

Балансировка сетевого трафика

Чемерицкий Евгений Викторович

к.ф.-м.н.

Балансировка нагрузки

в контексте компьютерных сетей

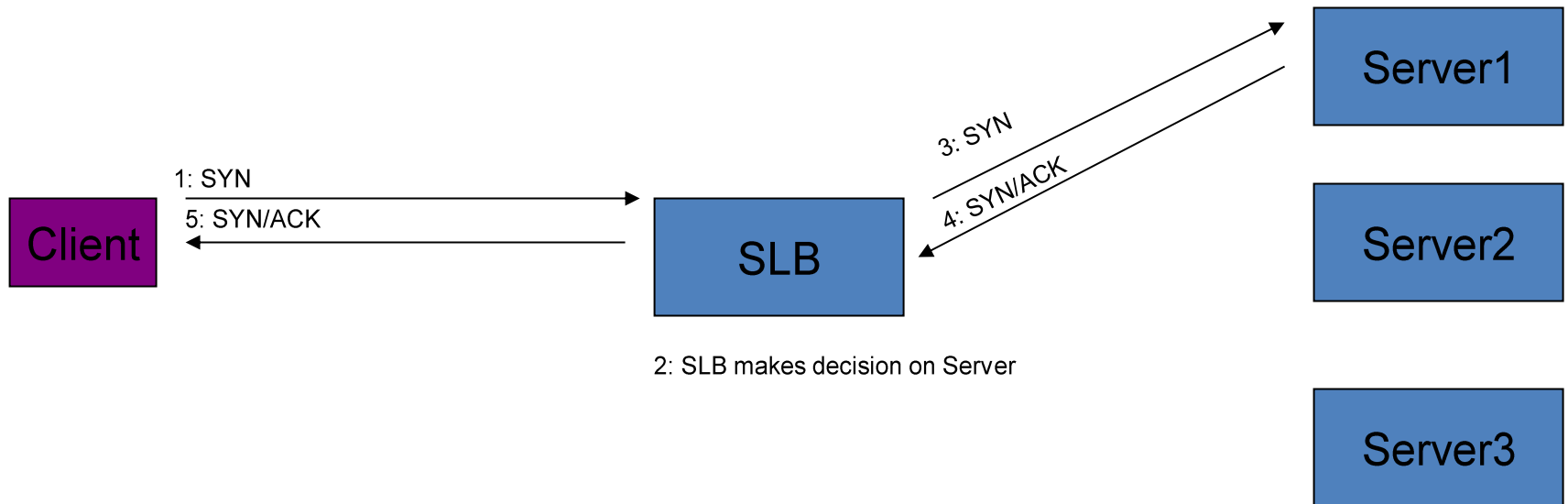
- Метод распределения трафика между несколькими хостами, линиями связи и коммутирующими устройствами
- Цели балансировки
 - Повышение степени утилизации ресурсов
 - Повышение пропускной способности
 - Уменьшение времени отклика системы
 - Борьба с отказами и перегрузками

Балансировка нагрузки

классификация

- По устройствам и по времени
 - Burstable – способ оплаты трафика, когда расчёты исходят из фактически потреблённой полосы в рамках выделенной
- Статическая и динамическая
 - Учёт нагрузки vs простота реализации
- Централизованная и распределённая
 - Выделенный контроль за балансировкой

Server Load Balancing (SLB)



Стратегии балансировки NGINX

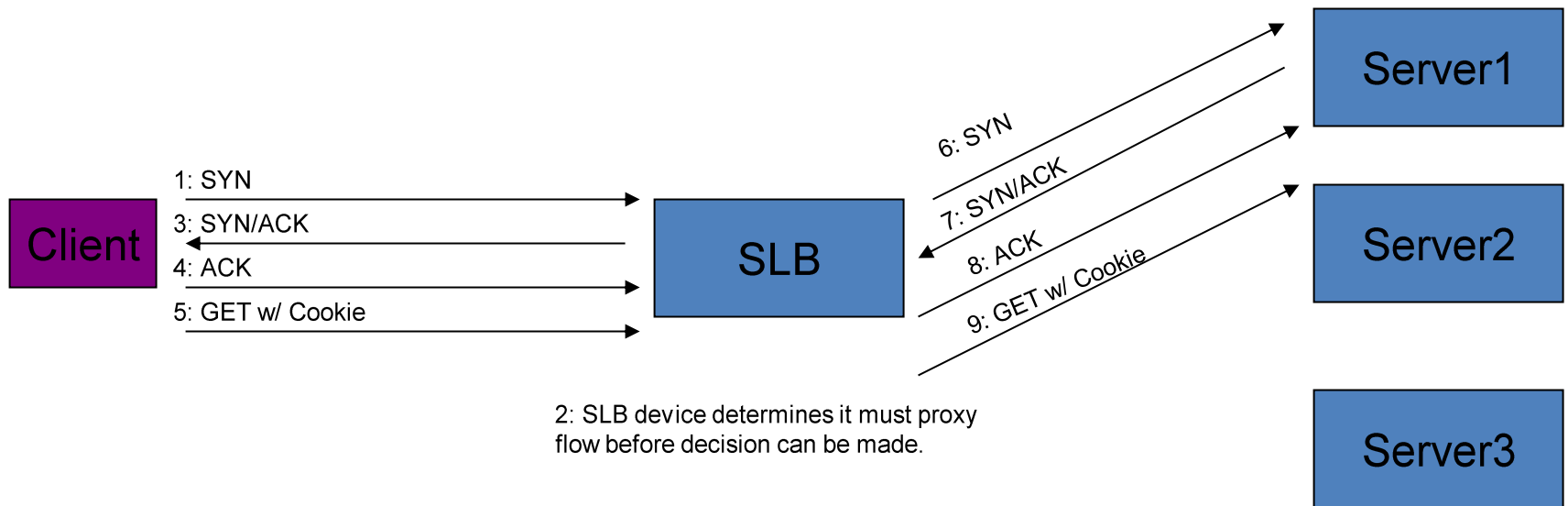
Статические:

- Round Robin
- Hash-based
 - Разные методы расчёта

Динамические:

