



RIPE NCC

RIPE NETWORK COORDINATION CENTRE

Studying Connectivity Between Russian Cities with RIPE Atlas



Round Trip Time (RTT)

$$RTT = d_1 + d_2 + d_3$$

$$d_1 = \sum_i (\textit{propagation delay})_i$$

$$d_2 = \sum_i (\textit{equipment delay})_i$$

$$d_3 = \sum_i (\textit{queue delay})_i$$

$$\min(d_3) = 0 \quad (!)$$

Channels

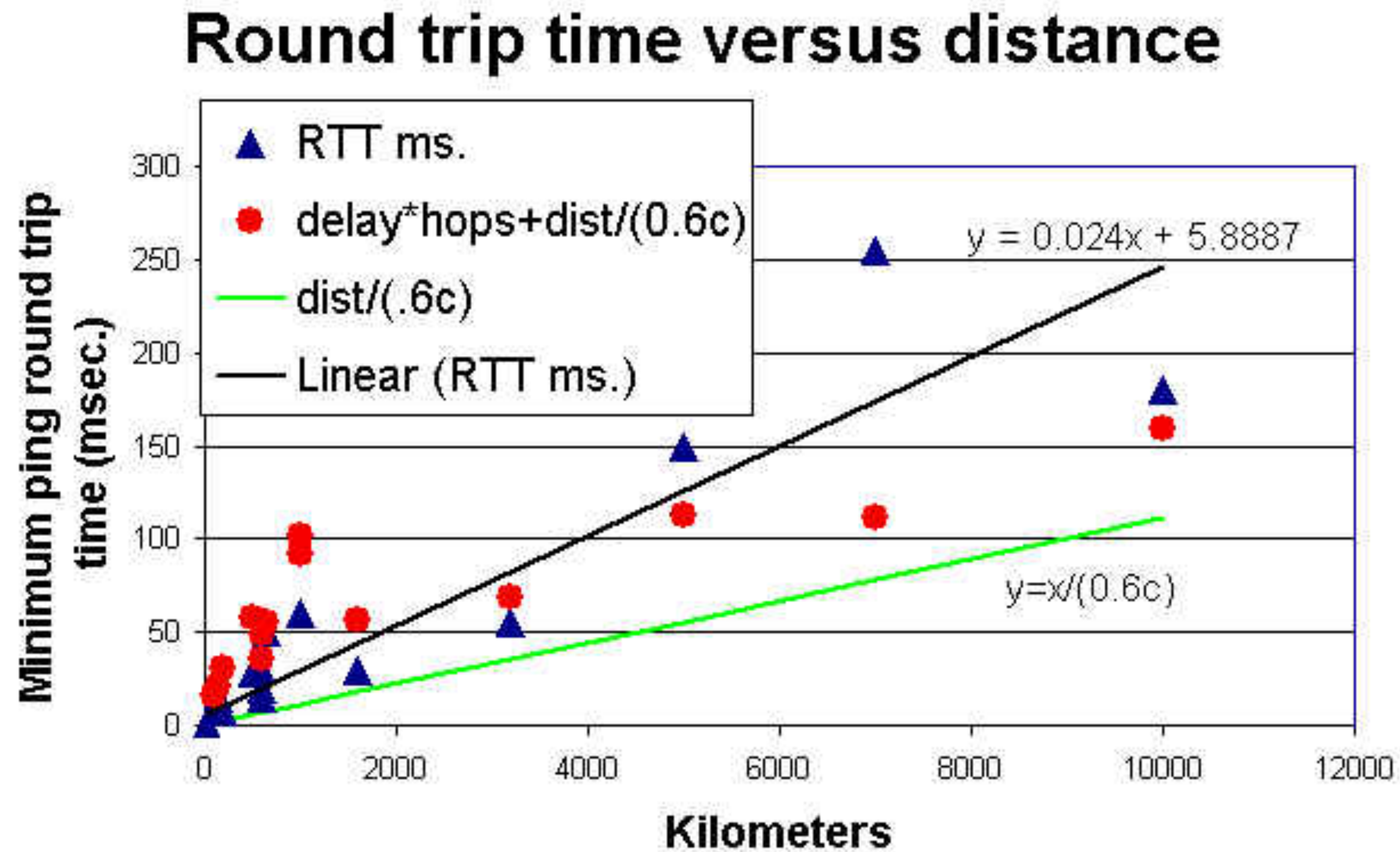
Hardware

Load



Зависимость RTT от расстояния

- Les Cottrell, Warren Matthews and Connie Logg из Stanford University, 2000-2002 гг.:



Зависимость RTT от расстояния



$$RTT = \min(\{RTT_i\}) = d_1 + d_2$$

$$d_1 = \sum_i \frac{distance_i}{c_{fibre\ i}}$$

Length of the fiber sector

Velocity of speed in this fibre

Скорость света в разном волокне отличается не очень сильно ($\pm 10\%$), поэтому:

