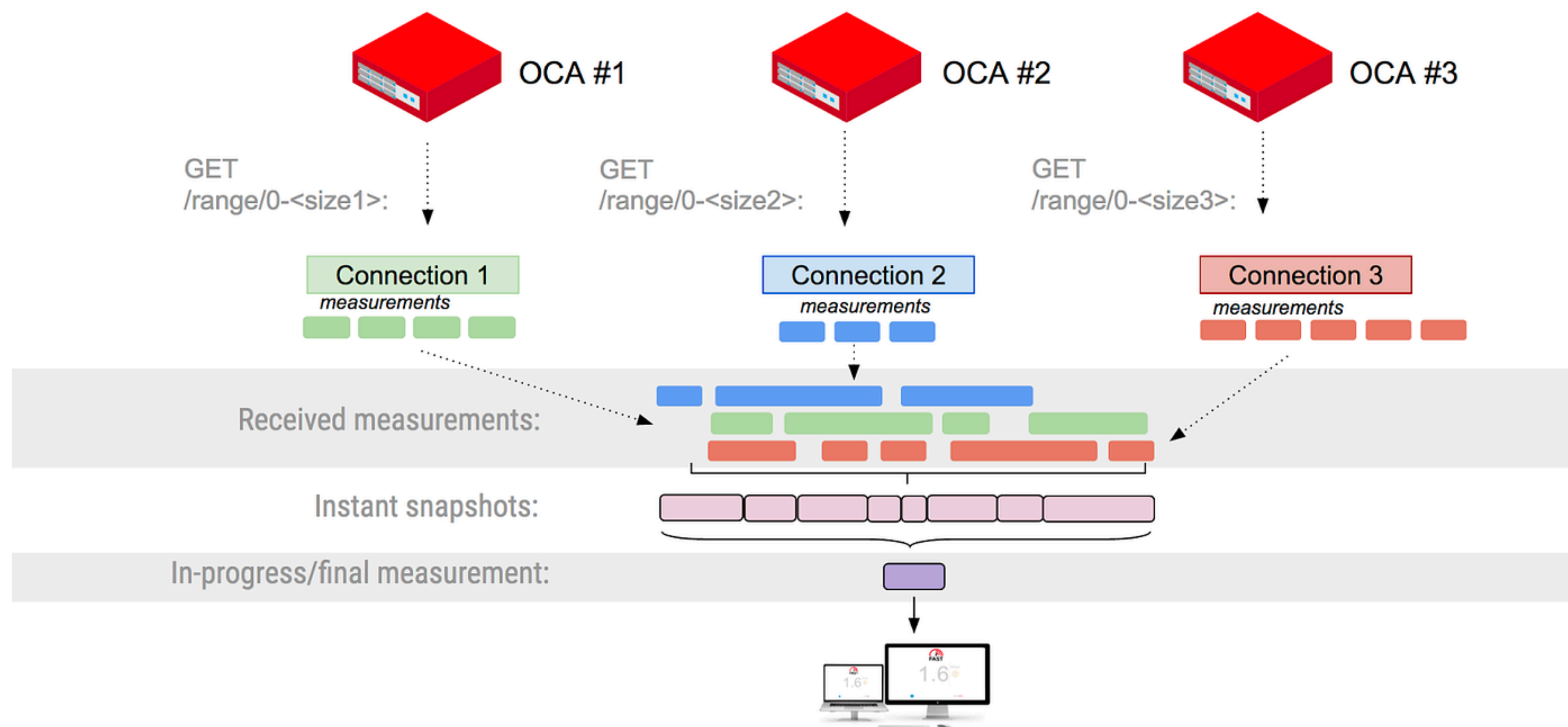
- スピードテストサイトから専用の計測サーバに誘導



source: <https://netflixtechblog.com/building-fast-com-4857fe0f8adb>

Fast.comの例：集計

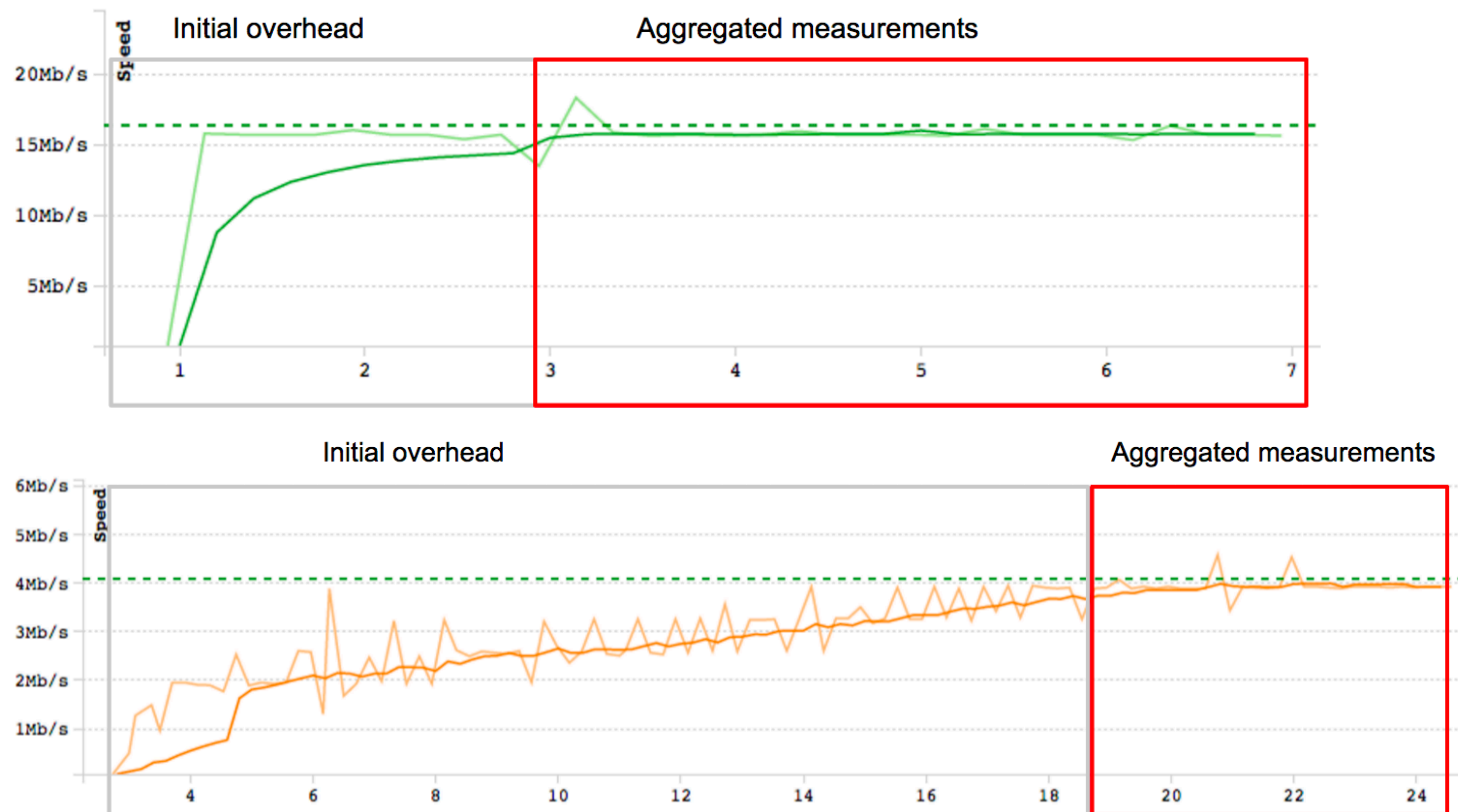
- 計測は複数コネクションのデータを集計



source: <https://netflixtechblog.com/building-fast-com-4857fe0f8adb>

Fast.comの例：測定値の収束

- 測定値の収束を検出して値を得る



source: <https://netflixtechblog.com/building-fast-com-4857fe0f8adb>

スピードテストの仕組み詳細

- 計測項目
 - ダウンロード, アップロード, 遅延とジッタ、IPv4/IPv6
- サーバ選択：正確な計測には近くのサーバが必要
- TCPセッション数：複数束ねて計測
- 計測時間
 - 長いとユーザは離脱（適切な途中経過フィードバックが大切）
 - 1Gbpsで10秒テストすると1.2GBぐらいになる（エコでない）

