

IAB Workshop Report: Measuring Network Quality for End-Users

Отчёт семинара IAB по измерению качества сети для конечных пользователей

Аннотация

Семинар Measuring Network Quality for End-Users был проведён IAB в виртуальном формате 14-16 сентября 2021 г. В этом документе приведены итоги семинара, рассмотренные вопросы и некоторые предварительные выводы, принятые в конце семинара.

Документ является отчётом о работе семинара. Приведённые здесь взгляды и позиции принадлежат участникам семинара и могут не отражать позиции IAB.

Статус документа

Документ не относится к категории Internet Standards Track и публикуется лишь для информации.

Документ является результатом работы Совета по архитектуре Internet (Internet Architecture Board или IAB) и содержит сведения, которые IAB считает нужным сохранить. Документ представляет согласованный взгляд IAB. Документ был одобрен для публикации IAB и не является кандидатом в Internet Standard, см раздел 2 RFC 7841.

Информацию о текущем статусе документа, ошибках и способах обратной связи можно найти по ссылке <https://www.rfc-editor.org/info/rfc9318>.

Авторские права

Авторские права (Copyright (c) 2022) принадлежат IETF Trust и лицам, указанным в качестве авторов документа. Все права защищены.

К документу применимы права и ограничения, указанные в BCP 78 и IETF Trust Legal Provisions и относящиеся к документам IETF (<http://trustee.ietf.org/license-info>), на момент публикации данного документа. Прочтите упомянутые документы внимательно. Фрагменты программного кода, включённые в этот документ, распространяются в соответствии с упрощённой лицензией BSD, как указано в параграфе 4.e документа IETF Trust Legal Provisions, без каких-либо гарантий (как указано в Simplified BSD License).

Оглавление

1. Введение.....	2
1.1. Пространство проблем.....	2
2. Программа семинара.....	2
3. Представленные документы.....	3
4. Темы и дискуссии семинара.....	4
4.1. Введение и обзор.....	4
4.1.1. Основные моменты из доклада Vint Cerf.....	4
4.1.2. Вступительное обсуждение.....	4
4.1.3. Ключевые точки.....	5
4.2. Показатели.....	5
4.2.1. Базовые показатели производительности.....	5
4.2.2. Показатели применимости.....	6
4.2.3. Показатели пропускной способности.....	7
4.2.4. Показатели задержки.....	7
4.2.5. Примеры измерений.....	7
4.2.6. Ключевые точки показателей.....	8
4.3. Межуровневые вопросы.....	8
4.3.1. Разделение ответственности.....	9
4.3.2. Вопросы безопасности и приватности.....	9
4.3.3. Вопросы измерения показателей.....	9
4.3.4. Пути улучшения межуровневой наблюдаемости.....	9
4.3.5. Взаимодействие оборудования с транспортными протоколами.....	9
4.3.6. Ключевые точки межуровневых измерений.....	10
4.4. Объединительная часть.....	10
4.4.1. Вопросы измерений и показателей.....	10
4.4.2. Представление показателей конечного пользователя.....	10
4.4.3. Ключевые точки объединения.....	11
5. Выводы.....	11
5.1. Общие заявления.....	11
5.2. Конкретные заявления о протоколах и методах.....	11
5.3. Заявления о проблемах и ответственности.....	11
5.4. Утверждения, по которым не достигнуто согласия.....	12
6. Последующая работа.....	12
7. Взаимодействие с IANA.....	12

8. Вопросы безопасности.....	12
9. Литература.....	12
Приложение А. Программный комитет.....	14
Приложение В. Руководители семинара.....	14
Приложение С. Участники семинара.....	14
Члены IAB на момент принятия документа для публикации.....	14
Благодарности.....	14
Участники работы.....	14
Адреса авторов.....	14

1. Введение

IAB время от времени организует семинары по рассмотрению долгосрочной стратегии и проблем Internet, а также по определению направления развития архитектуры Internet. Такое долгосрочное планирование IAB дополняет текущую инженерную работу IETF¹.

Семинар по измерению качества сети для конечных пользователей (Measuring Network Quality for End-Users) [WORKSHOP] проводился в виртуальном режиме 14-16 сентября 2021 г. В этом отчёте приведены итоги семинара, обсуждавшиеся темы и некоторые предварительные выводы, сделанные в конце семинара.

1.1. Пространство проблем

Internet в 2021 существенно отличался от того, что было 10 назад. Сегодня это важная часть жизни каждого селовека. Люди используют Internet в своей общественной жизни, для повседневной работы и покупок, для того, чтобы быть в курсе важных событий. Рост числа людей с доступом к гигабитным соединениям было трудно представить десять лет назад. Благодаря улучшениям в части безопасности, люди стали доверять Internet в финансовых операциях с банками, покупке товаров, повседневных платежей по счетам.

В то же время некоторые аспекты взаимодействия с конечными пользователями не обрели существенного улучшения. Многие пользователи сталкиваются с задержками соединений на уровне 10-летней давности. Несмотря на значительное повышение надёжности в средах центров обработки данных (ЦОД), конечные пользователи часто сталкиваются с перебоями в обслуживании. Несмотря на алгоритмические улучшения в теории управления, по-прежнему обнаруживается, что задержки в очередях оборудования последней мили превышают совокупные транзитные задержки. Улучшения в транспорте, такие как QUIC, Multipath TCP, TCP Fast Open, в некоторых сетях ещё поддерживаются не полностью. Точно так же не получили широкого распространения усовершенствования в сфере безопасности и приватности конечных пользователей, такие как шифрование DNS для локальных распознавателей.

Одним из основных факторов отсутствия прогресса является распространённое мнение о том, что пропускная способность часто является единственным показателем качества соединений Internet. С учётом этого семинар Measuring Network Quality for End-Users был нацелен на обсуждение указанных ниже тем.

- Какова задержка для пользователей в типичных условиях работы?
- Насколько надёжна связь в долгосрочной перспективе?
- Позволяют ли сети использовать широкий спектр протоколов?
- Какие службы могут запускать клиенты сети?
- Какие типы соединения IPv4, NAT, IPv6 предлагаются, применяются ли межсетевые экраны?
- Какие механизмы защиты доступны для локальных служб, таких как DNS?
- В какой степени обеспечивается приватность, конфиденциальность, целостность и подлинность пользовательских коммуникаций?

Улучшение этих аспектов качества сети вероятно будет зависеть от измерения и осмысленного представления показателей всем заинтересованным сторонам, включая конечных пользователей. Такие измерения и предоставление правильных показателей позволят поставщикам услуг и операторам сетей сосредоточиться на опыте своих пользователей и одновременно даст пользователям возможность выбрать провайдера (Internet Service Provider или ISP) в соответствии со своими потребностями.

- Каковы фундаментальные свойства сети, способствующие качественному взаимодействию с пользователем?
- Какие показатели количественно определяют эти свойства и как их собрать на практике?
- Каковы наилучшие методы интерпретации этих показателей и их включения в процесс принятия решений?
- Как лучше всего передать эти свойства поставщикам услуг и операторам сетей?
- Как эти показатели представить пользователям осмысленным способом?

2. Программа семинара

Семинар Measuring Network Quality for End-Users был разделен по нескольким темам (см. разделы 4 и 5):

- вводные обзоры и доклад Vint Cerf;
- показатели;
- межуровневые вопросы;
- объединительная часть;
- выводы.

¹Internet Engineering Task Force - комиссия по решению инженерных задач Internet.

3. Представленные документы

Ниже указаны документы, представленные участникам семинара. На странице семинара [WORKSHOP] содержится архив статей, презентаций и видеозаписей.