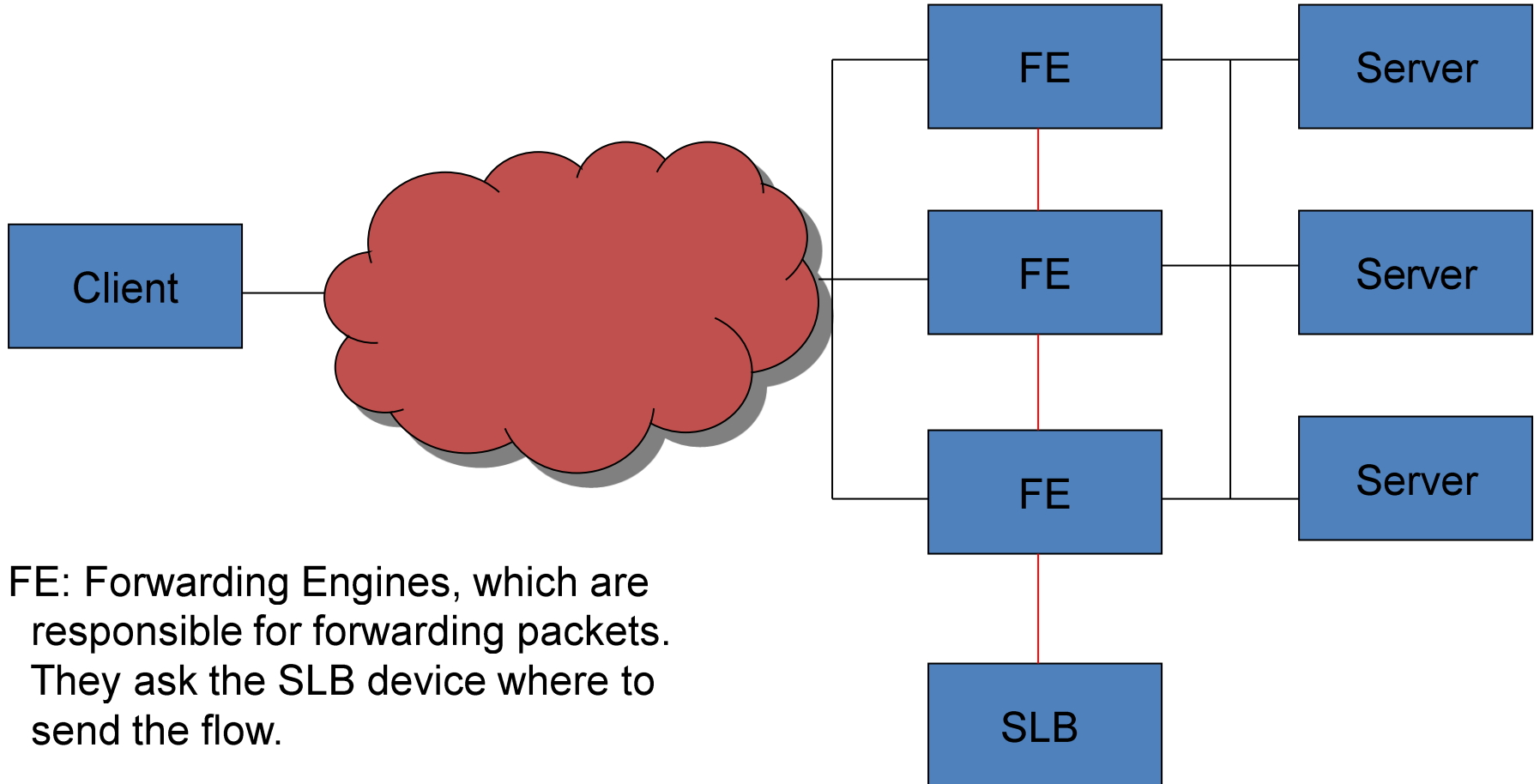
- Least Connections
 - Выбирается интерфейс, обслуживающий наименьшее количество соединений
- Lowest Response Time
 - Выбирается интерфейс с наименьшим временем завершения запроса

Балансировка на 5+

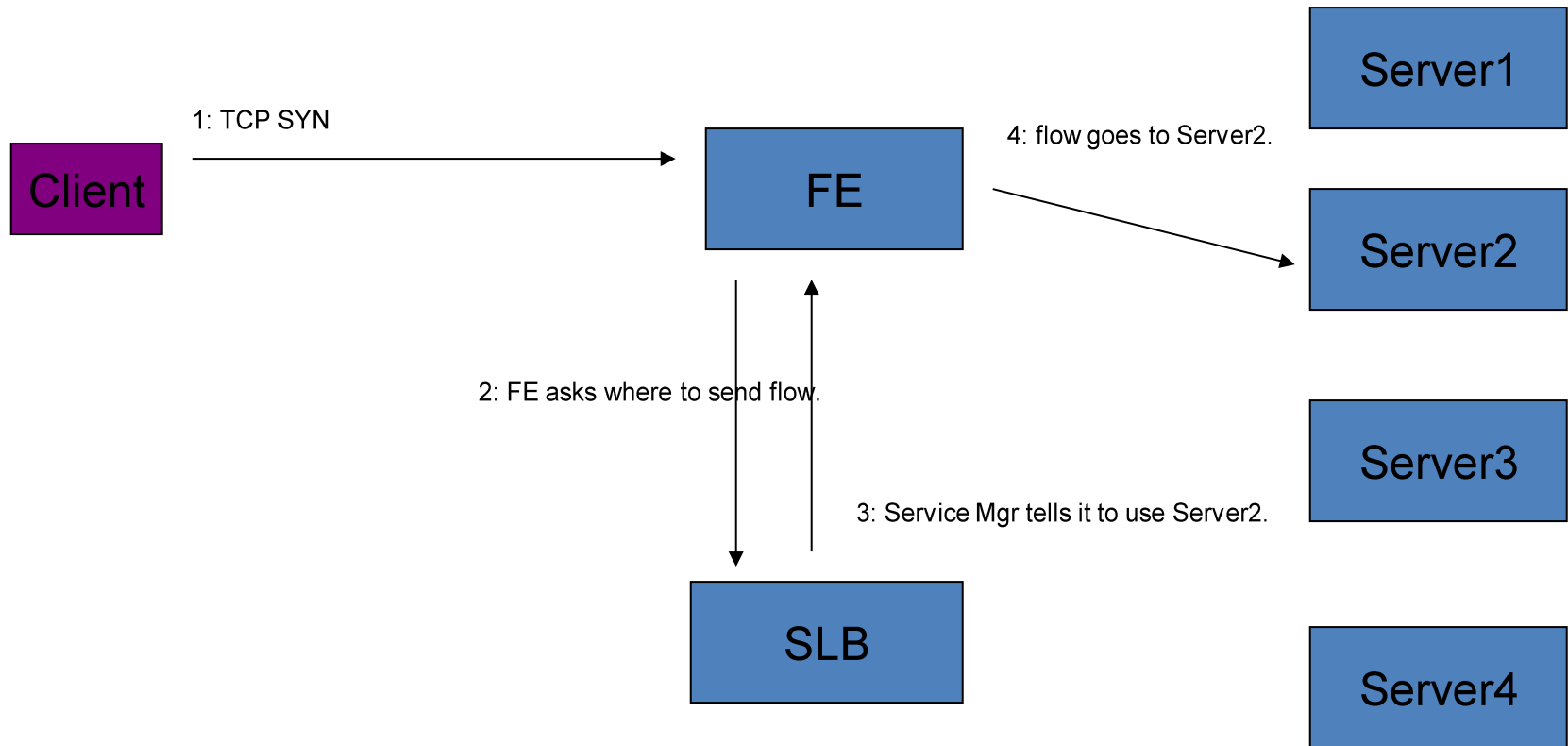


При работе на L5 балансировщик не может определить,
куда перенаправлять поток данных заранее =>
Он должен взять часть работы сервера на себя