スピードテストの設計

- オープンソースを例に: libspeed v5.2.5 (2022-01)
 - javascript Ajax (XHR 非同期データ送受信) ブラウザで動く (go versionもある)
 - サーバ選択: http pingを使って最小遅延のサーバを選択
 - テスト項目: http-ping (delay & jitter), DownLoad, UpLoad
 - duration: max 15sec, 速いと早期終了
 - tcp parallel sessions: 6 for DL, 3 for UL (同じサーバを使う)
 - measurement start offset: 1.5sec for DL, 3sec for UL
 - javascript performance APIを計測に使うか
 - overhead: x1.06 (HTTP+TCP+IPv4+ETH)

スピードテスト結果が示すもの

- 性能の良さを示す
 - 何度か計測して一番良い結果を選べば良い
- 品質を示す
 - 結果はばらつくので1回の計測では示せない
 - たくさんのサンプルを統計処理する必要
 - スピードテストに統計処理機能はない
- スピードテストは性能に満足している場合には良いツール、品質を示すには必ずしも適さない

スピードテストの簡単な歴史

- スピードテスト的なものはインターネット初期から存在
 - その昔はFTPで転送時間を測った
 - ベンチマークツール (netperf/iperfなど)：実験環境での性能検証
 - 可用帯域推測技術：帯域を埋めずに計測する工夫（次のスライド）
- 2000年代に入ってブロードバンドが普及
- 2000年代後半からモバイルデータ通信が普及
- 2006 ooklaがspeedtest.net 開設
 - スピードテストが一般に使われるようになる

pathchar: Van Jacobson, 1997

- 測定はtracerouteをパケットサイズを変えながら実行するイメージ
- 可用帯域推測技術：各サイズの最小遅延を線形回帰、傾きが帯域を示す
- 制約：片方向しか測れない、測定精度

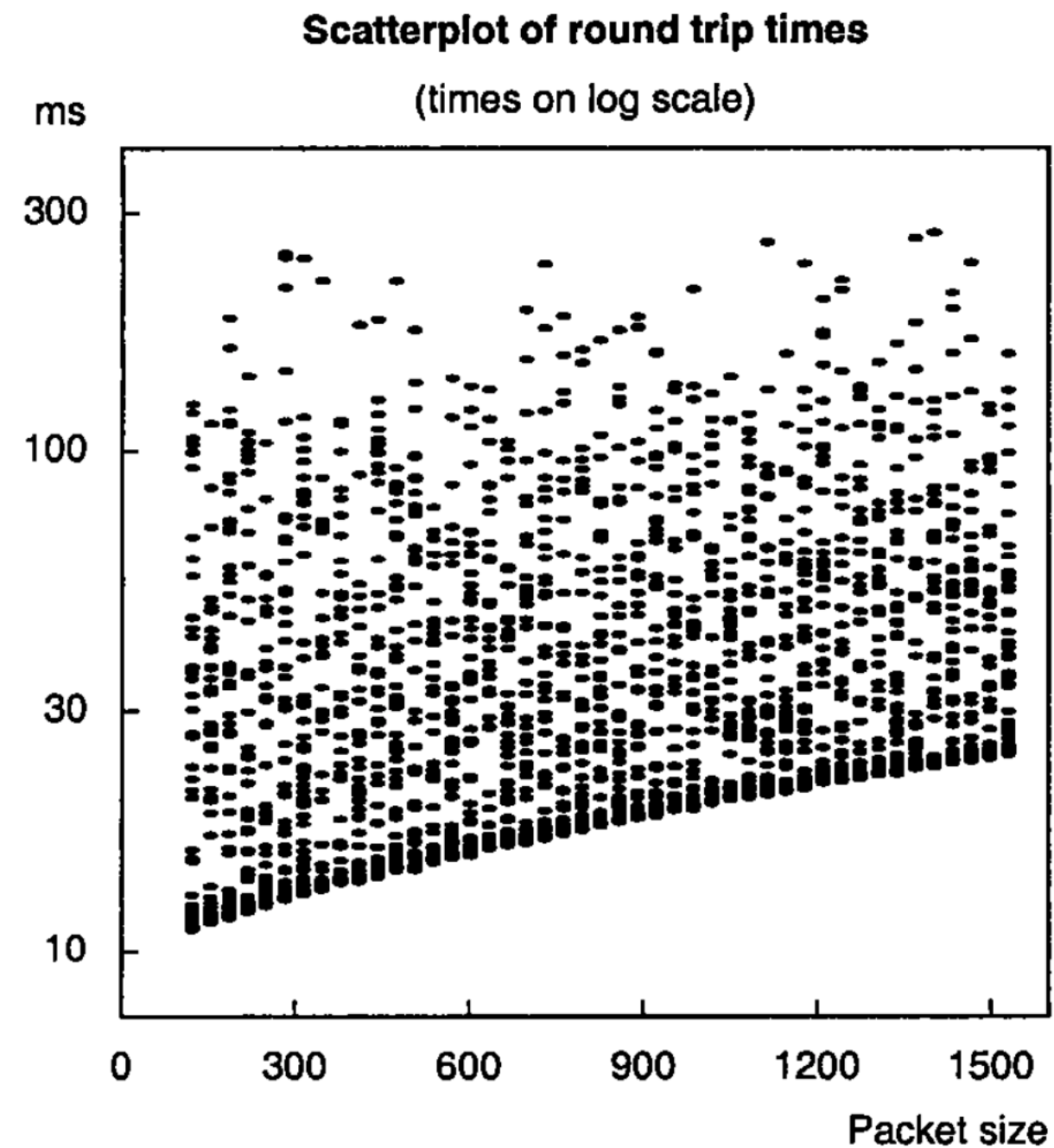


Figure 2: Scatterplot of round trip times versus packet size, 45 sizes, 64 probes each. The vertical axis is on a log scale, which is why the lower envelope appears convex, although it is a nearly straight line.