$$d_1 \approx \frac{total\ distance}{c_{fibre}}$$



Зависимость RTT от расстояния

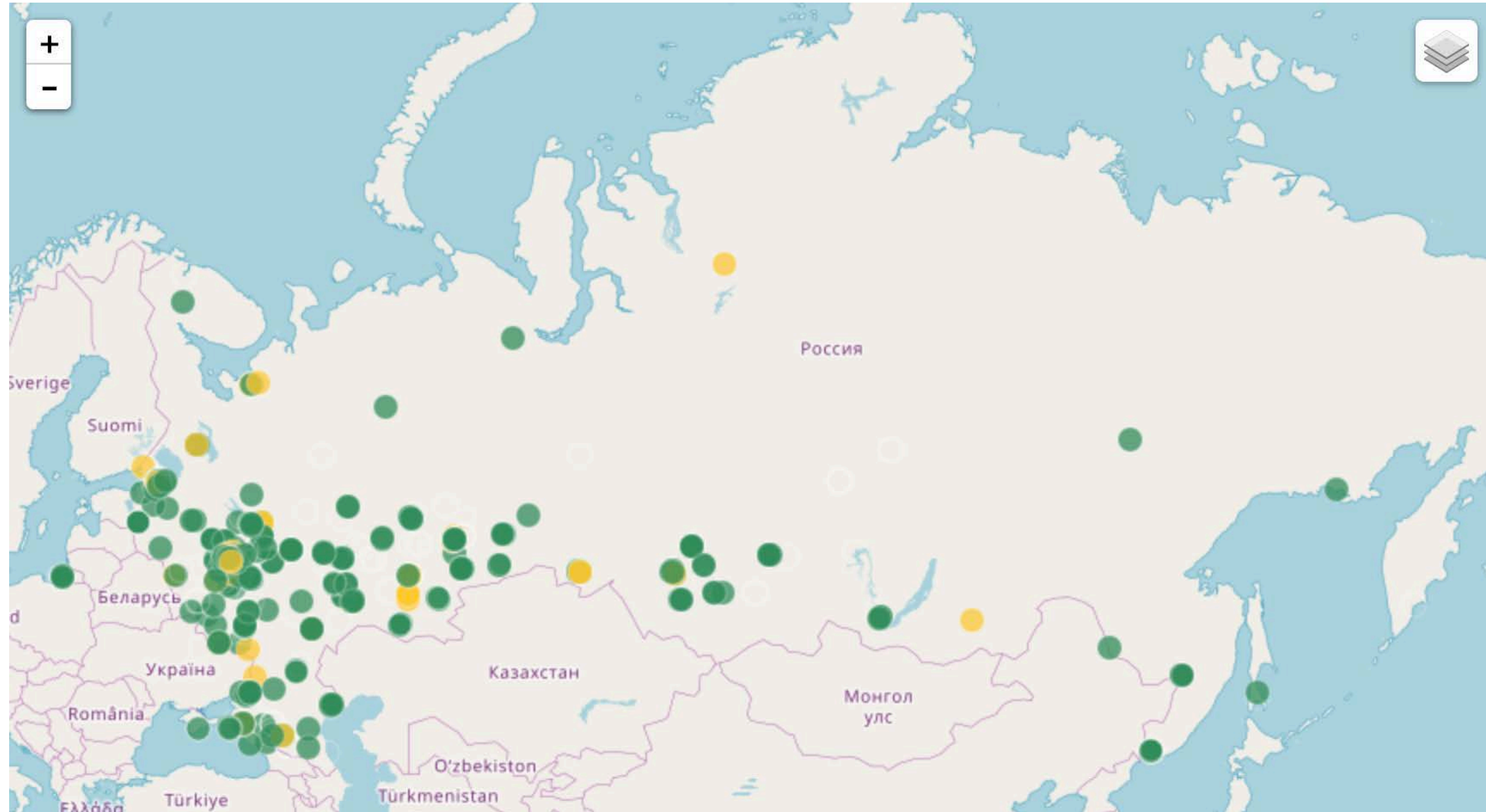
- В 2000 году (оборудование медленное, ASICов нет etc):

$$d_2 = 5.9 \text{ ms}$$

- В 2019 можно ожидать существенно меньшего значения.
- Итого, минимальное значение RTT между двумя узлами примерно пропорционально расстоянию между ними:

$$RTT \approx \frac{\text{total distance}}{c_{\text{fibre}}}$$

Чем может помочь RIPE Atlas?