Распределённый балансировщик



Распределённый балансировщик



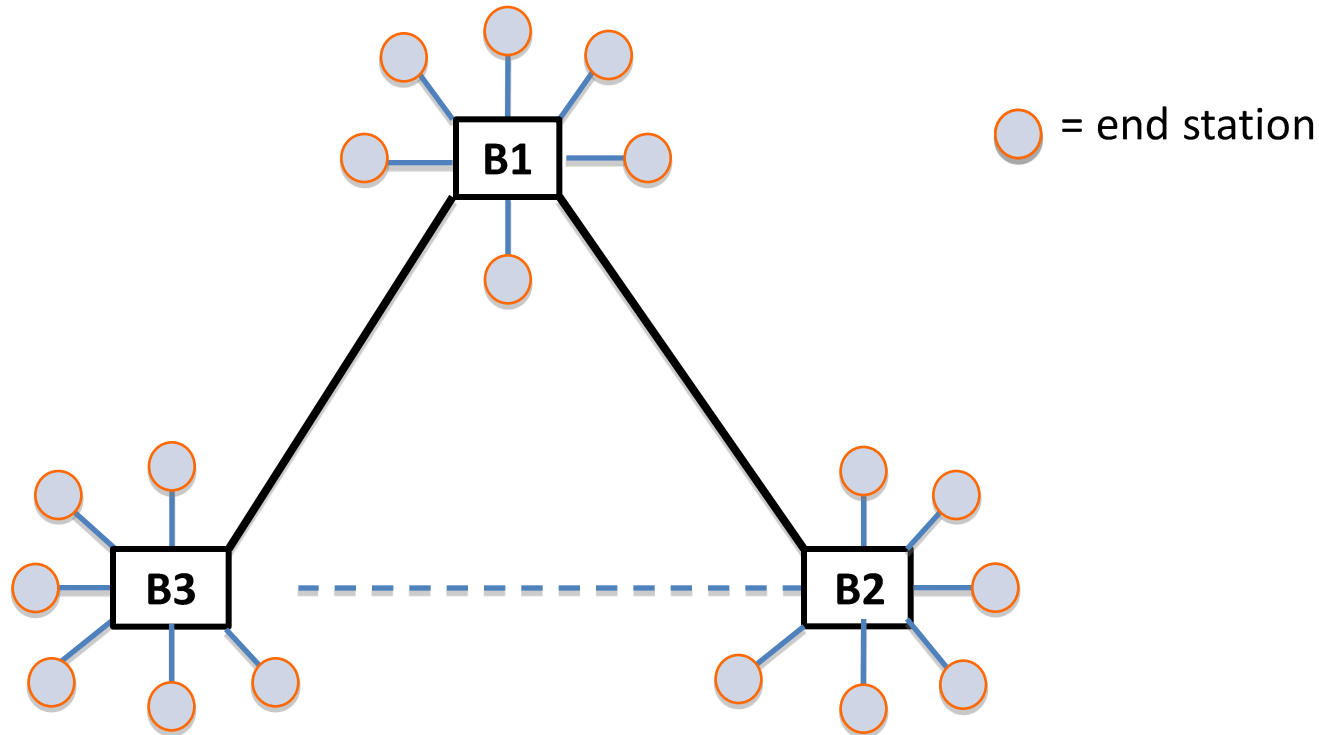
Subsequent packets flow directly from Client to Server2 thru the FE.
The FE must notify the SLB device when the flow ends.

Балансировка трафика по нескольким маршрутам

Недостатки передачи данных по одному пути:

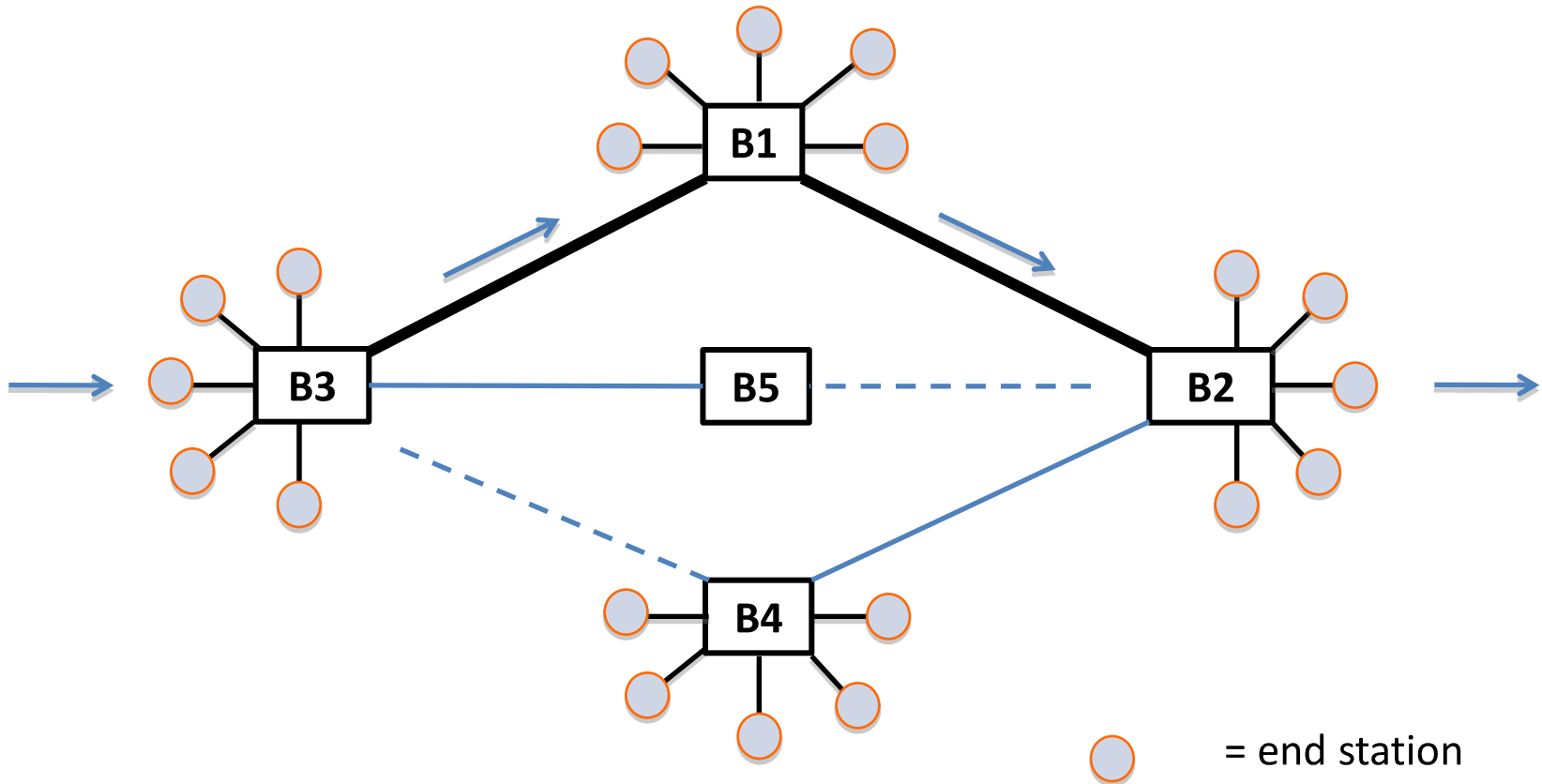
- Ограничение пропускной способности сети
 - Маршрутизация по одному пути не позволяет достигнуть оптимума функции утилизации
 - Производительность в угоду простоте
- Подверженность колебаниям
 - Ранний адаптивный протокол маршрутизации, который учитывал задержку на линиях связи оказался нестабильным