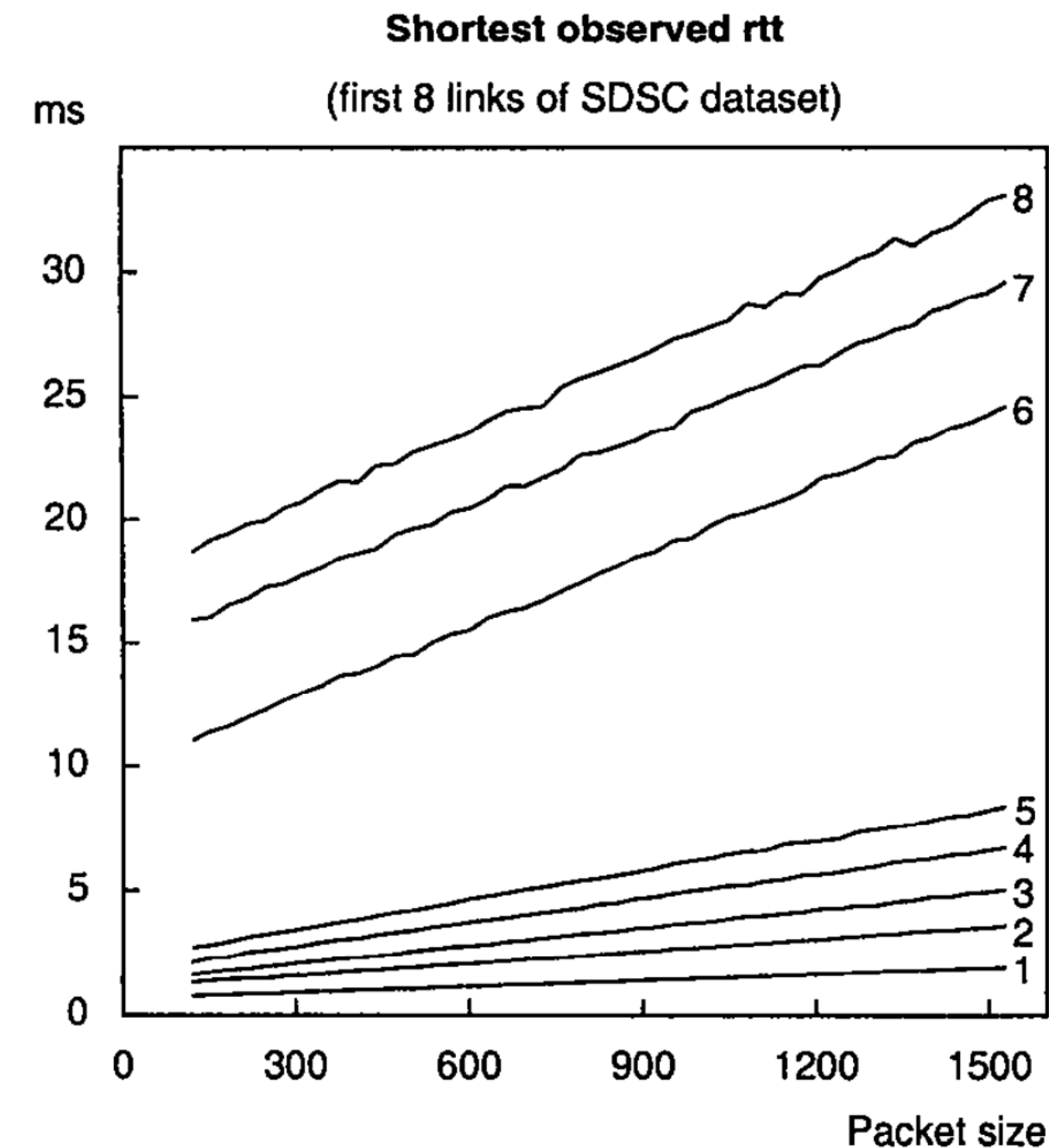
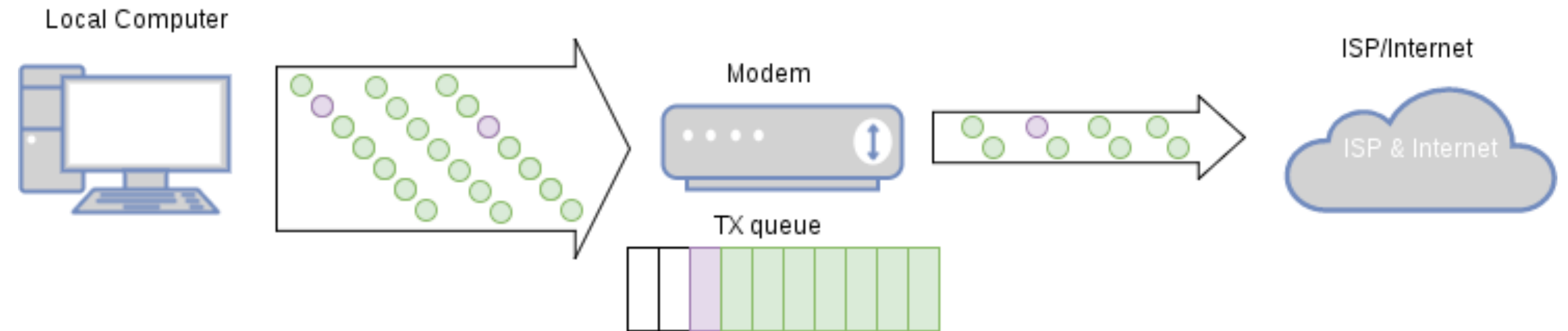


Figure 4: Shortest observed round trip times (SORTTs) for the first 8 links of the SDSC dataset.

source: <https://dl.acm.org/doi/pdf/10.1145/316188.316228>

バッファブloat問題

- 過剰なバッファによる遅延 (buffer bloat)
 - ホームルータで秒単位の遅延が生じる報告もある
 - アップロードの遅延はダウンロード性能にも影響
- アップロード時の遅延増加を検出するツール
 - 一部のスピードテスト (ping unloaded/loading) など



source: <https://wiki.edge.arista.com/index.php/Bufferbloat>

スピードテストと品質計測

- “スピードテスト = 品質計測” ではない

市場の成熟と品質意識

- 自動車にもかつては排気量神話があった
 - 接客サービスの品質意識もそろそろ変わるべき時期なのでは



わが国最高の貴賓用超豪華リムジン

ニッサン プリンス ロイヤル

O.H.C.V型 8 気筒6400cc最高速度160km/h

わが国で唯一の貴賓用リムジンとして、日産自動車が総力を結集、技術の粋を傾けた特別仕様車です。V型 8 気筒6400ccエンジンをはじめ、全長 6 m を越す気品にみちた車体に至るまで、ロイヤルならではの壮厳さがあふれています。

インターネットのサービス品質とは？

- 一般に「品質」は総合的にニーズを満たす度合い
 - 工学的には個別の性能指標やその他の要素を組み合わせで表す
 - 性能指標は特定の性能に関する評価尺度
 - パケット通信の性能指標：スループット、遅延、揺らぎ、パケットロス率など
- インターネットサービス：共有による影響
 - 他のユーザが使うとその影響を受け性能が低下する
 - 一般の工学では性能評価は回りからの影響を排除して行う
 - 定量化するには、異なる時間や場所での測定を沢山集めて統計的に処理する必要
 - 技術者が知りたい統計データは、かならずしもユーザが知りたいものではない

品質について技術者と利用者のギャップ

- 技術者視点
 - スループット、遅延、ばらつきなどの性能指標を組み合わせて定量化したもの
- 利用者視点
 - 宣伝どおり、期待どおりの性能か？
 - 使いたい時に快適に使えること
 - （マニア視点の品質は技術よりなので省く）
- 技術者視点の性能指標は利用者が求めるものではない

転送性能の二面性

- 転送性能：大きなファイルのダウンロードがいかに早く終わるかという性能
 - 実際のアプリの動作から乖離してきている
 - 回線容量に近いスループットが得られる期待
 - 外的要因による一時的な速度低下と設備不足による恒常的な速度低下
- 安定性向上：間接的には帯域の余裕を示し、余裕はシステムの安定性に繋がる
 - 余剰帯域が一時的なトラフィック変動を吸収
 - パケット交換網の安定動作に欠かせない
 - ビデオ会議や動画視聴などのアプリの安定動作に寄与
 - 必要以上に余裕を増やしてあまり効果はない
- 一般の人は性能を、通信技術者は品質安定を想定