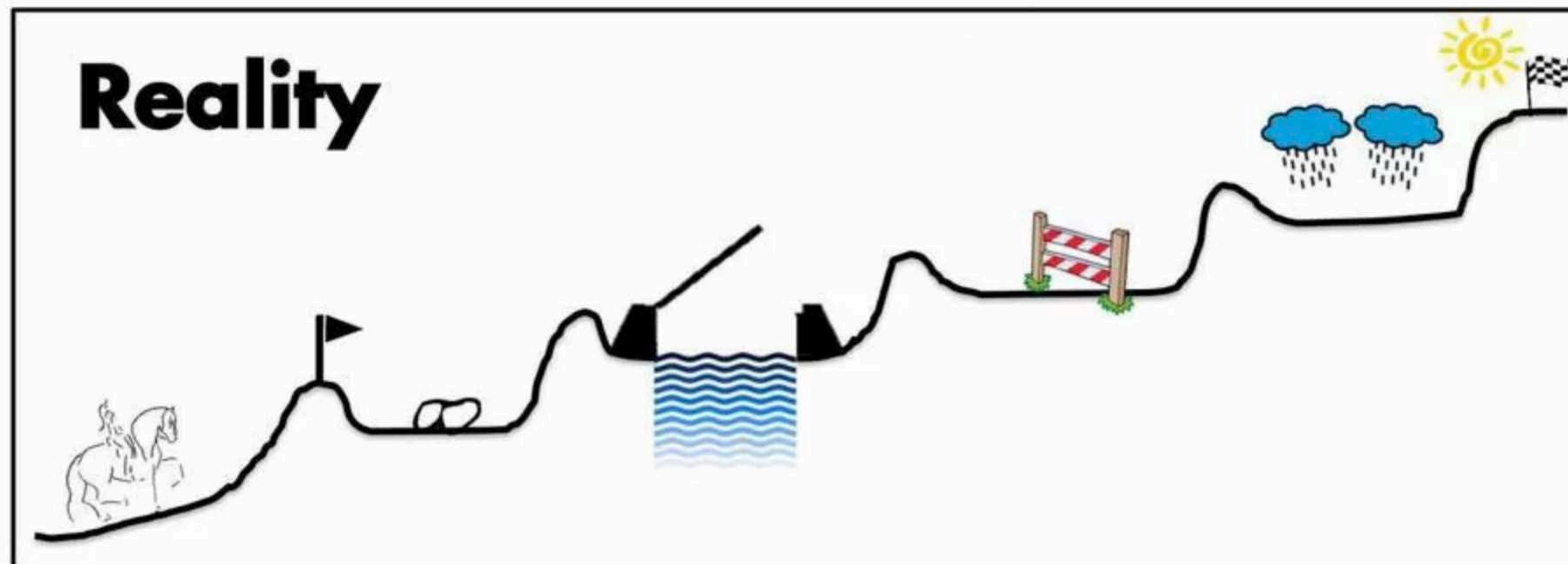
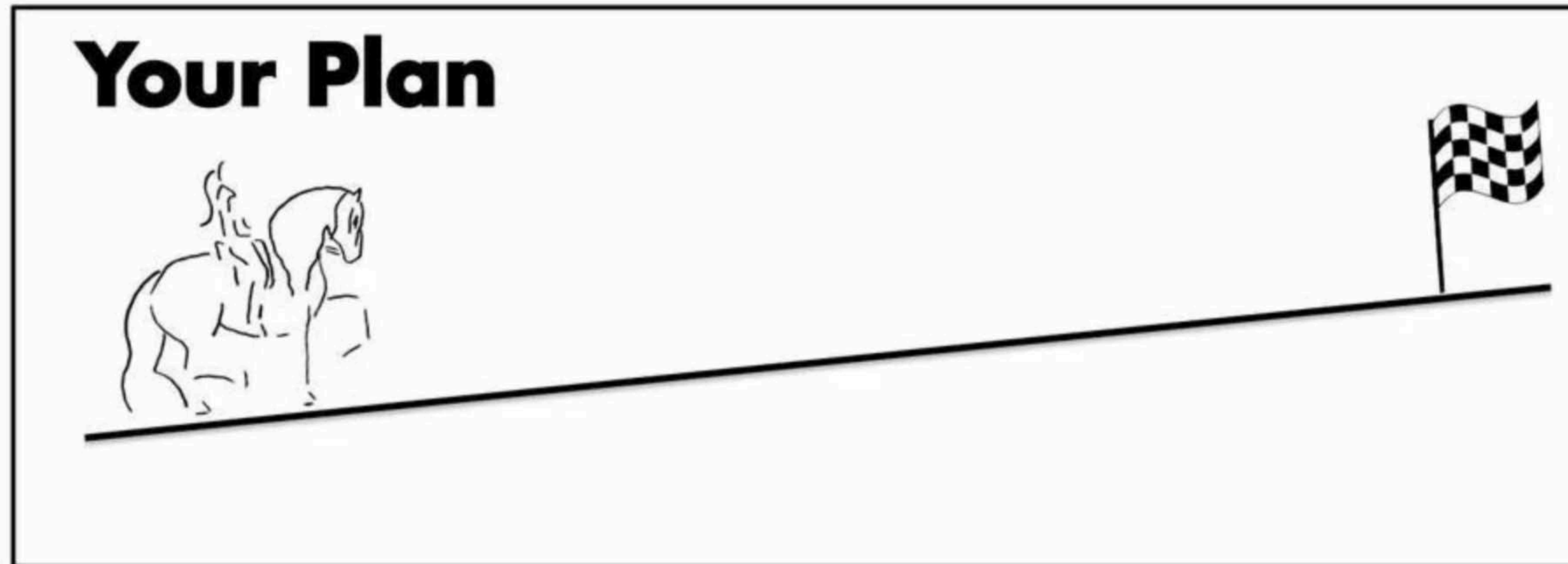




Чем может помочь RIPE Atlas?

- В России примерно 496 работающих проб
- Они распределены по всей территории страны
- Мы можем запустить измерения RTT между пробами, и построить матрицу задержки между городами
- Сравнивая задержку и расстояние, мы можем сделать вывод о географической оптимальности прокладки кабелей

Ожидания и реальность





Пререквизиты

- **Не использовались никакие “специальные возможности” сотрудника RIPE NCC для получения результатов**
 - Не уверен, что такие возможности вообще существуют
- **Измерения проводились с личного аккаунта на самостоятельно накопленные кредиты**
 - Да, их потребовалось много
- **Тем не менее, ничего бы не получилось без взаимодействия с командой RIPE Atlas**
 - Они отвечают на вопросы сообщества точно так же, как на мои!



Setup

- **3000 строк на Python3**
- **Активное использование библиотеки `ripe.atlas.cousteau` (с github)**
- **5 новых классов добавлено в эту библиотеку**
- **Найдены и исправлены командой:**
 - 1 ошибка в документации,
 - 2 ошибки в коде системы.
- **Сделано 68,147,264 измерений за полтора месяца.**



Дисклеймер

Исследование носит индикативный характер, его нельзя воспринимать как “истину в последней инстанции”:

- У нас нет проб даже в каждой автономной системе РФ, не говоря уж о каждом префиксе, поэтому мы “видим” не всё.
- Нельзя исключить, что какие-то явления носили временный характер, и уже сегодня картина поменялась.
- Несмотря на многократную проверку результатов, всё равно возможно влияние особенностей поведения Atlas-проб и узлов-целей.