Unicast Least Cost Paths



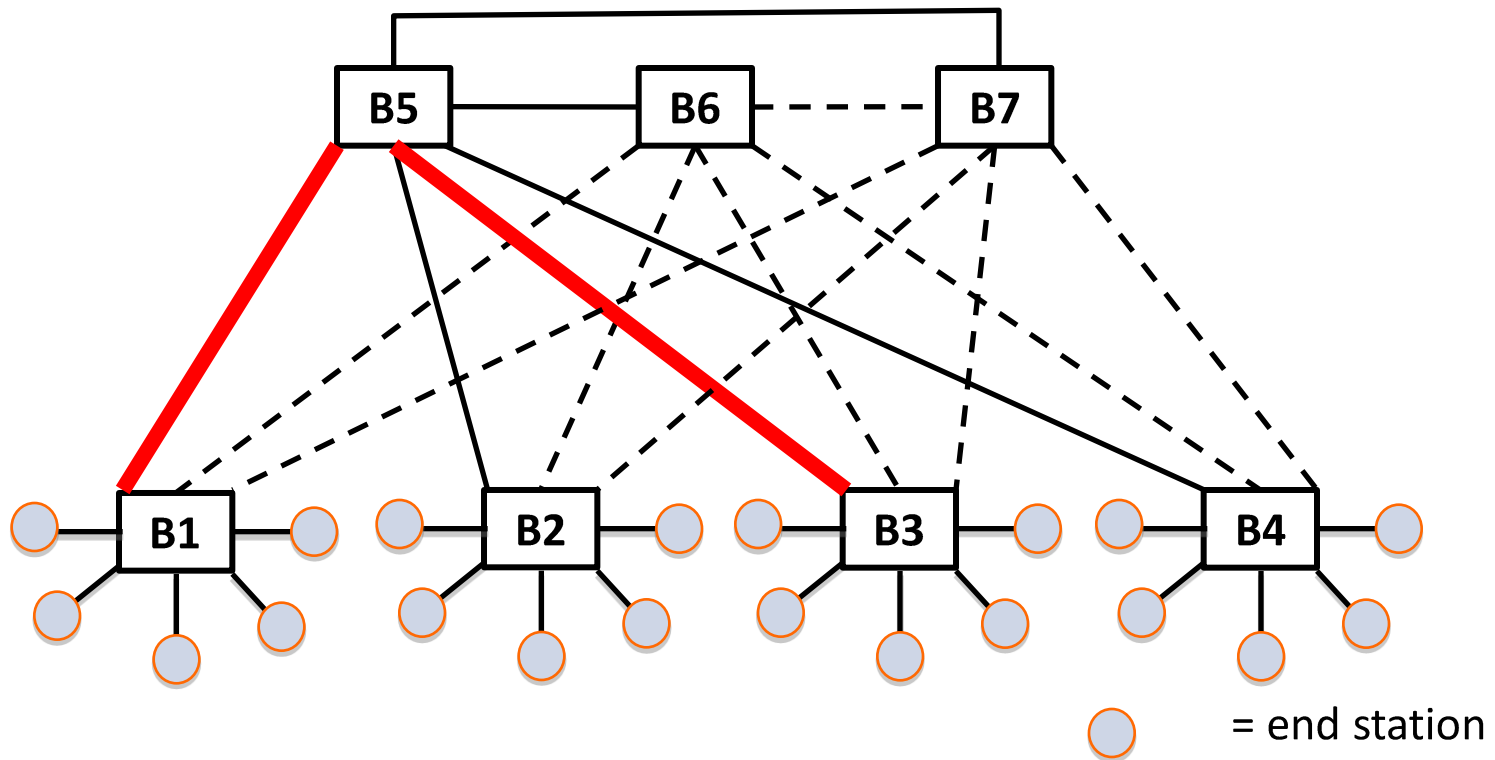
Spanning tree eliminates loops
by disabling ports

Unicast Multi-Pathing



Bridges limit traffic to one path

Multi-Pathing (Unicast)



Bridges limit traffic to one path

Протокол TRILL



Протокол TRILL

- Основан на простой идее:
 - Добавить к передаваемым пакетам новый L2 заголовок с hop-counter'ом
 - Маршрутизировать пакеты с помощью IS-IS
 - Удалить дополнительные заголовки перед тем, как доставить пакет
- Преимущества:
 - Передача по лучшему пути
 - Встроенная поддержка Equal Cost Multi Path
 - Поддержка передачи на несколько назначений

Предпосылки развития многопоточной маршрутизации

- Центры обработки данных
- Беспроводные среды передачи данных
- Устройства с несколькими беспроводными сетевыми интерфейсами
- Multihoming и гетерогенные среды
- Альтернативные пути внутри ISP

30- 80% трафика идёт неоптимальным маршрутом
40% трафика проходит через балансировщики

Меняется парадигма организации вычислений

- На смену клиент-сервисной архитектуре пришли Cloud Computing и концепция Software as a Service
- При высокой концентрации вычислительных ресурсов сеть становится узким местом