- Ahmed Aldabbagh. "Regulatory perspective on measuring network quality for end users" [Aldabbagh2021].
- Al Morton. "Dream-Pipe or Pipe-Dream: What Do Users Want (and how can we assure it)?" [Morton2021].
- Alexander Kozlov. "The 2021 National Internet Segment Reliability Research".
- Anna Brunstrom. "Measuring network quality - the MONROE experience".
- Bob Briscoe, Greg White, Vidhi Goel, and Koen De Schepper. "A Single Common Metric to Characterize Varying Packet Delay" [Briscoe2021].
- Brandon Schlinker. "Internet Performance from Facebook's Edge" [Schlinker2019].
- Christoph Paasch, Kristen McIntyre, Randall Meyer, Stuart Cheshire, and Omer Shapira. "An end-user approach to the Internet Score" [McIntyre2021].
- Christoph Paasch, Randall Meyer, Stuart Cheshire, and Omer Shapira. "Responsiveness under Working Conditions" [Paasch2021].
- Dave Reed and Levi Perigo. "Measuring ISP Performance in Broadband America: A Study of Latency Under Load" [Reed2021].
- Eve M. Schooler and Rick Taylor. "Non-traditional Network Metrics".
- Gino Dion. "Focusing on latency, not throughput, to provide better internet experience and network quality" [Dion2021].
- Gregory Mirsky, Xiao Min, Gyan Mishra, and Liuyan Han. "The error performance metric in a packet-switched network" [Mirsky2021].
- Jana Iyengar. "The Internet Exists In Its Use" [Iyengar2021].
- Jari Arkko and Mirja Kuehlewind. "Observability is needed to improve network quality" [Arkko2021].
- Joachim Fabini. "Network Quality from an End User Perspective" [Fabini2021].
- Jonathan Foulkes. "Metrics helpful in assessing Internet Quality" [Foulkes2021].
- Kalevi Kilkki and Benajamin Finley. "In Search of Lost QoS" [Kilkki2021].
- Karthik Sundaresan, Greg White, and Steve Glennon. "Latency Measurement: What is latency and how do we measure it?"
- Keith Winstein. "Five Observations on Measuring Network Quality for Users of Real-Time Media Applications".
- Ken Kerpez, Jinous Shafiei, John Cioffi, Pete Chow, and Djamel Bousaber. "Wi-Fi and Broadband Data" [Kerpez2021].
- Kenjiro Cho. "Access Network Quality as Fitness for Purpose".
- Koen De Schepper, Olivier Tilman, and Gino Dion. "Challenges and opportunities of hardware support for Low Queuing Latency without Packet Loss" [DeSchepper2021].
- Kyle MacMillian and Nick Feamster. "Beyond Speed Test: Measuring Latency Under Load Across Different Speed Tiers" [MacMillian2021].
- Lucas Pardue and Sreeni Tellakula. "Lower-layer performance not indicative of upper-layer success" [Pardue2021].
- Matt Mathis. "Preliminary Longitudinal Study of Internet Responsiveness" [Mathis2021].
- Michael Welzl. "A Case for Long-Term Statistics" [Welzl2021].
- Mikhail Liubogoshchev. "Cross-layer Cooperation for Better Network Service" [Liubogoshchev2021].
- Mingrui Zhang, Vidhi Goel, and Lisong Xu. "User-Perceived Latency to Measure CCAs" [Zhang2021].
- Neil Davies and Peter Thompson. "Measuring Network Impact on Application Outcomes Using Quality Attenuation" [Davies2021].
- Olivier Bonaventure and Francois Michel. "Packet delivery time as a tie-breaker for assessing Wi-Fi access points" [Michel2021].
- Pedro Casas. "10 Years of Internet-QoE Measurements. Video, Cloud, Conferencing, Web and Apps. What do we Need from the Network Side?" [Casas2021].
- Praveen Balasubramanian. "Transport Layer Statistics for Network Quality" [Balasubramanian2021].
- Rajat Ghai. "Using TCP Connect Latency for measuring CX and Network Optimization" [Ghai2021].
- Robin Marx and Joris Herbots. "Merge Those Metrics: Towards Holistic (Protocol) Logging" [Marx2021].
- Sandor Laki, Szilveszter Nadas, Balazs Varga, and Luis M. Contreras. "Incentive-Based Traffic Management and QoS Measurements" [Laki2021].
- Satadal Sengupta, Hyojoon Kim, and Jennifer Rexford. "Fine-Grained RTT Monitoring Inside the Network" [Sengupta2021].
- Stuart Cheshire. "The Internet is a Shared Network" [Cheshire2021].

- Toerless Eckert and Alex Clemm. "network-quality-eckert-clemm-00.4".
- Vijay Sivaraman, Sharat Madanapalli, and Himel Kumar. "Measuring Network Experience Meaningfully, Accurately, and Scalably" [Sivaraman2021].
- Yaakov (J) Stein. "The Futility of QoS" [Stein2021].

4. Темы и дискуссии семинара

Программа трехдневного семинара была разбита на 4 части, каждая из которых сыграла свою роль в организации дискуссий. Семинар начался с вводных презентаций и представления пространства проблем (4.1. Введение и обзор), затем обсуждались показатели (4.2. Показатели), межуровневые вопросы (4.3. Межуровневые вопросы) и была проведена общая дискуссия (4.4. Объединительная часть). По завершении этих разделов было проведено обсуждение и сделаны выводы, принятые участниками семинара (5. Выводы).

4.1. Введение и обзор

Семинар начался с широкого обсуждения качества обслуживания (Quality of Service или QoS) и работы (Quality of Experience или QoE) пользователей в современной сети Internet. Цель вводных бесед состояла в подготовке участников семинара к описанию пространства проблем, имеющихся решений и их ограничений.

В презентациях были показаны имеющиеся измерения QoS и QoE, а также их эффективность. Обсуждалось взаимодействие между пользователями в сети, а также между разными уровнями стека протоколов OSI. Vint Cerf выступил с основным докладом, посвященным истории и важности темы.

4.1.1. Основные моменты из доклада Vint Cerf

Мы можем работать в сетевом пространстве, разительно отличающемся по параметрам от пространства 30-летней давности. Это различие оправдывает не только важность одних показателей по сравнению с другими, но и пересмотр метафоры целиком.

Пришло время заняться экспертам не только настройкой TCP, но и исследованием более поздних протоколов, таких как QUIC. Важно свободное рассмотрение альтернатив TCP. Протокол TCP - это не плюшевый мишка и не следует бояться заменить его транспортным уровнем с лучшими свойствами, который будет полезней для пользователей.

Рассмотрим упражнение для определения желаемых свойств. Когда мы рассматриваем пространства параметров, можно отделить «желаемые» свойства от «фундаментальных», например, малую задержку. Примером от Агентства исследовательских проектов (Advanced Research Projects Agency или ARPA) служит желание знать, где находится ракета сейчас, а не где она была. Понимание управляет созданием и выбором конкретных параметров в пространстве проектирования.

Когда значения параметров, таких как связность, сильно меняются, появляются новые решения. Одним из примечательных примеров является межпланетный протокол, где ring больше не даёт чего-либо полезного. При рассмотрении чувствительности (responsiveness) не следует игнорировать связность.

К сожалению, совместимость с имеющимися решениями болезненна. Работа по проектированию IPv6 для перехода с IPv4 могла бы пройти лучше, если бы не требовалась совместимость с имеющимся протоколом. Для IPv6 уже слишком поздно, но ещё можно учесть этот аспект при решении возможных в будущем проблем.

IPv6 ещё не везде реализован достаточно полно. С момента начала работы в 1996 г. пройден длинный путь, но цель ещё не достигнута. В 1996 г. казалось, что внедрить IPv6 достаточно просто, но практика показала иное. В 1996 г. начался бум dot-com, когда быстро тратились большие деньги и момент не был уловлен вовремя, а рынок расширился экспоненциально. Это следует считать предостережением.

Ещё одно замечание относится к производительности пересылки через множество этапов (hop) в Internet. Не видно сквозных показателей для её оценки, поскольку довольно сложно измерить производительность сквозь разные сети и бизнес-границы. При разработке новых протоколов следует задавать вопрос о возможности работы протокола через множество узлов пересылки.

Сети с множеством пересылок (multi-hop) постепенно заменяются огромными плоскими сетями с достаточной связностью между операторами, где системы разделяет 1 или 2 этапа пересылки (hop), например, Google, Facebook, Amazon. Фундаментальная архитектура Internet меняется.

4.1.2. Вступительное обсуждение

Internet является общей (shared) сетью на основе протоколов IP, использующей коммутацию пакетов для соединения множества автономных сетей. Отход Internet от технологии коммутации каналов (circuit-switching) позволил выйти за пределы любого другого устройства сети. С другой стороны, отсутствие внутрисетевого регулирования осложняет обеспечение наилучшей работы для каждого пользователя.

Поскольку применение Internet продолжает расширяться, становится все труднее предсказать какие характеристики сети будут лучше для пользователей. Разные классы приложений, например, видеопотоки и телеконференции, могут влиять на работу пользователей сложным и трудно измеряемым образом. Загрузка Internet быстро меняется в течение каждого дня, недели и года, что дополнительно осложняет определение ключевых показателей, позволяющих надёжно прогнозировать взаимодействие с пользователями.