Выбор точек

- **Очевидный выбор - по размеру городов**
 - Топ-30, чтобы вошел Дальний Восток (Владивосток, Хабаровск)
- **Проблема: во многих городах не очень много Atlas-проб:**
 - Красноярск: 3 пробы
 - Уфа: 2 пробы
 - Ярославль: 3 пробы
 - Махачкала: ни одной (в результате город был исключен из исследования)
- **Города образуют агломерации**
 - Например, Самара-Тольятти - 70 км



Выбор точек

- **Итого, с учетом агломераций, 26 логических “областей”:**
 - Владивосток, Волгоград, Воронеж, Екатеринбург, Ижевск, Иркутск, Казань, Краснодар, Красноярск, Москва, Нижний Новгород, Новосибирск, Омск, Оренбург, Пермь, Ростов-на-Дону, Самара, Санкт-Петербург, Саратов, Томск, Тюмень, Ульяновск, Уфа, Хабаровск, Челябинск, Ярославль
- **Проблема:**
 - С учетом тех Atlas-проб, которые не отвечают на ICMP (например, находятся за NAT) удачных измерений оказывается недостаточно.



Выбор точек

- **$RTT(IP1, IP2) = RTT(IP2, IP1)$**
 - значит, надо найти узлы в каждом регионе, которые можно “пингать” с наших Atlas-проб
- **Первичные критерии:**
 - Узел отвечает по ICMP (для поиска использовалось сканирование с помощью *ntar*)
 - Географическая “привязка” в RIPE DB
 - Соответствие прямого и обратного разрешения имени
 - Дополнительные “бонусы”:
 - IP-адрес принадлежит локальному веб-сайту
 - Географическая привязка в доменном имени
- **На этом этапе отобрано 450 узлов**
- **Итого, 310 проб и 760 (310+450) целей**



Верификация результатов

- На каждом этапе результаты проходили перекрестную верификацию, и из исследования удалялись сомнительные узлы и пробы
- Примеры:
 - Проба, у которой RTT до самой себя был 5-10 мс
 - Вебсайты налоговых (RTT всегда больше 20 мс)
 - Узел, выдающий себя за челябинский, но расположенный в Москве
- В сумме исключено 80 адресов
- Некоторые другие найденные странности:
 - Проба в MOAS (Multiple-Origin Autonomous System)
 - Две разные пробы у разных пользователей в разных концах Москвы (но у одного ISP), между которыми постоянно скачет один IP-адрес



Обработка результатов

- Между каждой парой IP-адресов мы берем минимальный RTT
- Как сводить данные между разными точками или автономными системами?
 - Минимальное значение MinRTT
 - Медианное среднее MedRTT - это среднее арифметическое центральных элементов набора:

0.2, 0.4, 0.5, 1, **1, 1**, 1, 4, 10, 25

Median: **1**

Arithmetic mean: **5.4**



Локальный RTT внутри AS

- **Вырожденный случай: RTT внутри города в пределах каждой автономной системы**
- **Важно как этап промежуточной проверки**
- **71 результат**
- **Как и ожидалось:**
 - **везде MinRTT не превышает 3 мс**
 - **MedRTT нигде не превышает 5 мс, кроме одного случая:**
 - **внутри сети Сибирьтелекома (отделение Ростелекома, AS12389) в Томске MedRTT=12.6 мс.**



RTT внутри городов

- **Мы можем обработать данные между всеми возможными IP-адресами**
 - Этот подход в какой-то мере соответствует взгляду пользователей, которые как-то распределены между операторами
 - Эти колонки таблицы ниже помечены “between IP”
- **Мы можем сначала сгруппировать адреса по автономным системам, и не учитывать RTT внутри одной автономной системе**
 - Этот подход отражает операторский взгляд на связность
 - Эти колонки таблицы ниже помечены “between ASes”