ブロードバンド品質計測の取り組み

- 第三者機関による公的計測：消費者がサービスを選ぶ際の指標
- 2009 イギリスが最初の取り組み UK broadband report
- 2010 アメリカ FCC Measuring Broadband America
 - オープンデータ：産官学での本格的な取り組みを開始
 - <https://www.measuringbroadbandamerica.com/test-methodology>
- SamKnows社：FCC MBAプラットフォームを提供
 - イギリス、オーストラリア、ニュージーランド、香港、サウジアラビアなどからも受注
 - <https://samknows.com/solutions/regulator-solutions>
- 日本：2015 総務省がモバイル向けに計測ガイドラインを策定、主要キャリアが自主的に計測・公表
 - 2020 固定向けを検討する委員会ができたがあまり進捗はない
- 最近は専門家は遅延を重視し始めている
 - 例えば RFC 9318 - IAB Workshop Report: Measuring Network Quality for End-Users

最近の話題：FCCのラベリング

- FCC Broadband Consumer Labels
 - 現在策定プロセス最終段階
 - ISPのブローバンドサービス表示標準化
 - 価格や標準的なスピードや遅延の表示義務づけ
 - 消費者がチャレンジするプロセスも規定
- アメリカは国土が広いのでカバレッジ問題が深刻
 - ユニバーサルサービスとしてデジタルデバイド解消に大規模な予算割当

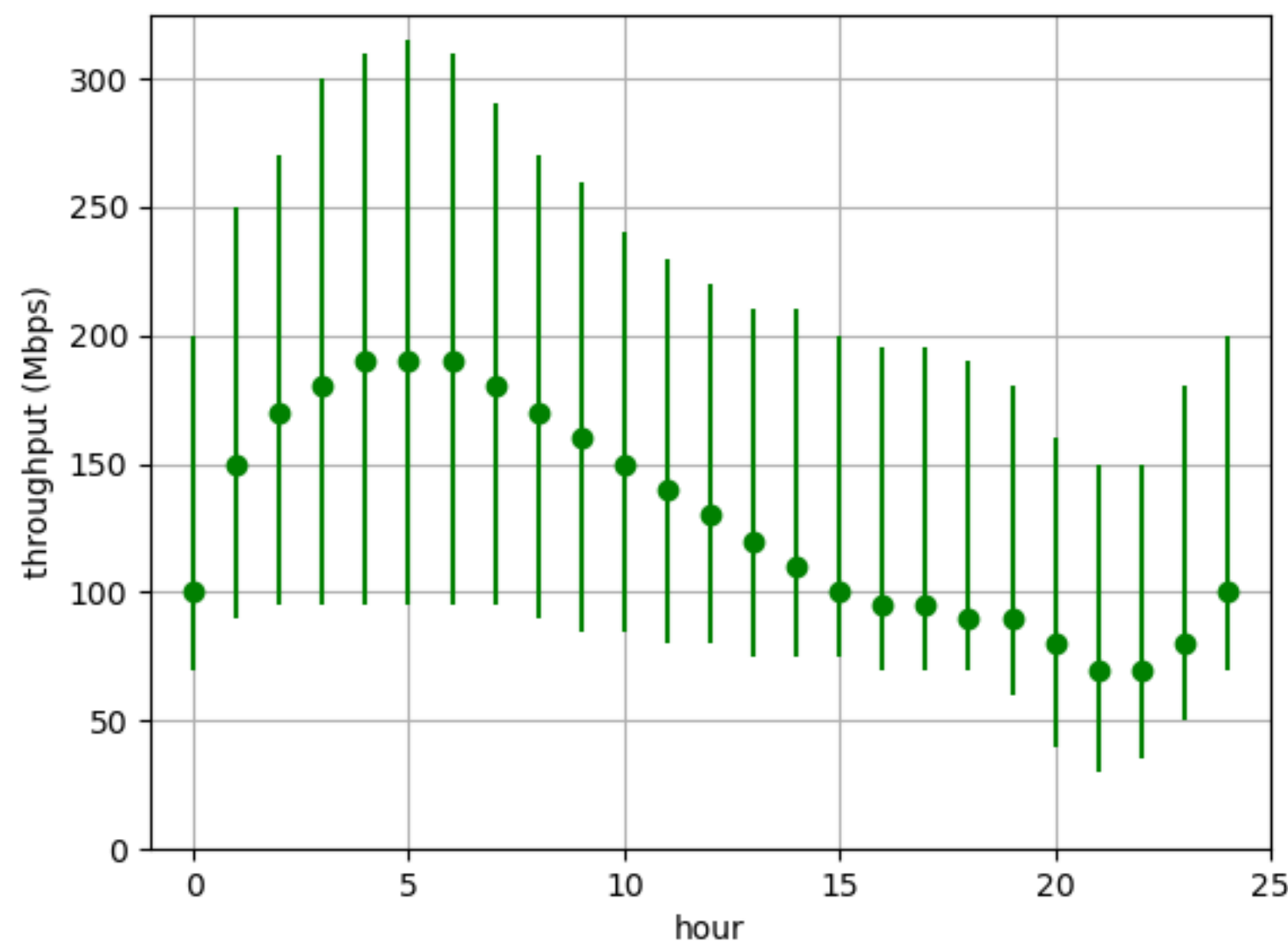
Broadband Facts	
Provider Name	
Service Plan Name and/or Speed Tier	
Fixed or Mobile Broadband Consumer Disclosure	
Monthly Price	[\$]
This Monthly Price [is/is not] an introductory rate. [if introductory rate is applicable, identify length of introductory period and the rate that will apply after introductory period concludes]	
This Monthly Price [does not] require[s] a [x year/x month] contract. [only required if applicable; if so, provide link to terms of contract]	
Additional Charges & Terms	
Provider Monthly Fees	[Itemize each fee] [\$]
One-time Fees at the Time of Purchase	[Itemize each fee] [\$]
Early Termination Fee	[\$]
Government Taxes	Varies by Location
Discounts & Bundles	
Click Here for available billing discounts and pricing options for broadband service bundled with other services like video, phone, and wireless service, and use of your own equipment like modems and routers. [Any links to such discounts and pricing options on the provider's website must be provided in this section.]	
Affordable Connectivity Program (ACP)	
The ACP is a government program to help lower the monthly cost of internet service. To learn more about the ACP, including to find out whether you qualify, visit affordableconnectivity.gov .	
Participates in the ACP	[Yes/No]
Speeds Provided with Plan	
Typical Download Speed	[] Mbps
Typical Upload Speed	[] Mbps
Typical Latency	[] ms
Data Included with Monthly Price	[] GB
Charges for Additional Data Usage	[\$/GB]
Network Management	Read our Policy
Privacy	Read our Policy
Customer Support	
Contact Us: example.com/support / (555) 555-5555	
Learn more about the terms used on this label by visiting the Federal Communications Commission's Consumer Resource Center.	
fcc.gov/consumer	
[Unique Plan Identifier Ex. F0005937974123ABC456EMC789]	