Мобильная революция



На каждого человека приходится по несколько мобильных устройств

Появились сервисы, ориентированные на мобильные устройства

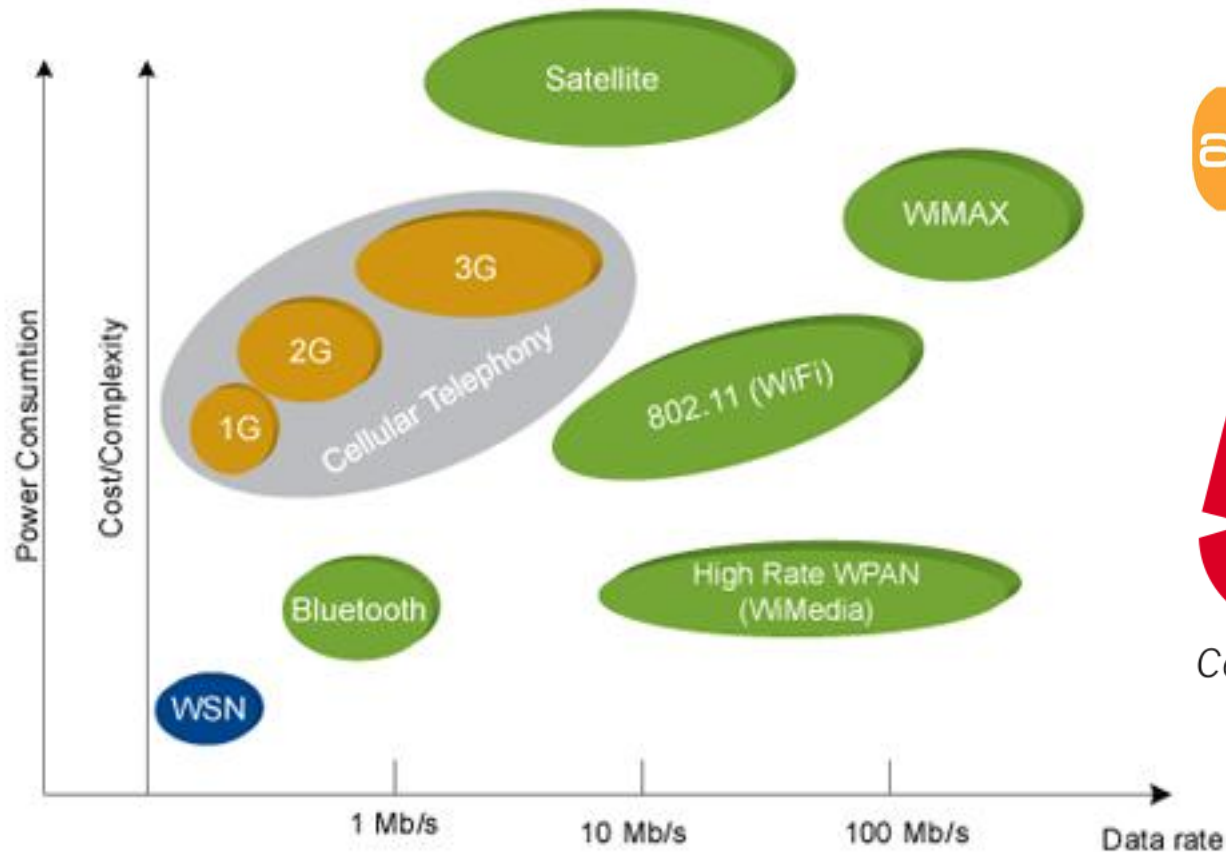


Intel встроила модуль WIFI в процессор

Выпущена версия Ubuntu для замены рабочей станции

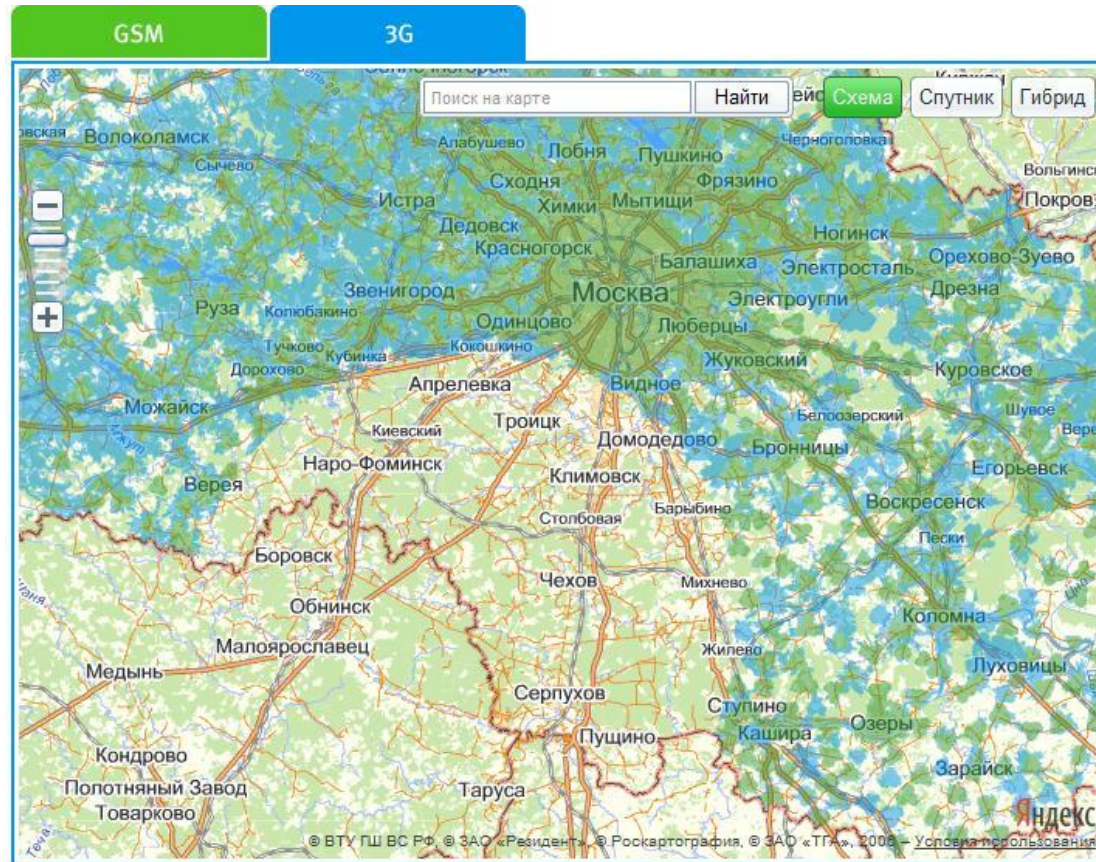


Постоянно появляются новые wireless технологии



Быстрее, дальше, дешевле!