RTT внутри городов

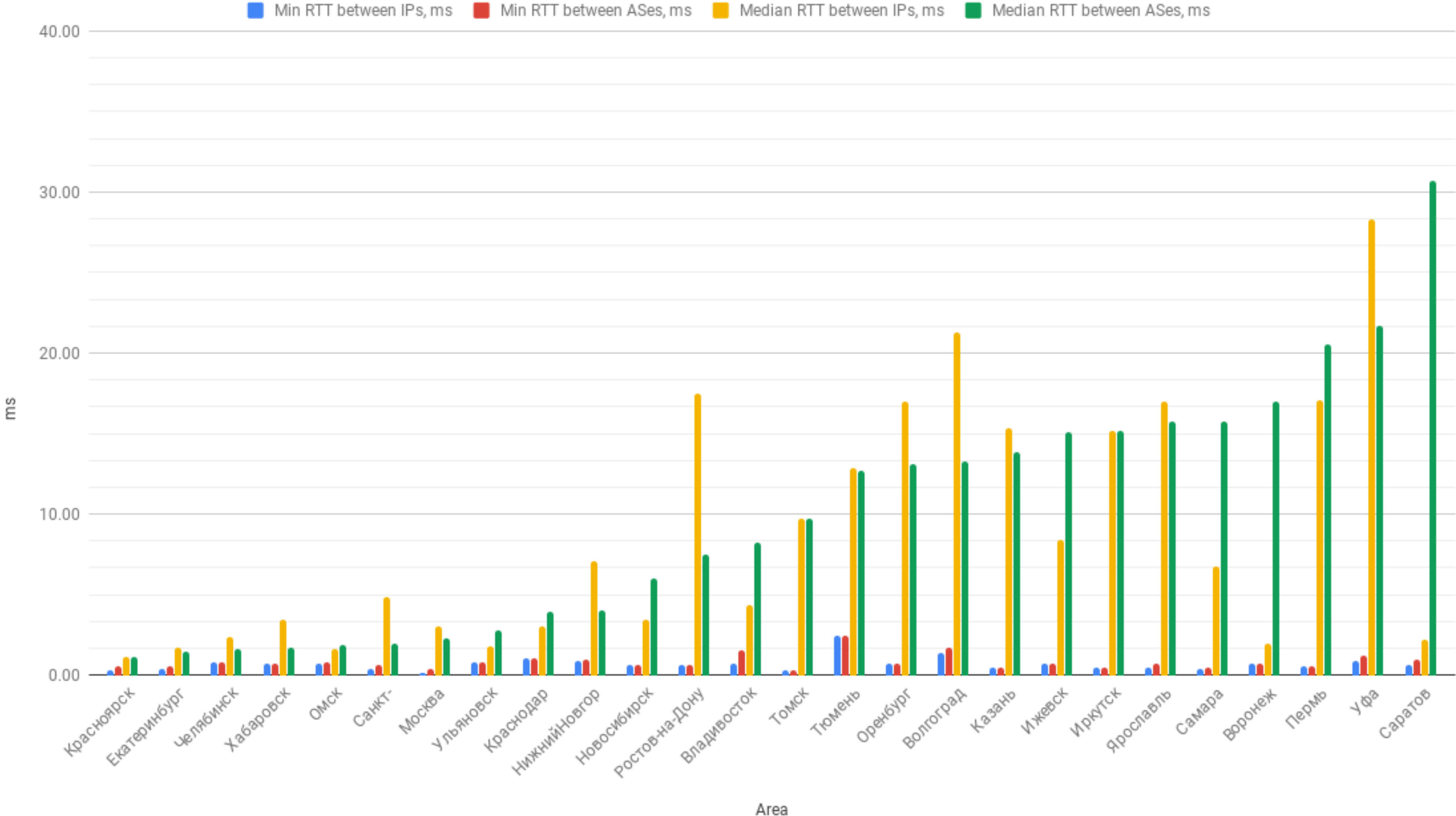


Area	Min RTT between IPs, ms	Min RTT between ASes, ms	Median RTT between IPs, ms	Median RTT between ASes, ms
Ярославль	0.49	0.75	17.00	15.71
Москва	0.18	0.41	3.03	2.28
Пермь	0.59	0.59	17.03	20.56
Новосибирск	0.60	0.60	3.41	6.00
Оренбург	0.70	0.70	16.99	13.12
Екатеринбург	0.35	0.55	1.70	1.44
Томск	0.27	0.33	9.75	9.69
Самара	0.42	0.46	6.76	15.71
Красноярск	0.32	0.56	1.09	1.13
Нижний Новгород	0.90	0.96	7.04	4.01
Ульяновск	0.77	0.77	1.82	2.80
Санкт-Петербург	0.39	0.66	4.84	1.93
Челябинск	0.83	0.83	2.33	1.60
Волгоград	1.38	1.72	21.26	13.26
Ижевск	0.71	0.71	8.38	15.07
Хабаровск	0.69	0.69	3.42	1.68
Иркутск	0.43	0.43	15.17	15.17
Тюмень	2.49	2.49	12.85	12.67
Владивосток	0.70	1.52	4.39	8.24
Воронеж	0.68	0.68	1.99	16.99
Саратов	0.66	0.95	2.24	30.66
Уфа	0.88	1.22	28.28	21.66
Омск	0.70	0.83	1.66	1.89
Ростов-на-Дону	0.63	0.63	17.48	7.47
Казань	0.51	0.51	15.36	13.81
Краснодар	1.05	1.05	3.03	3.92

RTT внутри городов



RTT внутри городов





RTT внутри городов: выводы

- В целом ситуация выглядит очень хорошо:
 - MinRTT везде < 3 мс
 - Подтверждение гипотезы, что $d_2 \approx 0$
 - MedRTT между автономными системами
 - в 10 случаях (40%) не превышают 5 мс
 - только в одном случае превышает 25 мс
 - всегда меньше 50 мс
- Тем не менее, есть куда улучшить ситуацию - в городах, где **MedRTT > 10 мс**
 - По крайней мере, в Перми, Саратове, Уфе
 - Возможно, стоит подумать о создании или развитии IXP?



RTT между городами

- Скорость света в волокне равна 60-70% скорости света в вакууме, т.е.