С помощью QoS пытались преодолеть эти трудности за счёт строгой приоритизации разных типов трафика. Однако показатели QoS не всегда коррелируют с пользовательской практикой. Полезность показателей QoS дополнительно ограничивается сложностями создания решений с желаемыми характеристиками QoS.

В QoE предпринимались попытки объединить психологические аспекты восприятия качества и создать статистические модели для оптимизации работы пользователей. Несмотря на значительные усилия по созданию моделей, подход QoE оказался полезным лишь для некоторых классов приложений. К сожалению, обобщение моделей оказалось затруднительным и вопрос о влиянии приложений друг на друга в одной сети остался нерешенным.

Ориентация отрасли на предоставление конечному пользователю большей пропускной способности (полосы) привела к значительным достижениям. Во многих местах по всему миру домашним пользователям ISP предоставляют гигабитную скорость. 10 лет назад такое сочли бы фантастикой. Однако рост пропускной способности был достигнут за счёт пренебрежения другим важным фактором - задержкой. В результате конечным пользователям, на работу которых значительная задержка оказывает негативное влияние, рекомендуют заменить своё оборудование на новое, которое вместо этого даёт большую пропускную способность. В [MacMillian2021] показано, что такое обновление может снизить задержку из-за коммерческих причин перевести клиента на использование более дорогих тарифных планов.

Поскольку в отрасли продолжается предоставление конечным пользователям большей пропускной способности без учёта факторов задержки, при разработке приложений начали применять методы сокрытия задержек и краткосрочных прерываний обслуживания. Например, производительность пользовательских web-браузеров тесно связана с содержимым локального кэша браузера. Хотя применяемые методы могут улучшить взаимодействие с пользователем при возможности использования лежалых (кэшированных) данных, это ещё сильнее отдаляет пользователя от фактических показателей.

За последние 10 лет усилия Dave Taht и сообщества раздувания буферов (bufferbloat) привели к существенному обновлению алгоритмов очередей для снижения задержки под нагрузкой по сравнению с простыми очередями FIFO. К сожалению, производители домашних маршрутизаторов ещё не реализовали эти алгоритмы в основном из соображений маркетинга и стоимости. Большинство производителей маршрутизаторов зависят от ускорителей SoC (System on a Chip) для создания продукции с желаемой пропускной способностью. Производители SoC выбирают простейшие алгоритмы и настойчивое (aggressive) агрегирование, считая, что микросхема с более высокой пропускной способностью будет иметь гарантированный спрос. Поскольку потребителям в основном предлагается выбор из высокоскоростных устройств, допущение о том, что рост пропускной способности повышает QoS, продолжает укрепляться.

Домашние маршрутизаторы - не единственное место, где можно получить преимущества от более чёткого указания приемлемой для пользователей производительности. Поскольку пользователи воспринимают Internet через «призму» приложений, важно призвать производителей приложений принять решения с меньшими задержками. К сожалению, скорость отклика измерить сложнее, чем пропускную способность. Многие приложения нашли наборы показателей, которые полезны для них, но не обобщены достаточно и не могут стать универсально применимыми. Кроме того, из-за высокой конкуренции в сфере приложений у производителей могут быть экономические причины не делиться своими полезными показателями.

4.1.3. Ключевые точки

1. Измерение пропускной способности необходимо, но само по себе недостаточно.
2. Во многих случаях пользователям Internet нужна не дополнительная, а лучшая пропускная способность, т. е. нужны иные улучшения связности.
3. Пользователи воспринимают качество своего соединения с Internet через приложения, на которые влияет комбинация факторов. Нет смысла показывать типичному пользователю весь спектр возможных причин низкой производительности с точки зрения их приложений.
4. Многие факторы, влияющие на работу пользователя, находятся за пределами сферы его контроля. Не очевидно, что предоставление таких сведений пользователю поможет тому понять состояние производительности его сети. Как правило, пользователи предпочитают простые, категоричные решения (например, хорошо, лучше, наилучший).
5. Рынок содержимого Internet отличается высокой конкуренцией и многие приложения разрабатывают свой «секретный соус».

4.2. Показатели

Во второй части программы семинара продолжалось обсуждение показателей, которые можно применять вместо или в дополнение к доступной пропускной способности. Несколько участников семинара представили подробные исследования методологии измерений.

4.2.1. Базовые показатели производительности

Полная потеря доступа в Internet, несомненно, является худшим случаем для пользователя. К сожалению, если перезагрузка домашнего маршрутизатора не приведёт к восстановлению связности, пользователь мало что может сделать, кроме обращения к своему поставщику услуг. Тем не менее систематический сбор показателей доступности на стороне клиента полезен и может помочь ISP пользователя в поиске и устранении проблемы, позволяя пользователю выбрать лучшего ISP. Доступность можно оценить напрямую, просто пытаясь соединиться со стороны клиента с интересующими удалёнными местами. Например, Ookla [Speedtest] использует большое число устройств Android для измерения доступности сети и сотовой связи по всему миру. Ookla собирает данные сотен миллионов точек в день и использует их для точной оценки доступности. Другой подход состоит в выводе доступности по частоте отказов и другим тестам. Например, в [FCC_MBA] and [FCC_MBA_methodology] используются тысячи маршрутизаторов с измерительными программами, разработанными [SamKnows]. Эти маршрутизаторы выполняют множество сетевых тестов и сообщают о доступности на основе успешности тестовых соединений.

Измерение полосы (capacity) может быть полезно для конечных пользователей, но оно ещё важнее для сервис-провайдеров и разработчиков приложений. Видеопотоки с высоким разрешением требуют существенно большей полосы, чем иные типы трафика. На момент проведения семинара видеопотоки составляли 90% всего трафика Internet и приносили 95% доходов от монетизации (подписка, сборы, реклама). В результате службы потокового видео, такие как Netflix, должны постоянно справляться с быстрым изменением доступной полосы. Способность измерения доступной полосы в реальном масштабе времени усиливается за счёт алгоритмов адаптивного сжатия (adaptive bitrate или ABR) для обеспечения наилучшего взаимодействия с пользователями. Измерение совокупной потребности в полосе позволяет ISP быть готовыми к скачкам трафика. Например, в новомодные каникулы глобальный спрос на полосу возрастает в 5-7 раз по отношению к иному времени. Для конечных пользователей знание потребности в полосе может помочь при выборе тарифного плана с учётом предполагаемого использования. Однако во многих

случаях у конечных пользователей имеется избыток полосы и повышение пропускной способности не нужно для них. Способность отличать пропускную способность от полезной пропускной способности (goodput) может быть полезна для определения перегрузки в сети.

При измерении качества сети задержка определяется временем прохождения пакета по сетевому пути от одного конца до другого. На момент написания отчёта пользователи во многих местах по миру имели доступ в Internet с достаточной пропускной способностью и доступностью. Для этих пользователей снижение задержки вместо повышения пропускной способности может вести к значительному росту QoE. Показателем задержки является время кругового обхода (round-trip time или RTT), обычно измеряемое в миллисекундах. Однако пользователи часто считают RTT неудобным показателем в отличие от других показателей производительности, поскольку высокое значение RTT указывает большую задержку, а пользователи обычно считают большие значения лучшими. Для решения этой проблемы в [Paasch2021] и [Mathis2021] предложен обратный показатель - число RTT в минуту (Round-trips Per Minute или RPM).

Важно различать задержку при бездействии (idle latency) и задержку в рабочих условиях. Первая измеряется при слабой загрузке сети и отражает наилучший вариант. Вторая измеряется при типичной загрузке сети. До недавнего времени типовые инструменты указывали загрузку при бездействии, что могло вводить в заблуждение. Например, данные, представленные на семинаре, показали, что при бездействии задержка может быть в 25 раз ниже, чем при типичной загрузке сети. Поэтому важно чётко различать представление этих значений задержки пользователю.

Измерения показывают, что быстрое изменение пропускной способности влияет на задержку. В [Foulkes2021] предпринята попытка количественной оценки частоты возникновения «нестабильности» сети (высокой задержки при очень низкой пропускной способности) при быстром изменении пропускной способности. Такие изменения могут быть вызваны отказами в инфраструктуре, но гораздо чаще связаны с событиями внутри сети, такими как изменение правил организации трафика или быстрые изменения перекрестного трафика (cross-traffic).