Для wireless устройств не хватает свободных частот

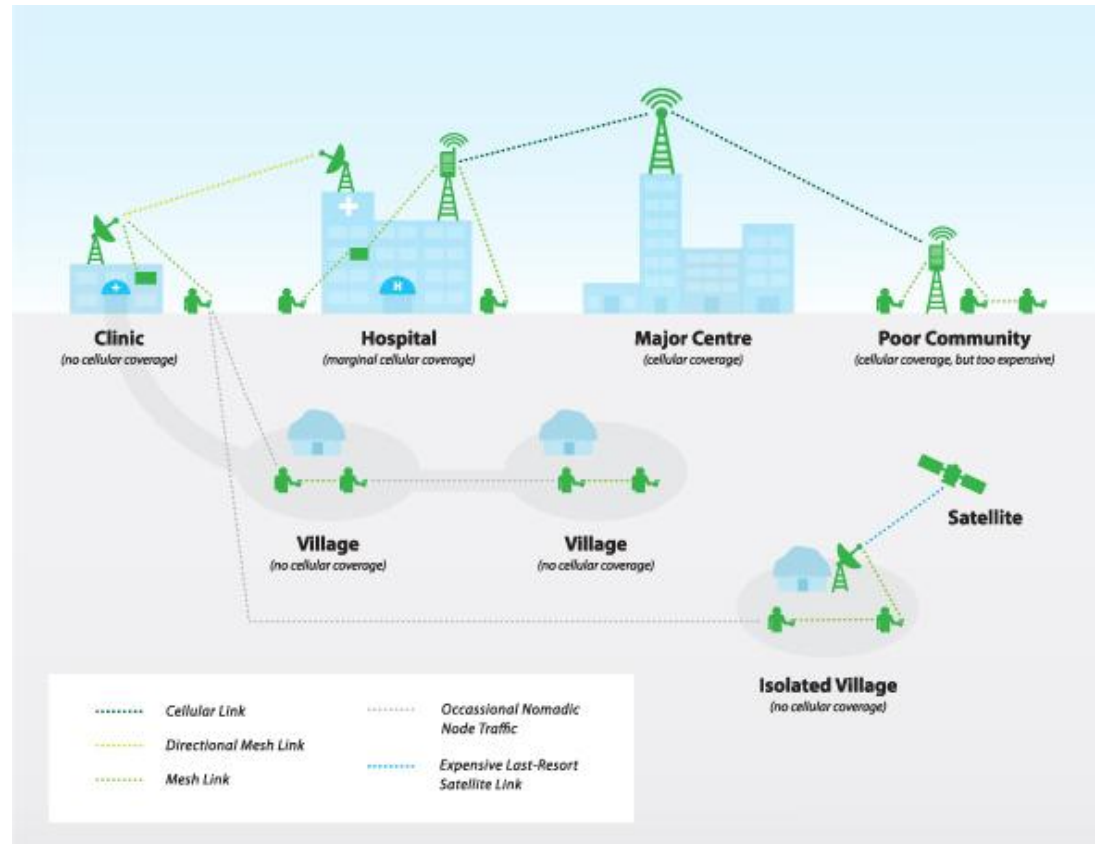


Зона покрытия сети Мегафон

Сети мобильных устройств

[Serval Project]

Идея – связать
мобильные
устройства с
базовыми
станциями
транзитивно



Увеличивает зону покрытия

Не может увеличить пропускную способность

При увеличении числа пользователей пропускная способность разделяется

Решение – увеличение
количества передающих
станций

wimax



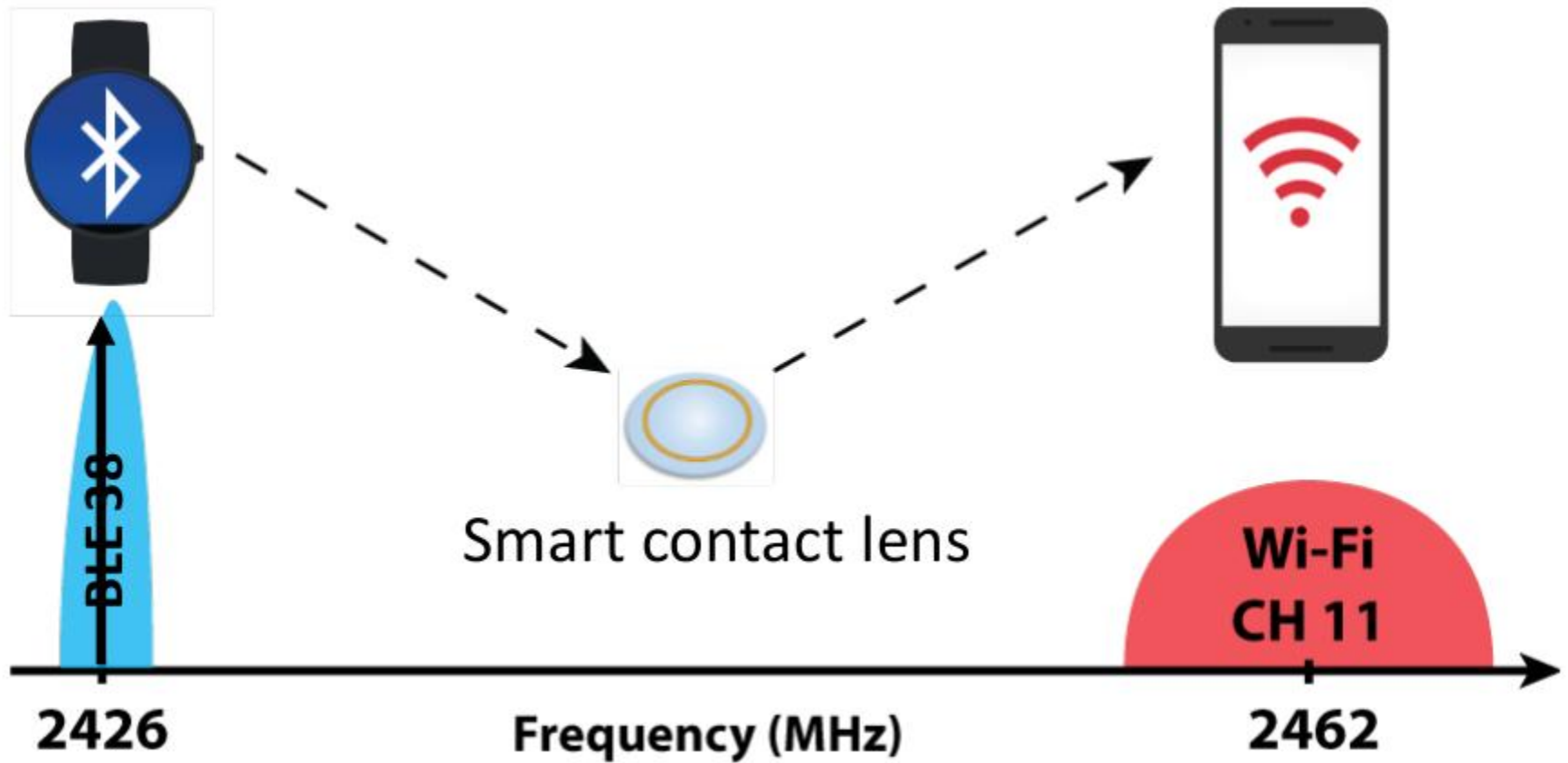
Слишком дорого!