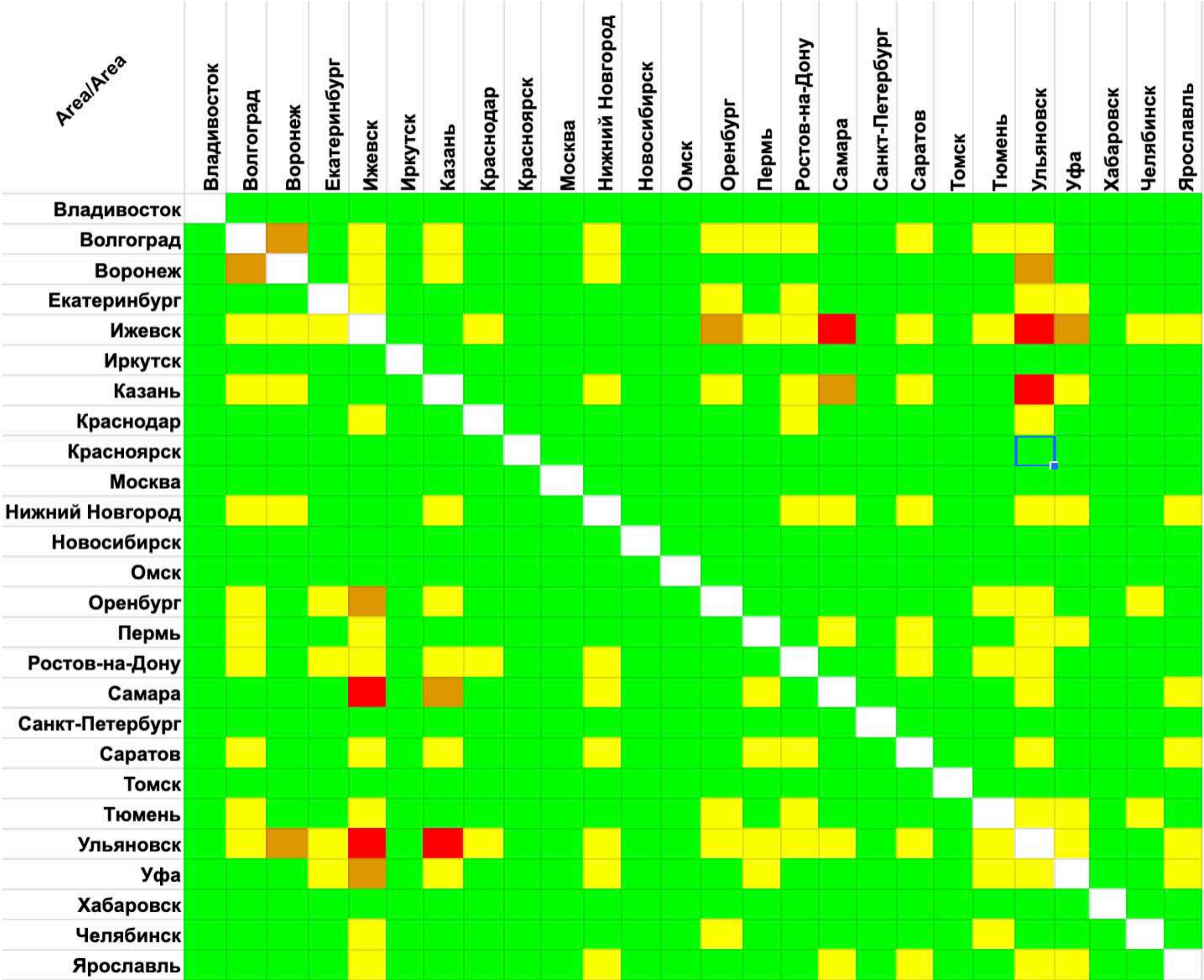
$$c_{\text{fibre}} \approx 200 \frac{\text{ms}}{\text{km}}$$

- Поэтому в случае идеально прямого волокна соотношение

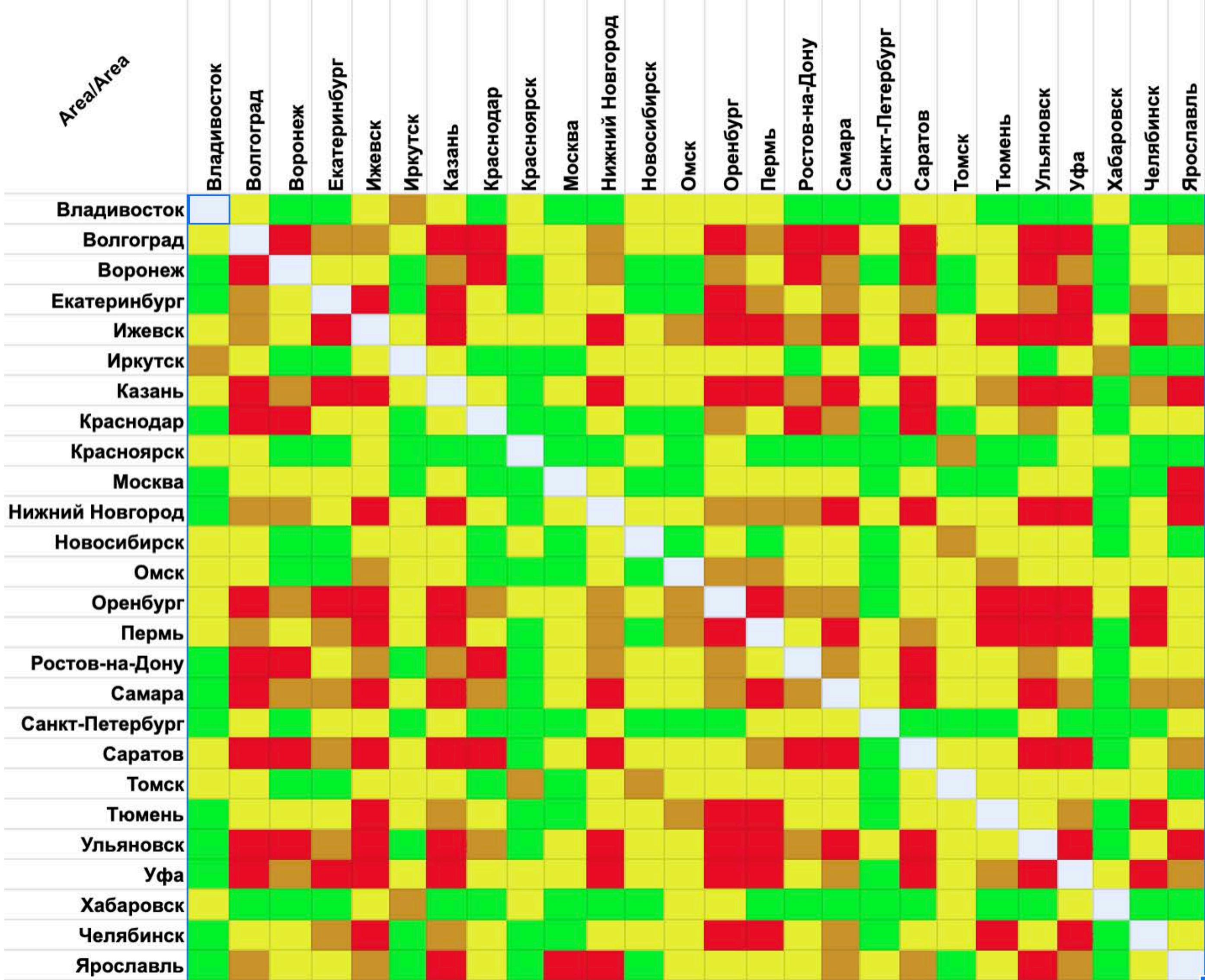
$$S = 100 \frac{RTT, \text{ms}}{\text{distance}, \text{km}} \text{ будет равно } 1.$$