Представленные на семинаре данные показывают, что у 36% измеренных линий показатель пропускной способности меняется больше чем на 10% в течение дня и нескольких дней. Эти различия связаны с многими переменными, включая локальные методы подключения (Wi-Fi и Ethernet), конкурирующий трафик ЛВС, загрузки и конфигурация устройства, время суток, пропускная способность локального соединения или транспорта. Эти изменения факторов осложняют измерение пропускной способности с использованием лишь устройства конечного пользователя или иного устройства оконечной сети. Маршрутизатор сети, просматривающий агрегированный трафик от нескольких устройств, обеспечивает лучшую точку измерения пропускной способности. Такой тест может учитывать весь локальный трафик и выполнить независимое тестирование пропускной способности. Однако различные факторы все равно могут ограничивать точность такого теста. Для аккуратного измерения пропускной способности нужно множество выборов.

Поскольку пользователь воспринимает Internet через приложения, сопоставление изменений пропускной способности и задержки с качеством работы пользователя может быть затруднительным. Например, web-браузеры полагаются на кэшированные версии страниц для более быстрой загрузки и смягчения потерь связности. Кроме того, приложения социальных сетей часто полагаются на предварительную загрузку элементов «ленты». Эти методы делают базовые показатели сети менее информативными в части удовлетворения пользователей и требуют сбора данных непосредственно от приложений конечного пользователя.

Полезно отличать приложения, работающие с «фиксированным бюджетом задержки», от приложений, устойчивых к вариациям задержки. Облако игр служит примером приложения, которому нужен «фиксированный бюджет задержки», оных серверов, поскольку всплеск задержки может определять решение «выигрыш-проигрыш» для игрока. Компании, конкурирующие на прибыльном рынке облачных игр, вкладывают значительные средства в инфраструктуру, такие как создание ЦОД ближе к пользователям. Эти ЦОД подчёркивают экономическую выгоду от снижения числа всплесков задержки, превышающую связанную с этим расходы на развёртывание. С другой стороны, более стойкие к всплескам задержки приложения могут продолжать нормальную работу даже при коротких всплесках. Тем не менее, даже такие приложения могут выиграть от неизменно низкой задержки при изменении загрузки. Например, приложения «видео по запросу» (Video-on-Demand или VOD) могут достаточно хорошо работать при линейном потреблении видеопотока, но при переключении пользователем канала или прокутке вперёд у пользователя могут возникнуть проблемы, если задержка недостаточно мала.

По мере развития приложений все большее значение приобретают их внутренние показатели. Например, приложения VOD могут оценивать QoE по зависимым от приложения показателям, таким как способность видеопроигрывателя использовать максимально возможное разрешение, определять плавность или «замораживание» изображения и т. п. Разработчики приложений могут эффективно применять эти показатели для определения приоритета будущих работ. Все популярные видео-платформы (YouTube, Instagram, Netflix и пр.) имеют схемы сбора и анализа показателей VOD. Одним из примеров является модель Scuba, применяемая в Meta [Scuba].

К сожалению внутренние показатели приложений могут быть слишком сложными для сравнительных исследований. Во-первых, разные приложения часто используют различные показатели для измерения одних и тех же явлений. Например, приложение А может измерять гладкость видео по среднему времени повторной буферизации, а приложение В может полагаться на вероятность повторной буферизации в течение 1 секунды для тех же целей. Другая проблема с внутренними показателями приложений состоит в том, что VOD является важным источником прибыли для YouTube, Facebook, Netflix, что создаёт препятствия обмену внутренними данными. Ещё одна проблема связана с вопросами приватности, связанными с внутренними показателями приложений, которые точно описывают действия и предпочтения конкретных пользователей.

4.2.2. Показатели применимости

Доступность определяется просто возможностью отправки пакета и получения соответствующего отклика. Наивно считают, что измерить доступность проще всего, но это оказывается сложнее, если учесть, что для повторяющихся мгновенных измерений требуется фиксировать малейшие перебои. Также трудно определить основную причину недоступности - была ли отключена линия пользователя, что-то произошло в сети или это отказ службы, к которой пользователь обратился.

4.2.3. Показатели пропускной способности

Если пропускная способность сети не соответствует потребностям пользователей, это влияет на качество сети. Когда пропускная способность соответствует запросам, её увеличение не ведёт к росту качества.

Фактическая пропускная способность сетевого соединения определяется оборудованием и линиями на пути через сеть и меняется в течение дня или нескольких дней. Исследования с линиями DSL в Северной Америке показывают, что более 30% этих линий имеют показатели пропускной способности меняющиеся более чем на 10% в течение дня или нескольких дней.

Ниже указаны некоторые факторы, влияющие на фактическую пропускную способность.

1. Наличие конкурирующего трафика в ЛВС или средах WAN. В ЛВС конкурирующий трафик образует несколько устройств, использующих одно соединение с Internet. В WAN конкурирующий трафик часто возникает из несвязанных сетевых потоков, использующих общий путь через сеть.
2. Возможности оборудования на пути сетевого соединения, включая скорость передачи данных и объем буферной памяти.
3. Активные средства управления трафиком, такие как формователи и ограничители трафика, зачастую применяемые провайдерами.

Имеются и другие факторы, негативно влияющие на фактическую пропускную способность линий.

Требования пользователей к трафику следуют картине использования и предпочтениям конкретного пользователя. Например, при большом переносе данных может использоваться вся доступная пропускная способность, тогда как потоковым приложениям для корректной работы достаточно ограниченной полосы. Для видеоконференций обычно требуется меньшая пропускная способность, чем для видеопотоков с высоким разрешением.

4.2.4. Показатели задержки

Сквозная задержка - это время, за которое конкретный пакет проходит путь через сеть от пользователя к адресату и обратно. Задержка состоит из нескольких частей.

1. Задержка распространения, отражающая протяжённость пути и технологии отдельных каналов (например, оптические или спутниковые). Распространение не зависит от загрузки сети, пока путь не меняется.
2. Задержка буферизации, отражающая интервалы времени, проведённого пакетом в сетевом оборудовании, соединяющем отдельные сетевые каналы, а также в памяти передающей конечной точки. Задержка буферизации зависит от загрузки сети, а также алгоритмов, управляющих буферизуемыми сегментами.
3. Задержки транспортных протоколов, отражающие время, затраченное на повтор передачи и сборку, а также время, когда транспорт был заблокирован (head-of-line blocked).
4. В некоторых материалах семинара явно упоминалась задержка приложений, отражающая неэффективность прикладного уровня.

Обычно сквозная задержка определяется при бездействии сети. Результаты таких измерений в основном отражают задержку распространения, но не другие части задержки. В этом отчёте используется термин «задержка при бездействии» (idle latency) для результатов, полученных во время бездействия сети.

Если измерения проводятся при типичной загрузке сети, результат отражает несколько частей задержки. В отчёте такая задержка называется рабочей (working latency). В других документах применяется термин «задержка под нагрузкой» (latency under load или LUL).

Представленные на семинаре документы показывают существенную разницу между задержкой при бездействии и под нагрузкой. В зависимости от направления трафика и технологии рабочая задержка может быть от 6 до 25 раз больше задержки при бездействии, как показано в таблице 1.

Таблица 1.

Направление	Технология	Задержка при работе	Задержка при простое	Разность задержек	Отношение работа/простой
Нисходящее	FTTH	148	10	138	15
Нисходящее	Cable	103	13	90	8
Нисходящее	DSL	194	10	184	19
Восходящее	FTTH	207	12	195	17
Восходящее	Cable	176	27	149	6
Восходящее	DSL	686	27	659	25

Хотя исторически средства измерения задержки были сосредоточены на задержках при бездействии, в отрасли наблюдается тенденция к измерению рабочей задержки (например, в Apple [NetworkQuality]).

4.2.5. Примеры измерений

Участники семинара предложили несколько конкретных методик измерения качества сети для конечных пользователей.

В [Paasch2021] представлена методика измерения рабочей задержки с точки зрения конечного пользователя. Предложенный метод постепенно добавляет сетевые потоки между пользовательским устройством и сервером, пока не возникнет «пробка» (bottleneck). Из этих измерений выводится время кругового обхода, сообщаемое конечному пользователю. Авторы решили выдавать результаты с метрикой RPM. Метод реализован в Apple macOS Monterey.