スピードテストの心理

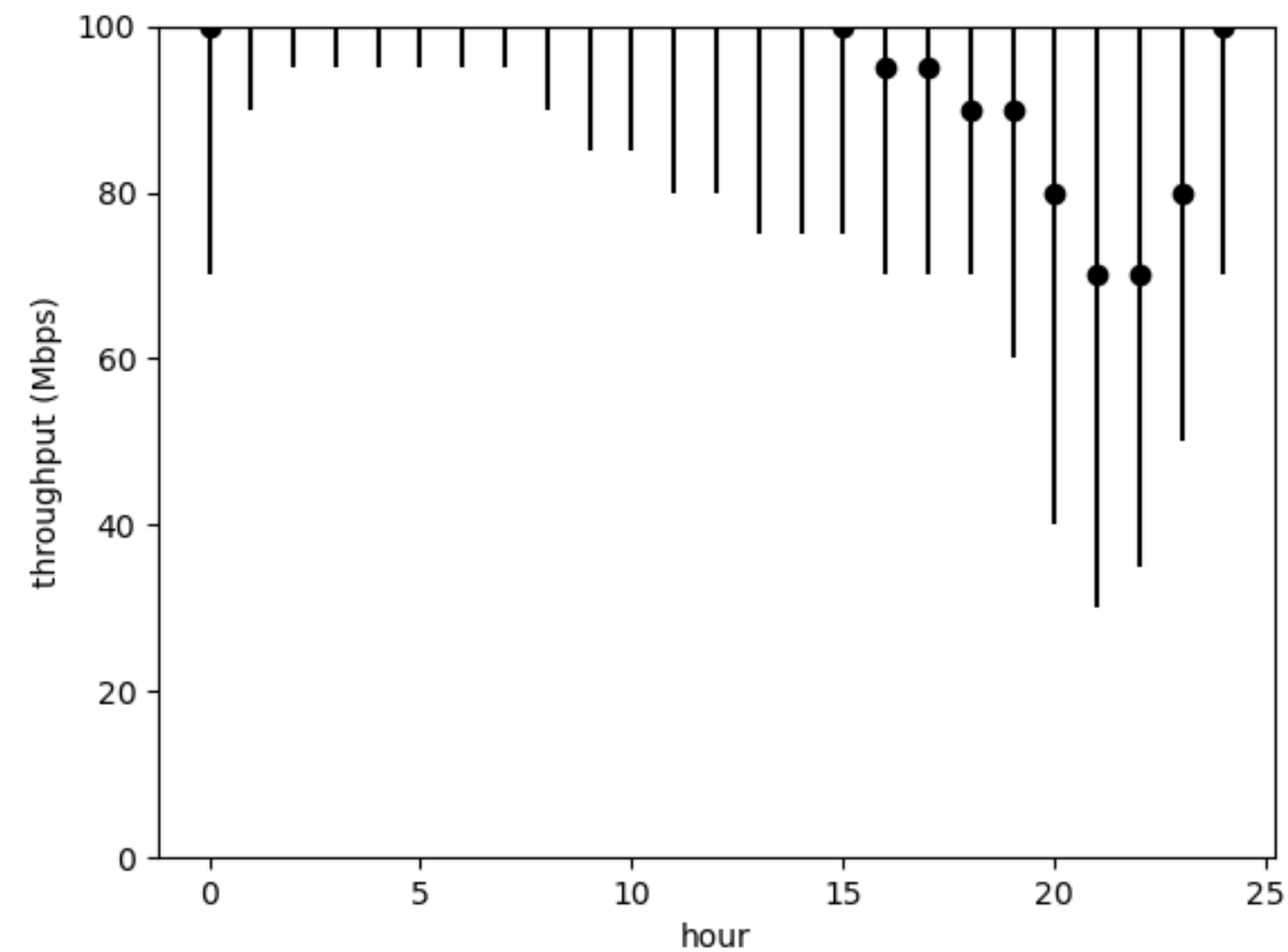
- 大きい値を求めてしまう
 - 最大値に着目しがち
 - 他の指標を示しても注目されない
- 最大速度1Gbpsと宣伝するサービスは近い値が出る期待
- 600Mbpsと800Mbpsのサービスがあれば後者を選びたくなる
- ユーザも事業者も、底上げよりもトップ争いに誘導される

可視化の方法でも印象は変わる

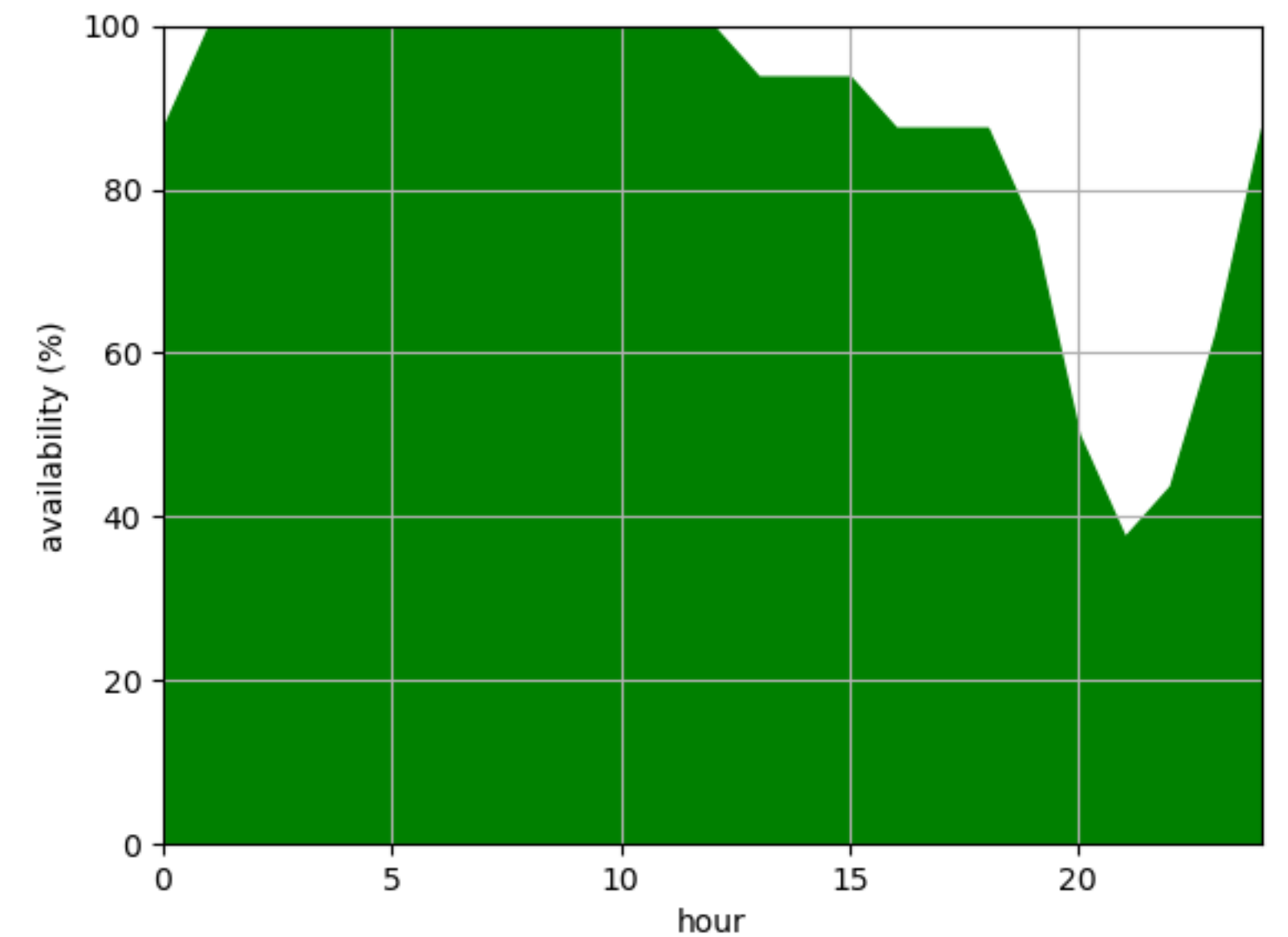
- スピードテスト結果の可視化
 - 100Mbps以上出ているか？



通常の技術プロット：
最大値に注目しがち



100Mbps以下だけ表示



充足部分を表示

ブロードバンドサービス全体の 品質向上に向けて

- 消費者のためになる情報とは？
 - ユーザが知りたいのはスピードテストの結果か？
 - サービスを選ぶ際に価格やサービス内容を比較したい？
 - サービスの品質に不満な場合、品質を示す手段が必要
- 広告の問題：回線容量と価格しか訴えるものがない？
- 健全なサービス競争を促すために
 - 長期的にサービス全体の品質向上に繋げるには？