- Вычислим S для MinRTT и MedRTT и построим “тепловые карты”

MinRTT между городами



MedRTT между городами





RTT между городами: наблюдения

- В целом ситуация тоже очень хорошая
 - minRTT - почти везде S близка к идеальной
 - medRTT - в 70% $S < 3$, что неплохо
- Есть несколько “патологических” значений medRTT, на которые явно необходимо обратить внимание:
 - Ижевск-Казань, $S=11.1$
 - Казань-Самара, $S=11.4$
 - Ижевск-Пермь, $S=15.5$
 - Казань-Ульяновск, $S=15.9$



RTT между городами: наблюдения

	Владивосток	Волгоград	Воронеж	Екатеринбург	Ижевск	Иркутск	Казань	Краснодар	Красноярск	Москва	НижнийНовгород	Новосибирск	Омск	Оренбург	Пермь	Ростов-на-Дону	Самара	Санкт-Петербург	Саратов	Томск	Тюмень	Ульяновск	Уфа	Хабаровск	Челябинск	Ярославль
Всего <2	13	1	8	7	0	10	2	9	17	11	3	9	7	1	3	4	3	15	3	6	5	4	2	17	6	6
Всего от 2 до 3	11	10	7	8	10	13	8	9	7	13	10	15	14	9	9	11	7	10	9	17	13	7	9	7	11	10
Всего свыше 3	1	14	9	9	14	1	14	6	1	1	11	1	4	14	12	9	14	0	12	2	7	13	13	1	8	8

5 лидеров:

1. Красноярск
1. Хабаровск
2. Санкт-Петербург
3. Владивосток
4. Москва

10 отстающих:

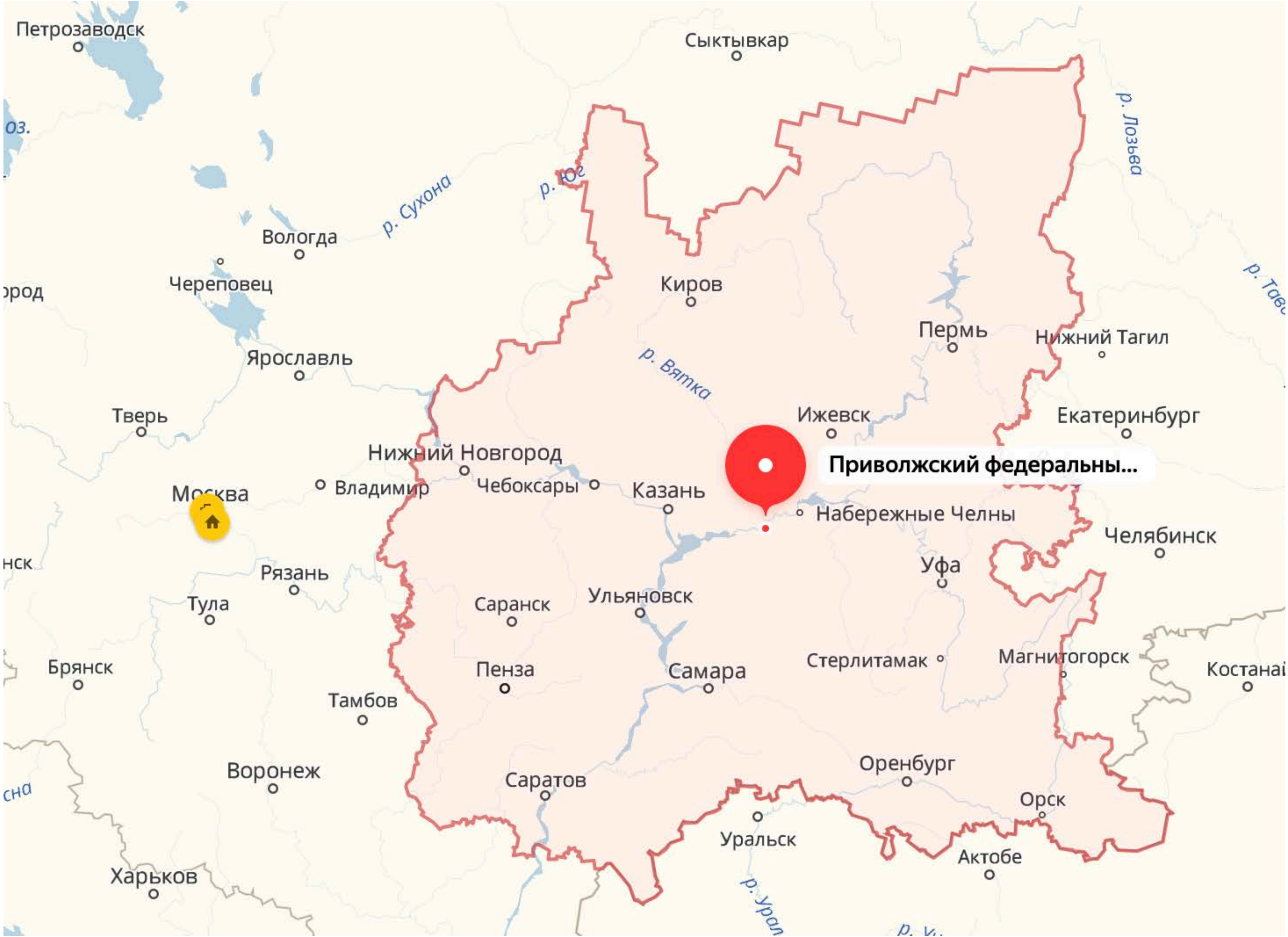
1. Волгоград
1. Ижевск
1. Казань
1. Оренбург
1. Самара
2. Ульяновск
2. Уфа
3. Пермь
3. Саратов
4. Нижний Новгород