В [Mathis2021] применена метрика RPM к результатам более 4 миллиардов тестов загрузки, которые были выполнены M-Lab в 2010-2021 г. За это время измерительная платформа Mпретерпела несколько обновлений, что позволило команде исследователей сравнить влияние различных механизмов контроля перегрузок TCP (congestion control algorithm или CCA) на измерение сквозной задержки. Исследование показало, что применение кубического CCA ведёт к росту рабочей задержки, связанному с использованием очередей большего размера.

В [Schlinker2019] представлено масштабное исследование с попыткой сопоставить полезную пропускную способность (goodput) и QoE в большой социальной сети. Авторы проводили измерения в нескольких ЦОД из которых передавались сегменты видео заданного размера потоками большому числу конечных пользователей. Авторы применяли показатели полезной и общей пропускной способности для обнаружения перегрузки отдельных путей.

В [Reed2021] представлен анализ измерений рабочей задержки собранных по программе Measuring Broadband America (MBA) Федеральной комиссии по связи (Federal Communication Commission или FCC). FCC не включает рабочую задержку в свои годовые отчёты, но предоставляет её в файлах необработанных данных. Авторы использовали часть необработанных данных для определения важных различий между рабочими задержками у разных ISP.

В [MacMillian2021] представлен анализ измерений рабочей задержки на нескольких уровнях обслуживания. Обнаружено (неудивительно), что пользователи уровня premium сталкиваются с меньшей рабочей задержкой по сравнению с уровнем value. Данные показывают, что рабочая задержка существенно меняется на каждом уровне, одним из возможных объяснений является установка в домах разного оборудования.

Эти исследования подчеркнули важность измерения рабочей задержки. На момент написания этого отчёта многие производители домашних маршрутизаторов полагались на аппаратное ускорение маршрутизации с использованием очередей FIFO. Сосредоточение внимания на измерениях рабочей задержки в таких устройствах и ознакомление потребителей о влиянии выбора тех или иных производителей может помочь в улучшении ситуации с домашними маршрутизаторами. Идеальным тестом была бы возможность определить рабочую задержку и места задержек (домашний маршрутизатор, ISP, серверная сторона, тот или иной узел в сети).

Ещё одной причиной высоких рабочих задержек являются сетевые маршрутизаторы с перекрестным трафиком. Как показано в [Schlinker2019], они могут насыщаться в пиковое время суток. Систематическое тестирование рабочей задержки в нагруженных маршрутизаторах может улучшить понимание задержек и влияния на них развернутой инфраструктуры.

4.2.6. Ключевые точки показателей

Показатели качества сети можно грубо разбить на 4 группы.

1. Показатели доступности, указывающие возможность доступа пользователя в сеть.
2. Показатели пропускной способности, указывающие, достаточно ли фактической пропускной способности для удовлетворения потребностей пользователя.
3. Показатели задержки, указывающие, может ли пользователь получить данные своевременно.
4. Показатели высокого порядка, включающие показатели сети, такие как интервал прибытия пакетов, и показатели приложения, такие как время между перебуферизациями видеопотоков.

Показатели доступности можно рассматривать как производные от пропускной способности (отсутствие пропускной способности означает недоступность) или задержки (бесконечная задержка означает недоступность).

В презентациях и дискуссиях было указано несколько ключевых точек.

1. Доступность и пропускная способность являются «гигиеническими» факторами - пока приложение не способно использовать дополнительную пропускную способность, конечные пользователи не увидят преимуществ от применения линий с избыточной скоростью.
2. Рабочая задержка сильнее влияет на работу пользователя, нежели задержка при бездействии сети и может превосходить последнюю на порядок.
3. Метрика RPM является стабильной, большие значения говорят о лучшем качестве и метрика более понятна конечному пользователю.
4. Связь между общей и полезной пропускной способностью может быть полезна при поиске точек насыщения не стороне клиента [Paasch2021] и сервера [Schlinker2019].
5. Рабочая задержка зависит от выбора алгоритмов контроля перегрузок для целевой конечной точки и очередей в маршрутизаторах.

Участники согласились с тем, что лучшими показателями являются те, которые позволяют действовать.

4.3. Межуровневые вопросы

В межуровневой части семинара были представлены материалы и прошли дискуссии о точно обнаружении проблемных мест. Обсуждение сосредоточилось на различиях между кабельными и беспроводными соединениями и сложности точного определения проблемных мест в случаях, когда качество зависит от множества разнотипных сегментов сети. Например, в [Kerepez2021] показано, что Wi-Fi 2,4 ГГц с ограниченной пропускной способностью чаще всего создаёт «пробки». С Wi-Fi на частоте 5 ГГц связано лишь 20% наблюдавшихся пробок.

Участники согласились с тем, что ни один элемент сетевого соединения не имеет всех данных, требуемых для оценки влияния производительности сети на качество взаимодействия с конечным пользователем.

- Приложения на устройствах конечных пользователей лучше всего знают о соответствующей производительности, но имеют ограниченное представление о поведении сети и не могут действовать на основе своего ограниченного представления.
- ISP хорошо представляют ситуацию с QoS, но не могут связать показатели QoS с работой конечных пользователей.
- Контент-провайдеры хорошо представляют общее поведение конечных пользователей, но не знают о том, какие аспекты производительности сети служат основными индикаторами поведения пользователей.

На семинаре выявлена потребность в стандартах и расширяемом способе обмена сведениями о производительности сети. Такому стандарту обмена следует учитывать по меньшей мере указанные ниже аспекты.

- Расширяемый способ фиксации производительности множества (потенциально, тысяч) конечных точек.
- Формату обмена данных следует предотвращать манипуляции с данными, чтобы участники обмена не могли обмануть механизмы.
- Сохранение приватности пользователей. В частности следует применять федеративный подход, чтобы не возникало центрального элемента с доступом ко всей картине.
- Модель, стимулирующая различных участников сетевого соединения делиться собранными данными о производительности.
- Сопутствующий набор инструментов для анализа данных.

4.3.1. Разделение ответственности

Обычно имеется тесная связь между сбором показателей производительности, их интерпретацией и предпринимаемыми на основе интерпретации действиями. К сожалению, такая модель не является лучшим решением для обмена межуровневыми данными поскольку:

- у субъектов, способных собирать определённые показатели производительности (например, TCP RTT) может не быть контекста, требуемого для осмысленной интерпретации;
- у субъектов, имеющих контекст и возможности расчётов и хранения для интерпретации показателей, может не быть возможности управлять поведением сети и/или приложений;
- у контролирующих поведение сетей и/или приложений субъектов может не быть доступа ко всем измеренным данным.

Участники семинара согласились с тем, что важно разделить 3 указанных аспекта, чтобы:

- участники, имеющие данные, но не способные их интерпретировать или предпринимать действия на их основе, публиковали полученные данные;
- участники, имеющие опыт интерпретации и синтеза данных производительности, публиковали результаты своей интерпретации.

4.3.2. Вопросы безопасности и приватности

Сохранение приватности конечных пользователей Internet является сложной задачей. Имеются внутренние противоречия между сбором данных о действиях пользователей и сохранением приватности. Участники семинара согласились с тем, что для точного измерения качества сети необходимы наблюдения на нескольких уровнях, но задача минимизации утечки приватных сведений остаётся не решенной.

4.3.3. Вопросы измерения показателей

- Указанные ниже показатели протокола TCP сочтены эффективными и доступными для пассивных измерений.
 - Задержка соединения TCP, измеряемая по времени SACK или ACK, а также время между повторными передачами TCP служат хорошими показателями для сквозных измерений RTT.
 - На платформах Linux структура tcp_info является фактическим стандартом для приложений, позволяющим проверить производительность сети в пространстве ядра. Однако в пользовательском пространстве эквивалентного фактического стандарта нет.
- Протоколы QUIC и MASQUE усложняют пассивные измерения производительности.
 - Для этих протоколов может оказаться более подходящим подход с федеративными измерениями и иерархическим агрегированием.
 - Формат QLOG представляется наиболее подходящим для такого обмена.

4.3.4. Пути улучшения межуровневой наблюдаемости

Владение Internet распределено между множеством административных доменов, что затрудняет получение данных о сквозной производительности. Кроме того, огромный масштаб Internet осложняет их объединение и анализ. В [Marx2021] представлен простой формат ведения журнала, который можно применять для сбора и агрегирования данных от различных уровней.