WiFi

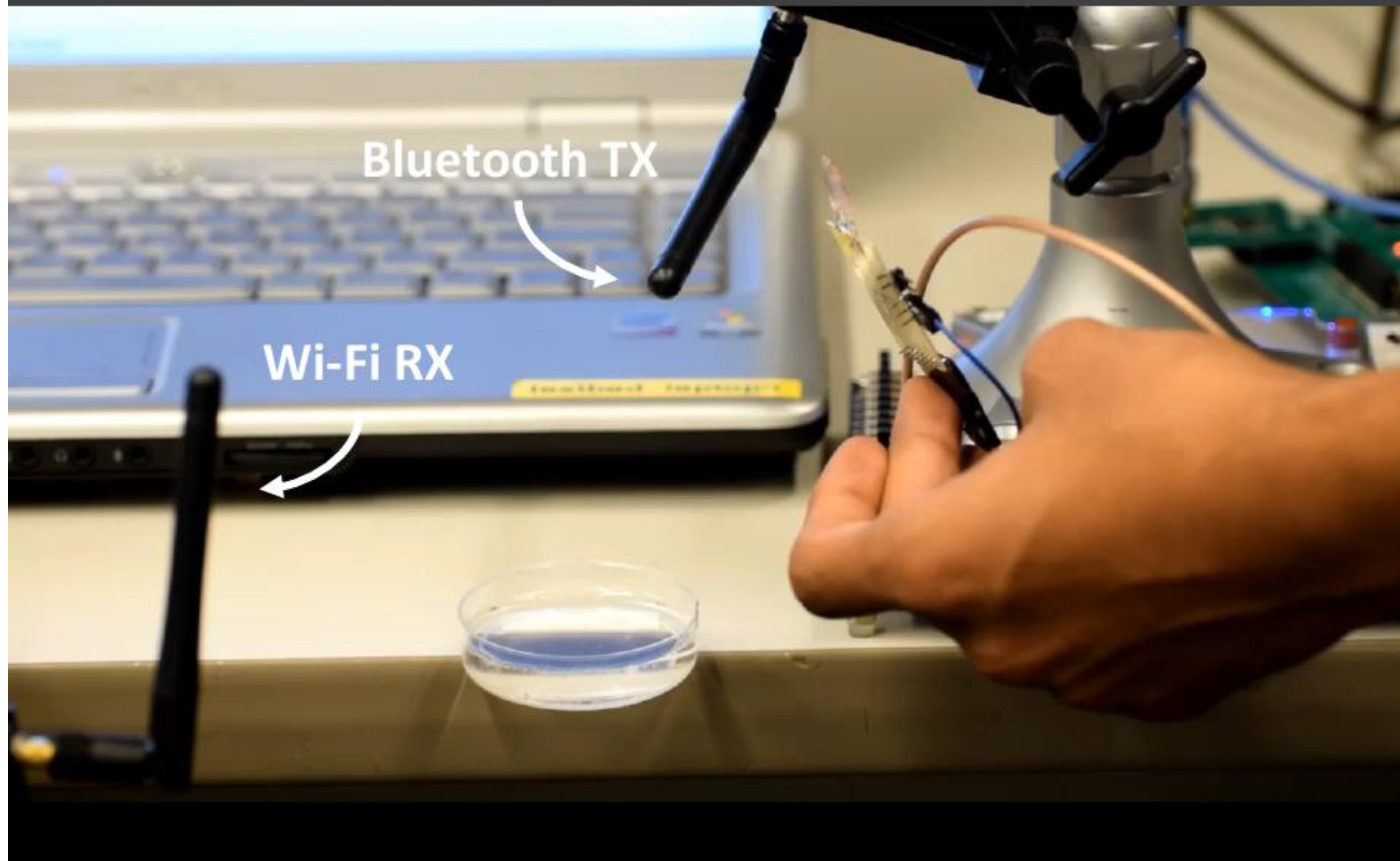


[Почему нельзя мультиплексировать
передачу данных через разные каналы?!]