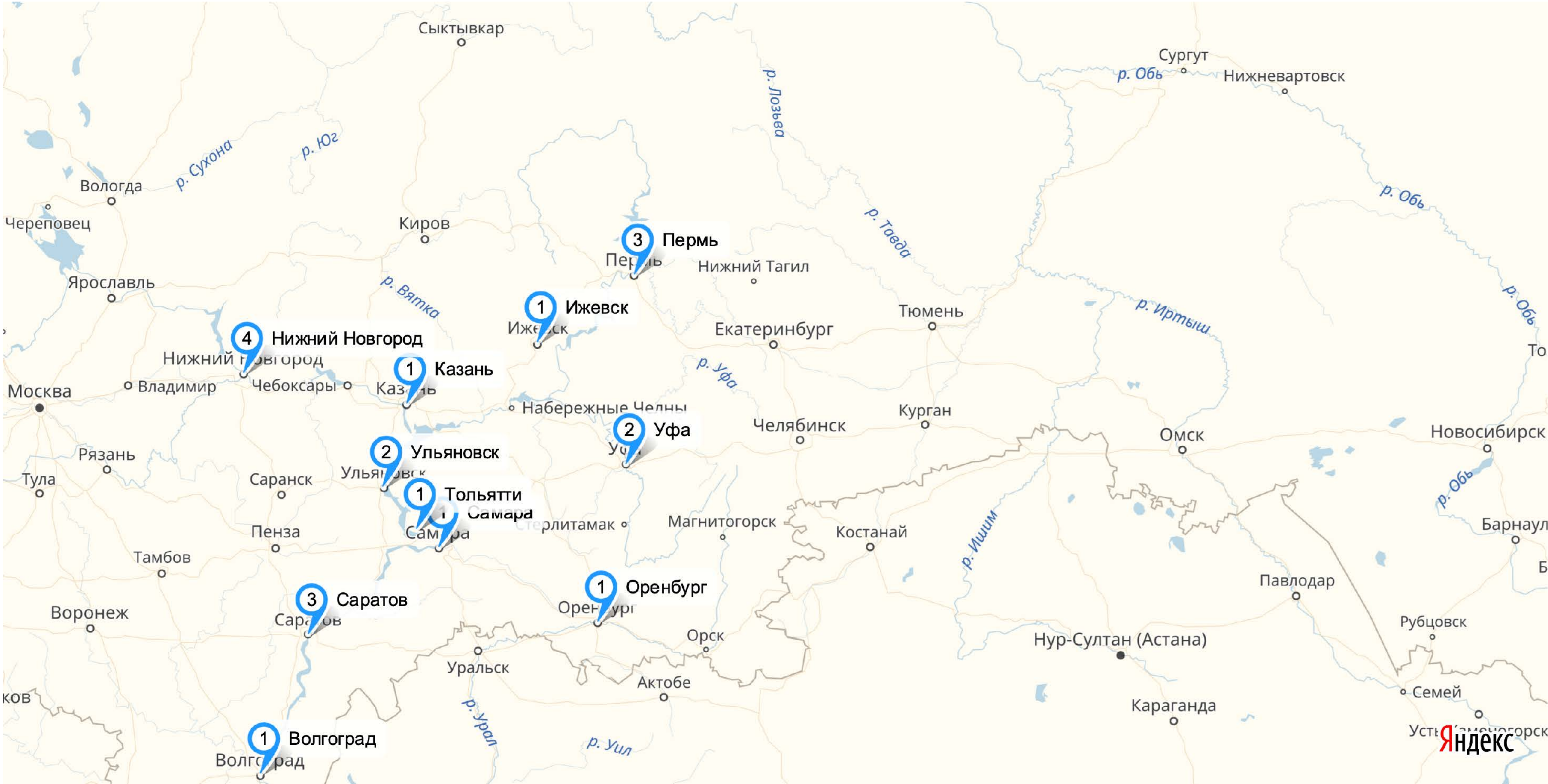
RTT между городами: секуундочку...

- Все города из второго списка находятся в одном географическом регионе!
- Два возможных (не исключających друг друга) объяснения:
 - Политико-административное: Поволжский Федеральный Округ
 - Географическое: реки Кама и Волга (самая крупная река Европы!)
- В целом, похоже, это всё - всего одна региональная проблема (отклонения от нормы, описанные ранее выше - тоже в этом же регионе!)

Политика



География





Мост через Волгу в Ульяновске (фото с сайта www.wikiwand.com)



Выводы

- Система RIPE Atlas дает возможность исследовать RTT в больших масштабах
- Внутри городов отсутствуют существенные проблемы связности
 - Исключение - один оператор (Сибирьтелеком) в одном городе (Томск)
- Показатели “прямоты” прокладки кабелей между городами в среднем очень хорошие
 - Исключение составляет только Поволжский федеральный округ — возможно, в силу наличия очень большой естественной водной преграды



Что дальше?

- Провести более глубокий анализ уже полученных результатов
- Внести улучшения в код (больше читаемости, меньше ручной работы в будущем)
- Автоматизировать больше работы по верификации результатов
- Провести аналогичные измерения для других стран
- На выверенных датасетах сделать эти измерения регулярными, а результаты - публичными



Acknowledgements

- Эмилю Абену - за инструмент IXP Country Jedi, результаты которого вдохновили меня на данное исследование
- Команде RIPE Atlas за разработанный инструментарий
- Виктору Наумову за консультации и помощь по ходу проекта



Questions