Ещё одна сложность межуровневого взаимодействия для измерений связана с тем, что большинство современных алгоритмов не представляют в явном виде данные о производительности, пригодные для межуровневого анализа. Сообщество IETF могло бы проявить усердие в определении ключевых идентификаторов производительности протоколов и их раскрытии в спецификации каждого протокола.

Несмотря на эти проблемы, возможно проводить исследования в ограниченных областях для лучшего понимания влияний взаимодействия различных элементов Internet на качество работы пользователей. Кроме того, недавние разработки федеративных алгоритмов обучения предполагают возможность проведения межуровневых измерений производительности с сохранением приватности пользователей.

4.3.5. Взаимодействие оборудования с транспортными протоколами

С появлением уведомления и контроля передрузок с малыми задержками и потерями, а также расширяемой пропускной (L4S) возросла потребность в согласовании работы транспортных протоколов и базового оборудования.

На момент проведения семинара типичные домашние маршрутизаторы использовали одну очередь FIFO достаточно большого размера для амортизации издержек, связанных с заголовками нижних уровней в различных транспортных PDU. Это хорошо работало с кубическим алгоритмом контроля перегрузок, однако новые алгоритмы способны работать с гораздо меньшими очередями. Чтобы обеспечить задержки меньше 1 мсек, домашний маршрутизатор

должен эффективно работать при последовательной передаче всего нескольких сегментов, а не быть оптимизированным для длительных пиков пакетов. Ещё одной конструктивной особенностью домашних маршрутизаторов является агрегирование пакетов для дополнительной амортизации издержек, связанных с заголовками нижних уровней. В частности, несколько дейтаграмм IP объединяются в один большой кадр передачи. Однако такое агрегирование может добавлять к задержке пакетов в устройстве до 10 мсек.

Следуя известной поговорке: «нельзя улучшить то, что невозможно измерить», важно выявить такие задержки агрегирования так, чтобы можно было найти причину пробок и сделать оборудование более подходящим для транспортных протоколов следующего поколения.

4.3.6. Ключевые точки межуровневых измерений

- Характеристики измеряемых показателей существенно различаются и нужна оптимизация для кабельных и беспроводных сетей.
- Выявление проблемных мест затруднено из-за измерений не множестве сегментов пути.
- Ни в одной точке пути нет всех данных, требуемых для оценки влияния производительности сети в целом на работу конечного пользователя.
- Для практических результатов требуется подходящий сбор и интерпретация данных.
- Важна координация между сетевыми провайдерами для лучшей оценки качества работы пользователей.
- Одновременное обеспечение точных измерений и сохранение приватности пользователей затруднено.
- Пассивные измерения в реализациях протоколов могут предоставить полезные данные.

4.4. Объединительная часть

Объединительная часть семинара включала презентации и обсуждение следующих шагов, которые могут потребоваться для продвижения. Особую озабоченность вызывают вопросы представления измерений так, чтобы они были понятны конечным пользователям, пытающимся выбрать опции подписки на сетевые услуги.

4.4.1. Вопросы измерений и показателей

Важным является вопрос принятия решений и выполнения действий на основе собранных показателей. Измерения нужно объединять с приложениями для того, чтобы те корректно видели перегрузки, поскольку измерения в иной инфраструктуре или через другие приложения могут давать неверные результаты. Перегрузка может быть временной проблемой и стратегия смягчения может меняться в зависимости от её продолжительности. Для краткосрочных перегрузок возникает проблема необходимости постоянных измерений для учёта критических моментов и продолжительных тенденций. Для таких перегрузок участники семинара обсудили вопрос, является ли исчезновение перегрузки проблемой или служит признаком приспособления и самовосстановления.

При определении показателей важно учитывать факторы, влияющие на результат. Это могут быть характеристики отдельных пакетов - для пакетов разного размера задержка обычно зависит от размера линейно. С учетом этого можно выделить задержки из-за протяжённости (географическое разделение), пакетизации и иных переменных факторов (шум). Каждая из этих долей задержки может различаться и измеряться отдельно в каждом сегменте многоэтапного пути. На переменную задержку могут существенно влиять внешние факторы, такие как буферизация, изменение маршрута, распределение нагрузки в сети и другие локальные или удалённые изменения производительности. Сетевые измерения, особенно тесты под нагрузкой, должны быть достаточно продолжительными, чтобы учесть все проблемы, связанные с буферизацией, очередями и т. п. Технологиям измерений следует различать восходящее (от пользователя) и нисходящее направление, а также результаты сквозных и сегментированных измерений.

4.4.2. Представление показателей конечного пользователя

Для определения потребностей конечных пользователей нужны информативные показатели и измерения. Как предоставить пользователю услуги, которые ему нужны? Могут ли пользователи эффективно выразить свои потребности? Пользователи обычно дают лишь простые ответы верхнего уровня, такие как надёжность, пропускная способность, комплексные услуги. Технические требования, понятные оператору, такие как малая задержка или предотвращение перегрузок, непонятны и не применяются конечными пользователями.

Примеры полезных для конечного пользователя показателей включают число поддерживаемых клиентов и приложений или потоков. Примером решения для борьбы с сетевыми проблемами является стратегия управления трафиком на основе стимулов (например, запрос приложением меньшей задержки может восприниматься как готовность принять меньшую пропускную способность). Нужно учитывать воспринимаемую пользователем задержку, а не только задержку в сети - взаимодействие с приложением, задержка на сервере и в сети могут учитывать лишь задержки нижнего уровня. Выбор правильного протокола для измерений очень важен для соответствия пользовательскому опыту (например, пользователь не передаёт данных по протоколу ICMP, хотя это один из важных инструментов измерения).

Измерениям в приложениях следует учитывать тип приложения, например, видеопотоки, обмен файлами, групповая игра или голосовая связь. Возможно, было бы полезно спросить пользователя, на какие компромиссы он готов пойти - например, нужна ему малая задержка или большая пропускная способность? Для игры могут потребоваться иные решения, чем для домашнего офиса или распространения содержимого. Как пользователи могут принять компромисс, не мешая другим пользователям: Имеется противоречие между решением этих проблем и готовностью клиентов к затратам на будущее улучшение.

Проблемы с предоставлением пользовательскому трафику более высокого приоритета связаны с возможностью сетей воспринимать запросы клиентов как стимул, даже когда это не вполне соответствует коммерческим интересам. В средах общего пользования часто возникает «переподписка», когда числа пользователей, которых может поддерживать сеть, достаточно при незагруженной сети, но недостаточно в условиях высокой загрузки. На отдельные показатели также влияют домашние устройства от недорогих маршрутизаторов до микроволновых печей и поведение

пользователей во время тестов. Показатель сам по себе и одиночное измерение без контекста не помогут оператору или пользователю найти источник проблем.

Понимание сети пользователями остаётся проблемой. Многие участники семинара предлагали одно значение (возможно, рассчитанное со взвешенным агрегированием) или небольшое число измерений за ожидаемое время использования (например, оценка для игры или доступа к содержимому). Многие согласились с тем, что часть пользователей может предпочесть упрощённый или цветовой рейтинг (например, хорошо-лучше-отлично, красный-желтый-зеленый или бронза-золото-платина).

4.4.3. Ключевые точки объединения

- Некоторые предложенные показатели:
 - число круговых обходов в минуту (Round-trips Per Minute или RPM);
 - число пользователей на сеть;
 - задержка;
 - 99% задержки и пропускная способность.
- Измерения медианного и среднего значения отвлекают от реальных проблем.
- Сеть с общим доступом существенно влияет на качество.
- Нужны продолжительные измерения для фиксации всех аспектов возможных пробок в сети.
- Для успешных исследований требуется лучшее финансирование.
- Конечные пользователи лучше поймут упрощённую систему оценки или ранжирования.

5. Выводы

В заключительном часе 3-дневного семинара были собраны заявления, которые группа сочла резюмирующими. Позднее все спорные утверждения были отброшены (они указаны ниже для полноты). Для этого документа авторы взяли исходный список и грубо разделили его на категории, внесли предложенные в почтовой конференции изменения и отредактировали для ясности и обеспечения контекста.