ひとつの方向性（提案）

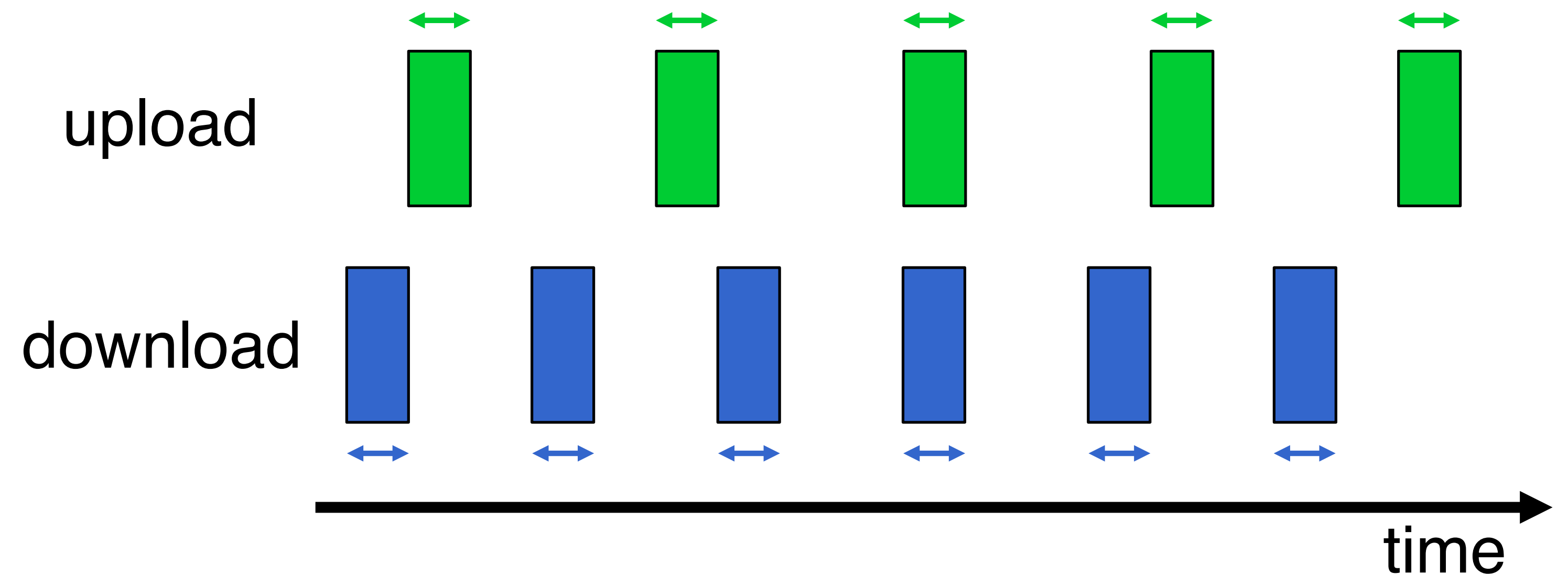
- ブロードバンド品質の概念をより成熟させる啓蒙活動
 - スピード性能ではなく安定した性能を重視する
- 適切なツールの提案（議論を進めるために叩き台が必要）
 - 主な目的は基準品質以下のサービスを特定し改善を促すこと
 - 標準サービスレベルの達成率を品質指標に用いる
 - （あくまで一例：他の指標も実装方法も色々あり得る）

目的適合性テスト(fit for purpose test)

- 利用目的に適合しているかを評価
 - ほとんどの家庭の利用で十分な性能を標準サービスレベルとして定義
 - 合否判定のテストを実施し、達成率（99.9%など）を示す
 - （誤用に注意しつつ技術者むけの情報も提供する）
- 標準サービスレベル
 - 固定ブロードバンドの利用目的はさまざま
 - コロナ禍での変化：家族がそれぞれ同時にビデオ会議やビデオストーリーリング
 - 現状で一般利用に十分な性能は定義できそう（必要に応じて見直すもの）
 - ビデオ会議 2本とビデオストーリーミング 2本を同時にできる程度
 - ダウンロード、アップロード、遅延、揺らぎの総合評価

適合性テストのプロトタイプ

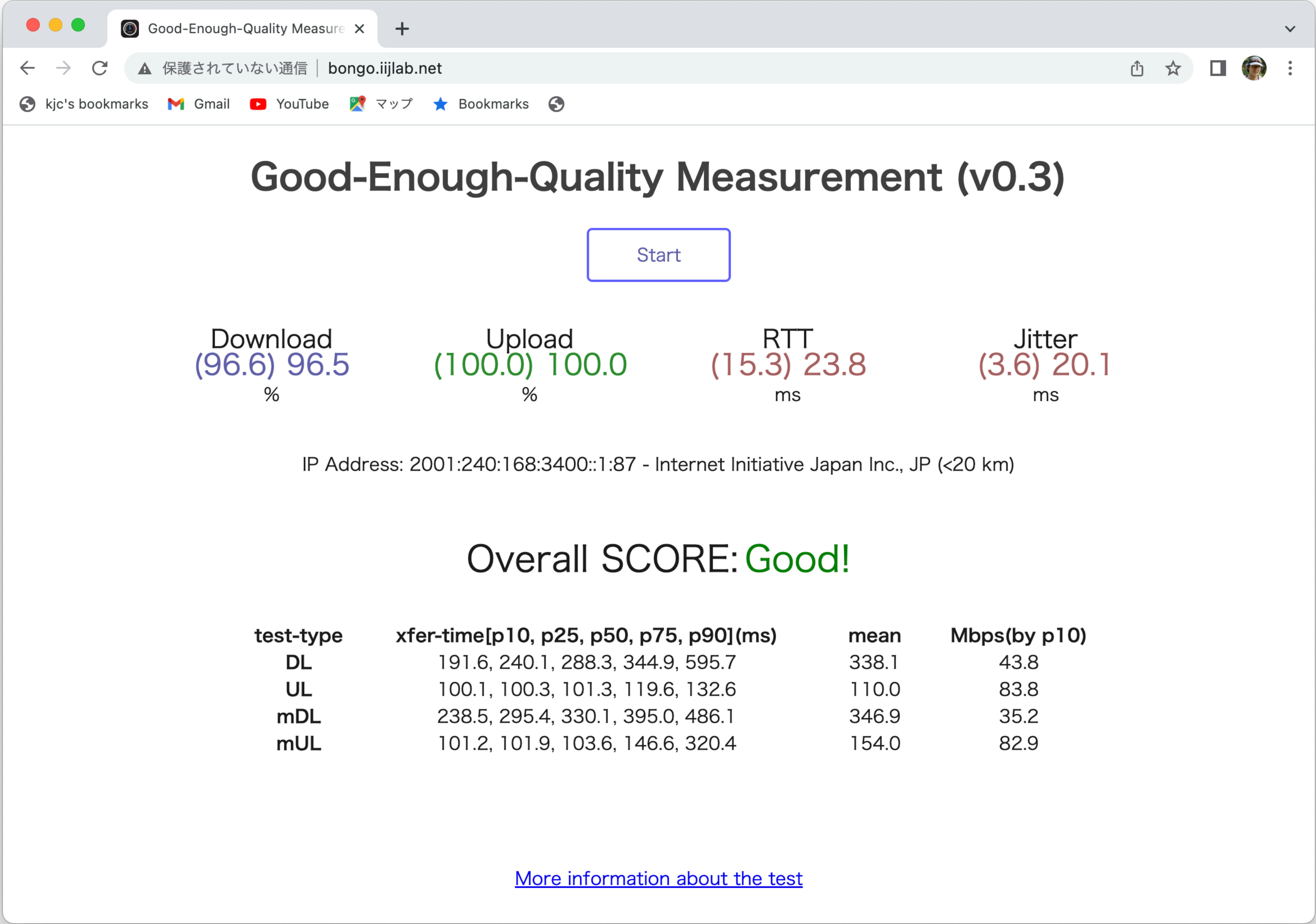
- ダウンロードとアップロードを並列実行
 - DL: 28Mbps (1MB/300ms)
 - UL: 21Mbps (1MB/400ms)
- 各ファイルの転送時間が安定しているかを評価
 - ダウンロード、アップロード、遅延、揺らぎの総合評価