5.1. Общие заявления

1. Пропускная способность (полоса) необходима, но не достаточна.
2. Во многих случаях пользователям Internet нужно не больше полосы, а лучшая пропускная способность, т. е. требуются иные улучшения связности.
3. Нужны активные и пассивные измерения, пассивные могут обеспечить «историческую отладку».
4. Пассивные измерения (включая надёжность и связность) должны быть непрерывными, архивируемыми и доступными для запроса.
5. Для пользователей значимо, будет ли их приложение корректно работать или откажет из-за отсутствия нужных характеристик сети.
6. Полезные показатели для ценности (goodness) должны фактически стимулировать ценность - хорошим показателям следует быть действенными в плане совершенствования отрасли.
7. Снижение задержки в Internet принесёт пользу всем конечным пользователям.

5.2. Конкретные заявления о протоколах и методах

1. Число круговых обходов в минуту (RPM) является полезным и востребованным показателем.
2. Нужны полезные инструменты, заполняющие зазоры между тестами доступности сети, задержки и скорости.
3. Конечным пользователям, желающим участвовать в QoS, нужно дать возможность заявить свои потребности.
4. Нужны приложения, выполняющие хорошие измерения и отчёты о качестве для обнаружения недостатков доступа в сеть.
5. Проведённые регуляторами исследования показывают, что пользователи (потребители) предпочитают простые показатели для приложения, указывающие способность приложения работать должным образом.
6. Новые измерения и методы QoS и QoE не должны зависеть лишь от заголовков TCP.
7. От разработчиков интерактивных приложений и операторов сетей известно, что задержка является важным фактором для QoE пользователей. Однако для прямой поддержки этого утверждения нет показателей.

5.3. Заявления о проблемах и ответственности

1. Измерения медианного и среднего значения задержки отвлекают от более важных измерений.
2. Разочаровывает измерение сетевых служб без одновременного роста качества обслуживания.
3. Заинтересованные стороны не ждут лёгких побед и нужны стимулы для улучшения доступа в общественные сети. Измерения могут быть одним из шагов к созданию конкурентных рыночных стимулов.
4. Для перспективных сетей важно учитывать влияние на экологию применяемых материалов и энергии.
5. Нет неопровержимых свидетельств большей важности какого-либо показателя (например, задержки или скорости), чтобы убедить производителей устройств сосредоточиться на его оптимизации.

5.4. Утверждения, по которым не достигнуто согласия

Рассмотрены и записаны утверждения, по которым не было выработано единого мнения.

1. Нет неопровержимых свидетельств того, что раздувание буферов является распространённой проблемой.
2. Измерения должны поддерживать локализацию отчётов для нахождения проблем.
 - Обнаружения проблемы недостаточно, если не найдено место её возникновения.
 - Нужна поддержка не только английского языка, но и других языков.
3. Заинтересованные стороны не настроены на лёгкие победы в этом направлении.

6. Последующая работа

На семинаре обсуждались вопросы будущей работы. Предложено часть работ продолжать в имеющихся рабочих группах IETF (например, IPPM, DetNet, RAW), а длительные исследования могут потребовать создания групп IRTF.

7. Взаимодействие с IANA

Этот документ не требует действий IANA.

8. Вопросы безопасности

На семинаре обсуждалось несколько связанных с безопасностью тем, включая:

- методы приоритизации, которые могут работать без вторжения в приватность взаимодействующих сторон;
- как сети с превышением трафика (oversubscribed) могут казаться находящимися под DDoS-атакой.

9. Литература

- [Aldabbagh2021] Aldabbagh, A., "Regulatory perspective on measuring network quality for end-users", September 2021, <<https://www.iab.org/wp-content/IAB-uploads/2021/09/2021-09-07-Aldabbagh-Ofcom-presentation-to-IAB-1v00-1.pdf>>.
- [Arkko2021] Arkko, J. and M. Kühlewind, "Observability is needed to improve network quality", August 2021, <<https://www.iab.org/wp-content/IAB-uploads/2021/09/iab-position-paper-observability.pdf>>.
- [Balasubramanian2021] Balasubramanian, P., "Transport Layer Statistics for Network Quality", February 2021, <<https://www.iab.org/wp-content/IAB-uploads/2021/09/transportstatsquality.pdf>>.
- [Briscoe2021] Briscoe, B., White, G., Goel, V., and K. De Schepper, "A Single Common Metric to Characterize Varying Packet Delay", September 2021, <<https://www.iab.org/wp-content/IAB-uploads/2021/09/single-delay-metric-1.pdf>>.
- [Casas2021] Casas, P., "10 Years of Internet-QoE Measurements Video, Cloud, Conferencing, Web and Apps. What do we need from the Network Side?", August 2021, <https://www.iab.org/wp-content/IAB-uploads/2021/09/net_quality_internet_qoe_CASAS.pdf>.
- [Cheshire2021] Cheshire, S., "The Internet is a Shared Network", August 2021, <<https://www.iab.org/wp-content/IAB-uploads/2021/09/draft-cheshire-internet-is-shared-00b.pdf>>.
- [Davies2021] Davies, N. and P. Thompson, "Measuring Network Impact on Application Outcomes Using Quality Attenuation", September 2021, <<https://www.iab.org/wp-content/IAB-uploads/2021/09/PNSol-et-al-Submission-to-Measuring-Network-Quality-for-End-Users-1.pdf>>.
- [DeSchepper2021] De Schepper, K., Tilmans, O., and G. Dion, "Challenges and opportunities of hardware support for Low Queuing Latency without Packet Loss", February 2021, <<https://www.iab.org/wp-content/IAB-uploads/2021/09/Nokia-IAB-Measuring-Network-Quality-Low-Latency-measurement-workshop-20210802.pdf>>.
- [Dion2021] Dion, G., De Schepper, K., and O. Tilmans, "Focusing on latency, not throughput, to provide a better internet experience and network quality", August 2021, <<https://www.iab.org/wp-content/IAB-uploads/2021/09/Nokia-IAB-Measuring-Network-Quality-Improving-and-focusing-on-latency-.pdf>>.
- [Fabini2021] Fabini, J., "Network Quality from an End User Perspective", February 2021, <<https://www.iab.org/wp-content/IAB-uploads/2021/09/Fabini-IAB-NetworkQuality.txt>>.
- [FCC_MBA] FCC, "Measuring Broadband America", <<https://www.fcc.gov/general/measuring-broadband-america>>.
- [FCC_MBA_methodology] FCC, "Measuring Broadband America - Open Methodology", <<https://www.fcc.gov/general/measuring-broadband-america-open-methodology>>.
- [Foulkes2021] Foulkes, J., "Metrics helpful in assessing Internet Quality", September 2021, <https://www.iab.org/wp-content/IAB-uploads/2021/09/IAB_Metrics_helpful_in_assessing_Internet_Quality.pdf>.
- [Ghai2021] Ghai, R., "Using TCP Connect Latency for measuring CX and Network Optimization", February 2021, <<https://www.iab.org/wp-content/IAB-uploads/2021/09/xfinity-wifi-ietf-iab-v2-1.pdf>>.
- [Iyengar2021] Iyengar, J., "The Internet Exists In Its Use", August 2021, <<https://www.iab.org/wp-content/IAB-uploads/2021/09/The-Internet-Exists-In-Its-Use.pdf>>.
- [Kerpez2021] Shafiei, J., Kerpez, K., Cioffi, J., Chow, P., and D. Bousaber, "Wi-Fi and Broadband Data", September 2021, <<https://www.iab.org/wp-content/IAB-uploads/2021/09/Wi-Fi-Report-ASSIA.pdf>>.

- [Kilkki2021] Kilkki, K. and B. Finley, "In Search of Lost QoS", February 2021, <<https://www.iab.org/wp-content/IAB-uploads/2021/09/Kilkki-In-Search-of-Lost-QoS.pdf>>.
- [Laki2021] Nadas, S., Varga, B., Contreras, L.M., and S. Laki, "Incentive-Based Traffic Management and QoS Measurements", February 2021, <https://www.iab.org/wp-content/IAB-uploads/2021/11/CamRdy-IAB_user_meas_WS_Nadas_et_al_IncentiveBasedTMwQoS.pdf>.
- [Liubogoshchev2021] Liubogoshchev, M., "Cross-layer Cooperation for Better Network Service", February 2021, <<https://www.iab.org/wp-content/IAB-uploads/2021/09/Cross-layer-Cooperation-for-Better-Network-Service-2.pdf>>.
- [MacMillian2021] MacMillian, K. and N. Feamster, "Beyond Speed Test: Measuring Latency Under Load Across Different Speed Tiers", February 2021, <https://www.iab.org/wp-content/IAB-uploads/2021/09/2021_nqw_lul.pdf>.
- [Marx2021] Marx, R. and J. Herbots, "Merge Those Metrics: Towards Holistic (Protocol) Logging", February 2021, <https://www.iab.org/wp-content/IAB-uploads/2021/09/MergeThoseMetrics_Marx_Jul2021.pdf>.
- [Mathis2021] Mathis, M., "Preliminary Longitudinal Study of Internet Responsiveness", August 2021, <<https://www.iab.org/wp-content/IAB-uploads/2021/09/Preliminary-Longitudinal-Study-of-Internet-Responsiveness-1.pdf>>.
- [McIntyre2021] Paasch, C., McIntyre, K., Shapira, O., Meyer, R., and S. Cheshire, "An end-user approach to an Internet Score", September 2021, <<https://www.iab.org/wp-content/IAB-uploads/2021/09/Internet-Score-2.pdf>>.
- [Michel2021] Michel, F. and O. Bonaventure, "Packet delivery time as a tie-breaker for assessing Wi-Fi access points", February 2021, <https://www.iab.org/wp-content/IAB-uploads/2021/09/camera_ready_Packet_delivery_time_as_a_tie_breaker_for_assessing_Wi-Fi_access_points.pdf>.
- [Mirsky2021] Mirsky, G., Min, X., Mishra, G., and L. Han, "The error performance metric in a packet-switched network", February 2021, <<https://www.iab.org/wp-content/IAB-uploads/2021/09/IAB-workshop-Error-performance-measurement-in-packet-switched-networks.pdf>>.
- [Morton2021] Morton, A. C., "Dream-Pipe or Pipe-Dream: What Do Users Want (and how can we assure it)?", Work in Progress, Internet-Draft, draft-morton-ippm-pipe-dream-01, 6 September 2021, <<https://datatracker.ietf.org/doc/html/draft-morton-ippm-pipe-dream-01>>.
- [NetworkQuality] Apple, "Network Quality", <<https://support.apple.com/en-gb/HT212313>>.
- [Paasch2021] Paasch, C., Meyer, R., Cheshire, S., and O. Shapira, "Responsiveness under Working Conditions", Work in Progress, Internet-Draft, draft-cpaasch-ippm-responsiveness-01, 25 October 2021, <<https://datatracker.ietf.org/doc/html/draft-cpaasch-ippm-responsiveness-01>>.
- [Pardue2021] Pardue, L. and S. Tellakula, "Lower-layer performance is not indicative of upper-layer success", February 2021, <<https://www.iab.org/wp-content/IAB-uploads/2021/09/Lower-layer-performance-is-not-indicative-of-upper-layer-success-20210906-00-1.pdf>>.
- [Reed2021] Reed, D.P. and L. Perigo, "Measuring ISP Performance in Broadband America: A Study of Latency Under Load", February 2021, <https://www.iab.org/wp-content/IAB-uploads/2021/09/Camera_Ready_-_Measuring-ISP-Performance-in-Broadband-America.pdf>.
- [SamKnows] "SamKnows", <<https://www.samknows.com/>>.
- [Schlinker2019] Schlinker, B., Cunha, I., Chiu, Y., Sundaresan, S., and E. Katz-Basset, "Internet Performance from Facebook's Edge", February 2019, <<https://www.iab.org/wp-content/IAB-uploads/2021/09/Internet-Performance-from-Facebooks-Edge.pdf>>.
- [Scuba] Abraham, L. et al., "Scuba: Diving into Data at Facebook", <<https://research.facebook.com/publications/scuba-diving-into-data-at-facebook/>>.
- [Sengupta2021] Sengupta, S., Kim, H., and J. Rexford, "Fine-Grained RTT Monitoring Inside the Network", February 2021, <https://www.iab.org/wp-content/IAB-uploads/2021/09/Camera_Ready_Fine-Grained_RTT_Monitoring_Inside_the_Network.pdf>.
- [Sivaraman2021] Sivaraman, V., Madanapalli, S., and H. Kumar, "Measuring Network Experience Meaningfully, Accurately, and Scalably", February 2021, <<https://www.iab.org/wp-content/IAB-uploads/2021/09/CanopusPositionPaperCameraReady.pdf>>.
- [Speedtest] Ookla, "Speedtest", <<https://www.speedtest.net>>.
- [Stein2021] Stein, Y., "The Futility of QoS", August 2021, <<https://www.iab.org/wp-content/IAB-uploads/2021/09/QoS-futility.pdf>>.
- [Welzl2021] Welzl, M., "A Case for Long-Term Statistics", February 2021, <https://www.iab.org/wp-content/IAB-uploads/2021/09/iab-longtermstats_cameraready_docx-1.pdf>.
- [WORKSHOP] IAB, "IAB Workshop: Measuring Network Quality for End- Users, 2021", September 2021, <<https://www.iab.org/activities/workshops/network-quality>>.
- [Zhang2021] Zhang, M., Goel, V., and L. Xu, "User-Perceived Latency to Measure CCAs", September 2021, <https://www.iab.org/wp-content/IAB-uploads/2021/09/User_Perceived_Latency-1.pdf>.

Приложение А. Программный комитет

Jari Arkko	Toke Hoiland-Jorgensen	Randall Meyer
Olivier Bonaventure	Geoff Huston	Kathleen Nichols
Vint Cerf	Cullen Jennings	Christoph Paasch
Stuart Cheshire	Katarzyna Kosek-Szott	Tommy Pauly
Sam Crowford	Mirja Kühlewind	Greg White
Nick Feamster	Jason Livingood	Keith Winstein
Jim Gettys	Matt Mathis	

Приложение В. Руководители семинара

Wes Hardaker
Evgeny Khorov
Omer Shapira

Приложение С. Участники семинара

Ahmed Aldabbagh	Joris Herbots	Levi Perigo
Jari Arkko	Geoff Huston	David Reed
Praveen Balasubramanian	Toke Høiland-Jørgensen	Alvaro Retana
Olivier Bonaventure	Jana Iyengar	Roberto
Djamel Bousaber	Cullen Jennings	Koen De Schepper
Bob Briscoe	Ken Kerpez	David Schinazi
Rich Brown	Evgeny Khorov	Brandon Schlinder
Anna Brunstrom	Kalevi Kilkki	Eve Schooler
Pedro Casas	Joon Kim	Satadal Sengupta
Vint Cerf	Zhenbin Li	Jinous Shafiei
Stuart Cheshire	Mikhail Liubogoshchev	Shapelez
Kenjiro Cho	Jason Livingood	Omer Shapira
Steve Christianson	Kyle MacMillan	Dan Siemon
John Cioffi	Sharat Madanapalli	Vijay Sivaraman
Alexander Clemm	Vesna Manojlovic	Karthik Sundaresan
Luis M. Contreras	Robin Marx	Dave Taht
Sam Crawford	Matt Mathis	Rick Taylor
Neil Davies	Jared Mauch	Bjørn Ivar Teigen
Gino Dion	Kristen McIntyre	Nicolas Tessaes
Toerless Eckert	Randall Meyer	Peter Thompson
Lars Eggert	François Michel	Balazs Varga
Joachim Fabini	Greg Mirsky	Bren Tully Walsh
Gorry Fairhurst	Cindy Morgan	Michael Welzl
Nick Feamster	Al Morton	Greg White
Mat Ford	Szilveszter Nadas	Russ White
Jonathan Foulkes	Kathleen Nichols	Keith Winstein
Jim Gettys	Lai Yi Ohlsen	Lisong Xu
Rajat Ghai	Christoph Paasch	Jiankang Yao
Vidhi Goel	Lucas Pardue	Gavin Young
Wes Hardaker	Tommy Pauly	Mingrui Zhang

Члены IAB на момент принятия документа для публикации

Jari Arkko	Mallory Knodel	Russ White
Deborah Brungard	Mirja Kühlewind	Qin Wu
Lars Eggert	Zhenbin Li	Jiankang Yao
Wes Hardaker	Tommy Pauly	
Cullen Jennings	David Schinazi	

Благодарности

Авторы признательны участникам семинара, членам IAB и программному комитету за интересные дискуссии.

Участники работы

Спасибо участникам работы над этим документом:

Erik Auerswald
Simon Leinen
Brian Trammell

Адреса авторов

Wes Hardaker
Email: ietf@hardakers.net

Omer Shapira
Email: omer_shapira@apple.com

Перевод на русский язык

Николай Малых
nmalykh@protokols.ru