プロトタイプ実装

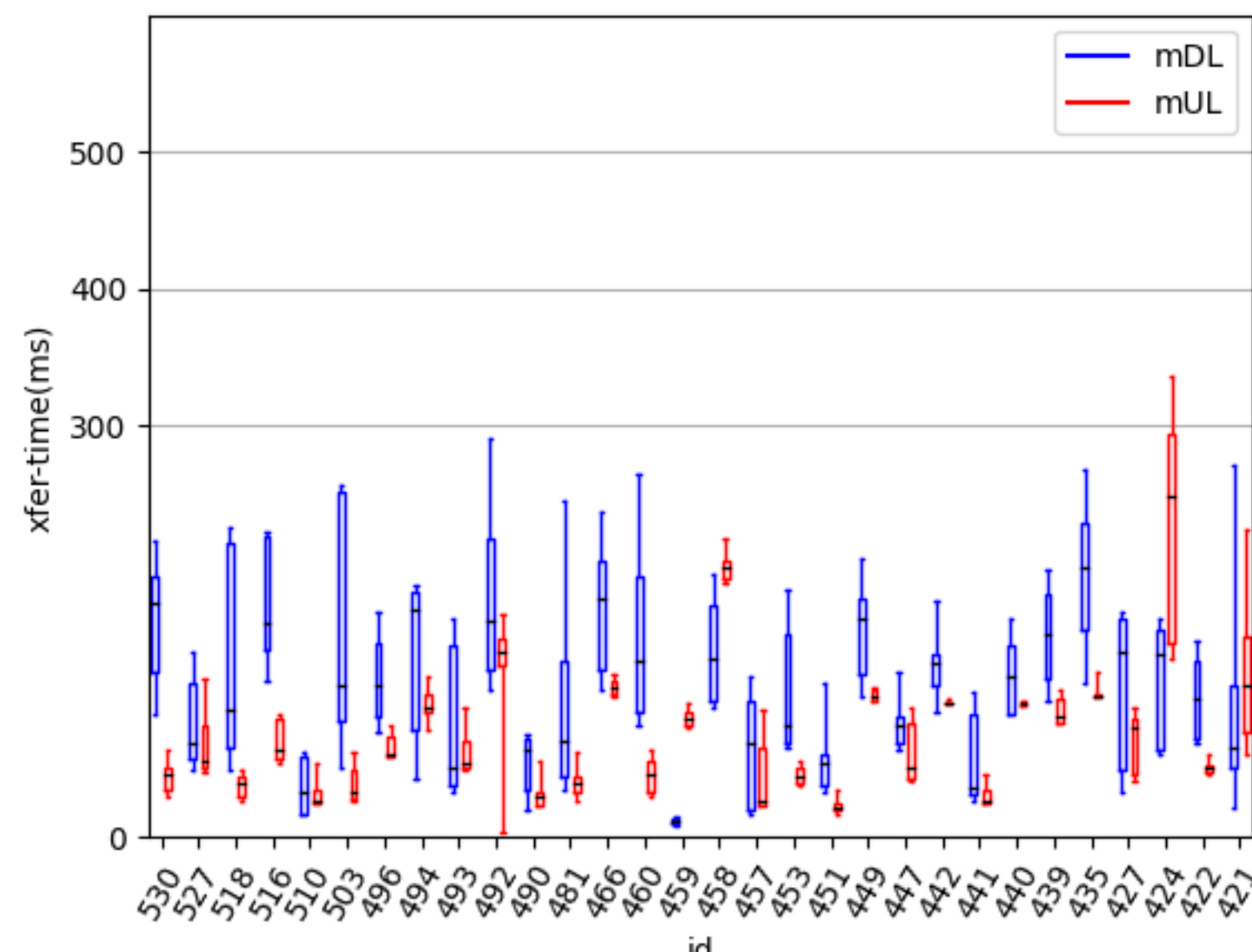
- オープンソースのlibrespeedを改造
 - HTTPで1MBの転送時間を計測
 - 転送間隔: 300ms for DL (28Mbps), 400ms for UL (21Mbps)
 - 転送時間の分布を記録 [p10, p25, p50, p75, p90]
 - 3フェーズ: DL, UL, mixed DL/UL
- 結果のスコア（固定接続を想定、モバイルにはハードル高すぎ）
 - Excellent: mDL_p90 < 300ms and mUL_p90 < 400ms
 - Good: mDL_p75 < 500ms and mUL_p75 < 500ms
 - Fair: mDL_throughput > 10Mbps and mUL_throughput > 5Mbps
 - Poor: otherwise

プロトタイプUI

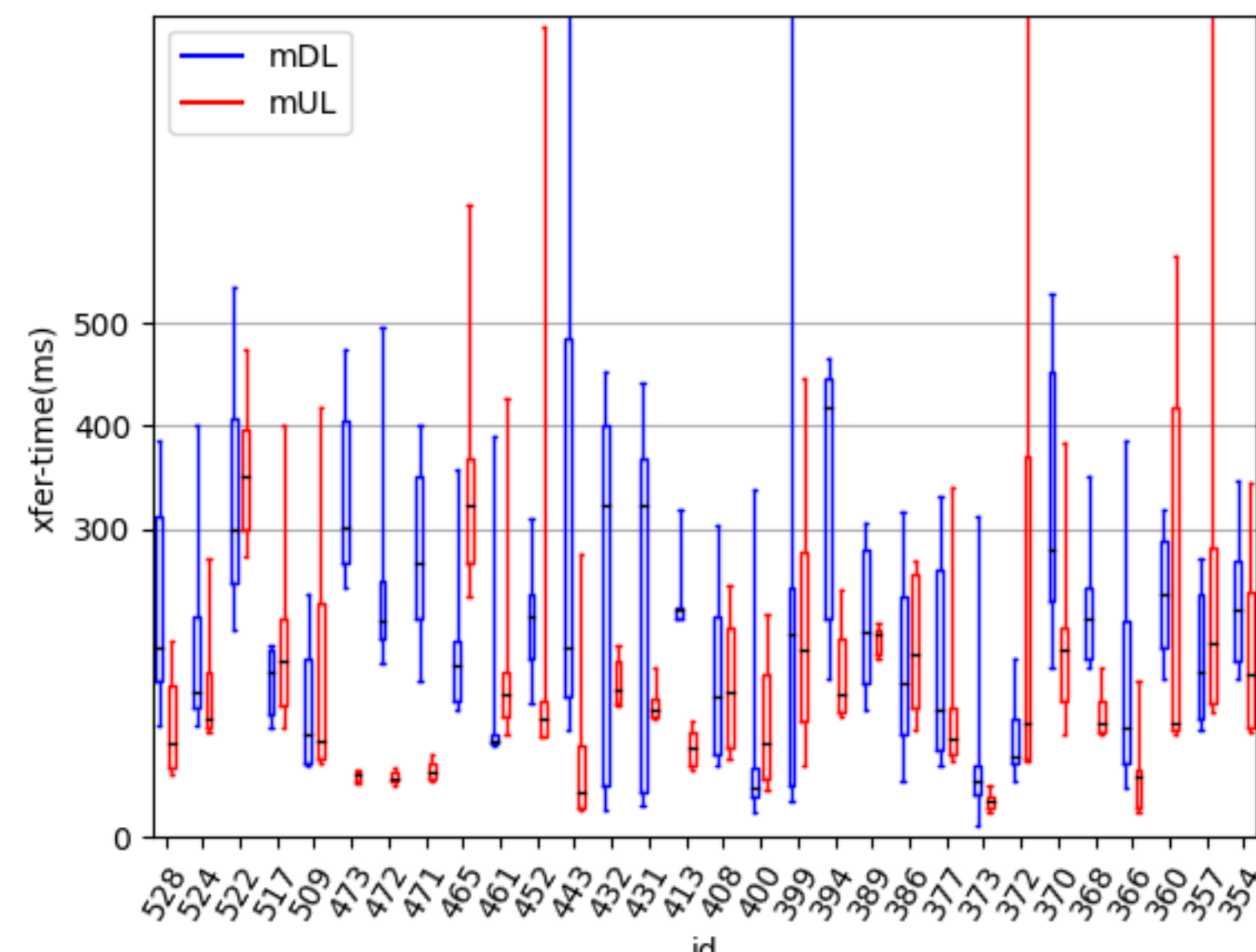


フィールドテスト (2022.10)

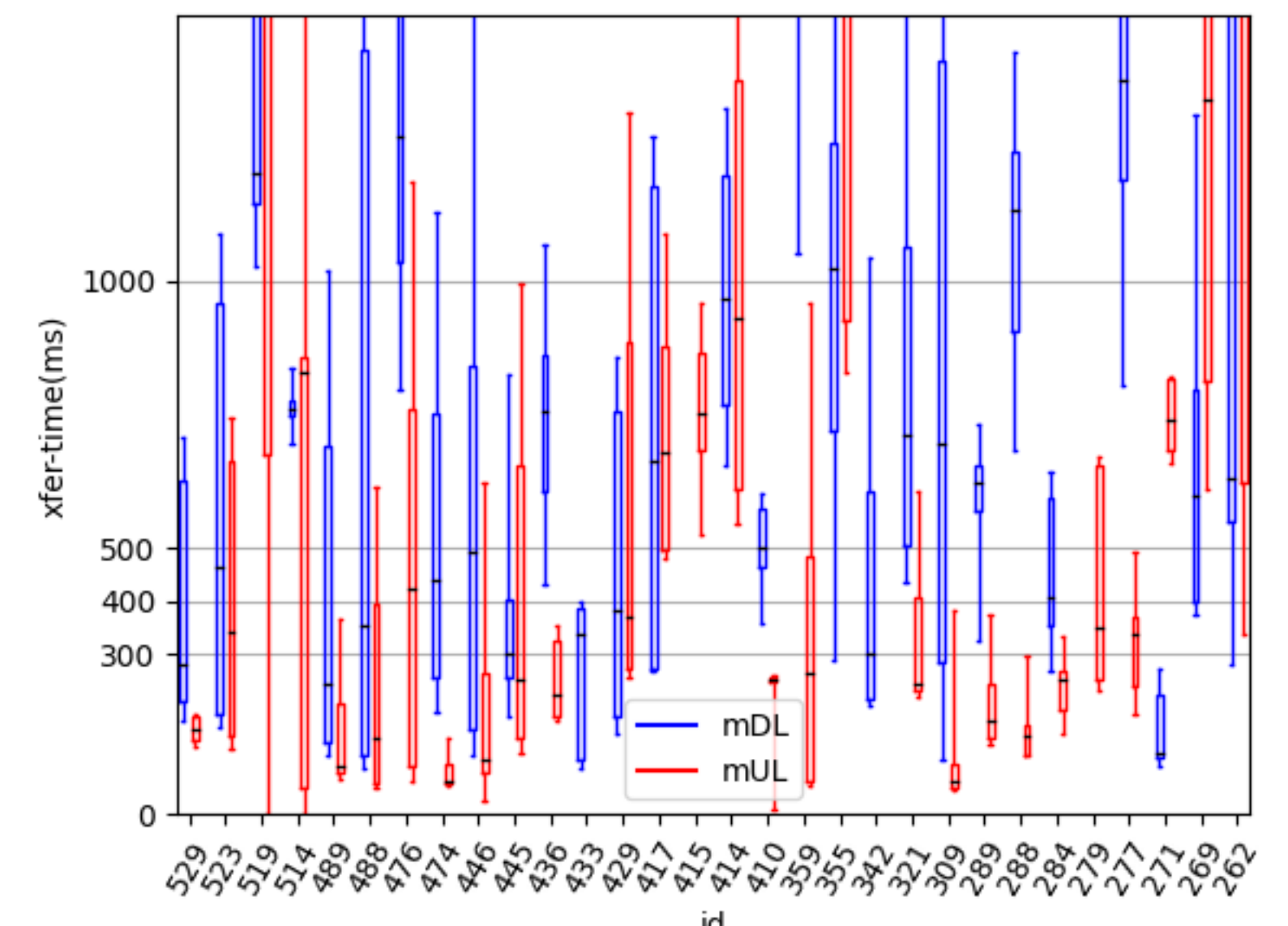
- IIJとWIDEのメンバーにテストを呼びかけ、約500回の計測
- 各ユニークIPアドレスの最新値を100個抽出してプロット
 - Excellent:53, Good:11, Fair:20, Poor:16
- [p10, p25, p50, p75, p90]をボックスプロット表示



Excellent



Good



Fair

現在の開発状況

- 実証実験の結果：期待できるのではないか
 - サービスの安定度を大まかに捉えることはできている
 - Wi-FiやVPN経由は（予想どおり）ばらつきが大きい
 - 一般むけの簡素な情報と技術者が必要な詳細情報のバランス
- 次のステップ
 - 継続して自動計測するCLIツール
 - ベース：CLI-librespeed (go-lang)
 - 一般向けには配布方法が課題
 - 結果の表示方法の工夫
 - 認知度向上へ向けた活動

まとめ

- スピードテストの仕組みを解説
- スピードテストはかならずしも品質計測には適さない
 - 接続サービスの品質改善につながらないのではないか
- 接続サービスの品質の概念をより成熟させたい
- スピードテストの代わりになるツールが必要
 - 目的適合性テストを提案しプロトタイプを作って実験実施
- 方向性についていろいろな立場のひとと議論したい