Interscatter Communication



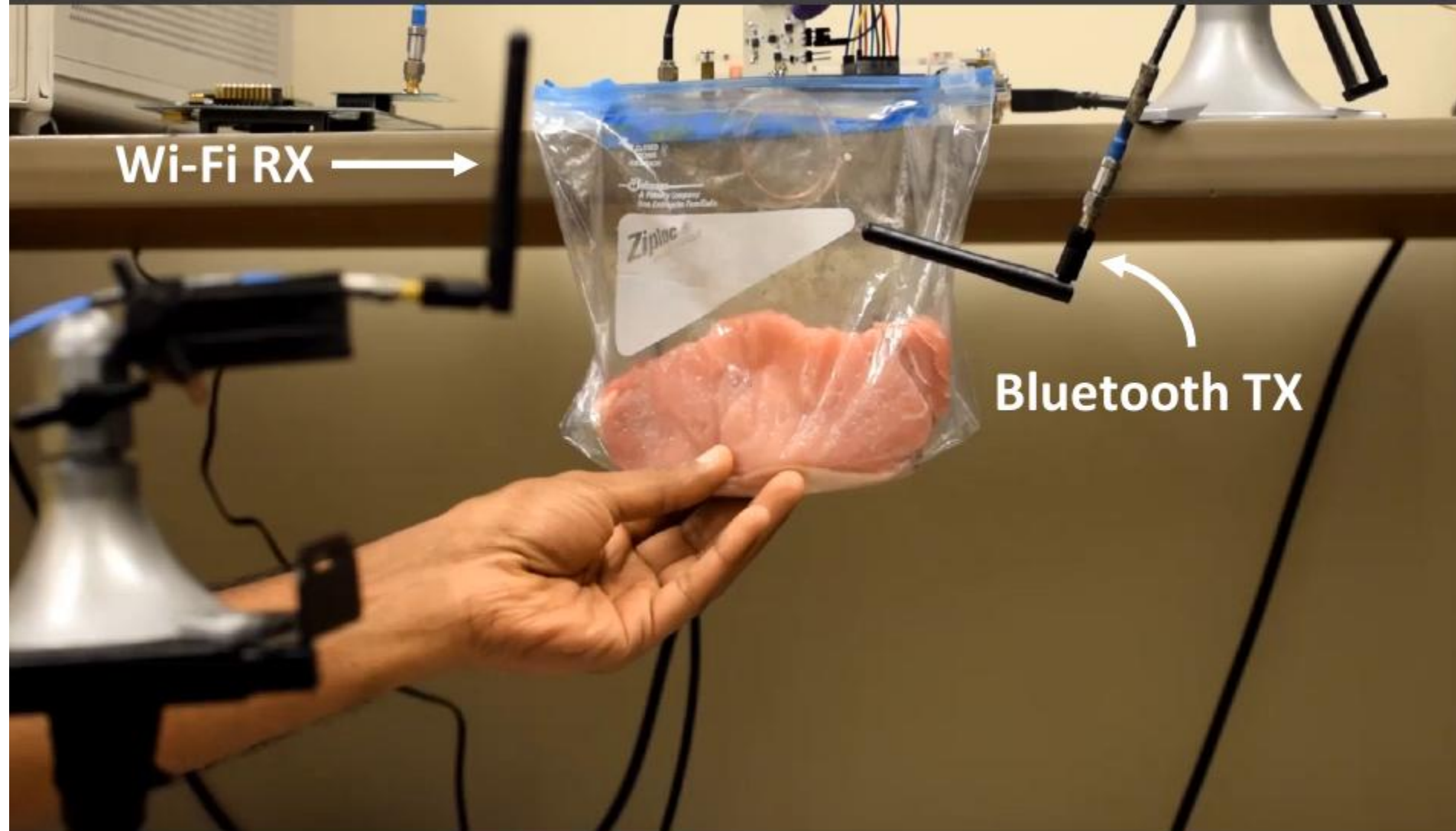
Contact Lens Prototype



BMI Prototype

Wi-Fi RX →

Bluetooth TX



Diversity of radio architectures

Active:
Symmetric
Radio

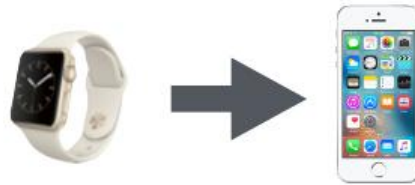


WiFi/
Bluetooth



Bluetooth®

Backscatter:
Low power
transmitter



RFID Tag



Passive:
Low power
receiver



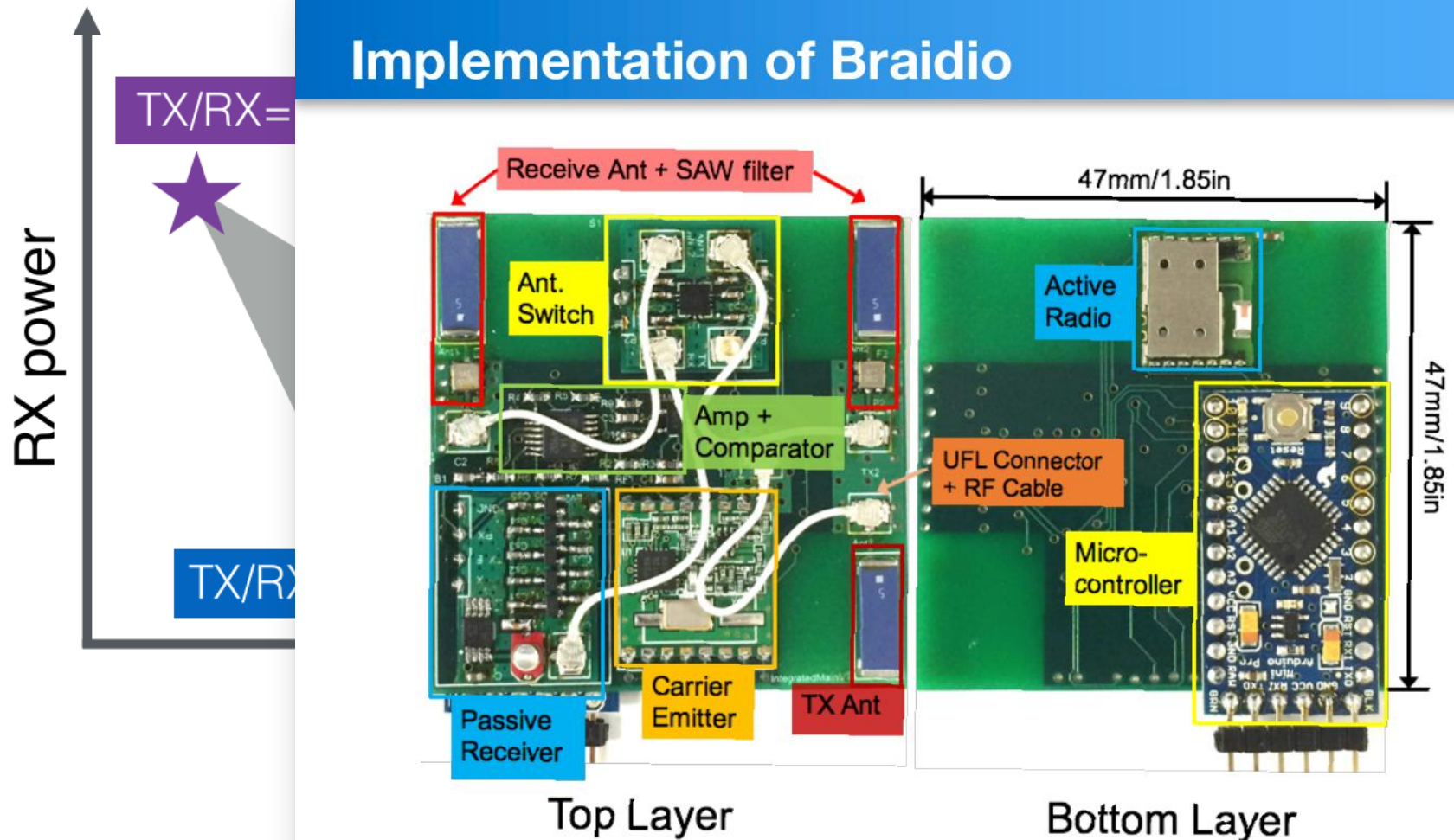
AM receiver



Architecture of radios

Achievable
region

Implementation of Braidio



Braidio: Performance gain over active radio

Receiver

Nike Fuel Band	1.43	2.45	3.51	6.63	24.7	39.1	49.1	251	276	350
Pebble Watch	2.57	1.43	1.76	2.97	9.98	15.5	19.4	97.7	107	136
Apple Watch	3.68	1.85	1.43	2.11	6.51	10.0	12.4	61.6	67.9	85.8
Pivthead	6.97	3.12	2.21	1.43	3.45	5.12	6.29	29.8	32.8	41.4
iPhone 6S	25.9	10.4	6.1	3.12	1.43	1.43	1.43	8.64	10.7	13.6
iPhone 6 Plus	41.0	16.3	10.4	6.1	3.12	1.43	1.43	5.65	6.99	8.64
Nexus 6P	51.6	20.4	13.6	8.64	5.12	3.12	1.43	4.61	5.68	6.99
Surface Book	263	102	64.7	31.3	8.29	5.44	4.46	1.43	1.43	1.63
MacBook Pro 13	290	112	71.3	34.4	9.07	5.94	4.85	1.50	1.43	1.54
MacBook Pro 15	368	143	90.1	43.4	11.3	7.34	5.96	1.71	1.62	1.43

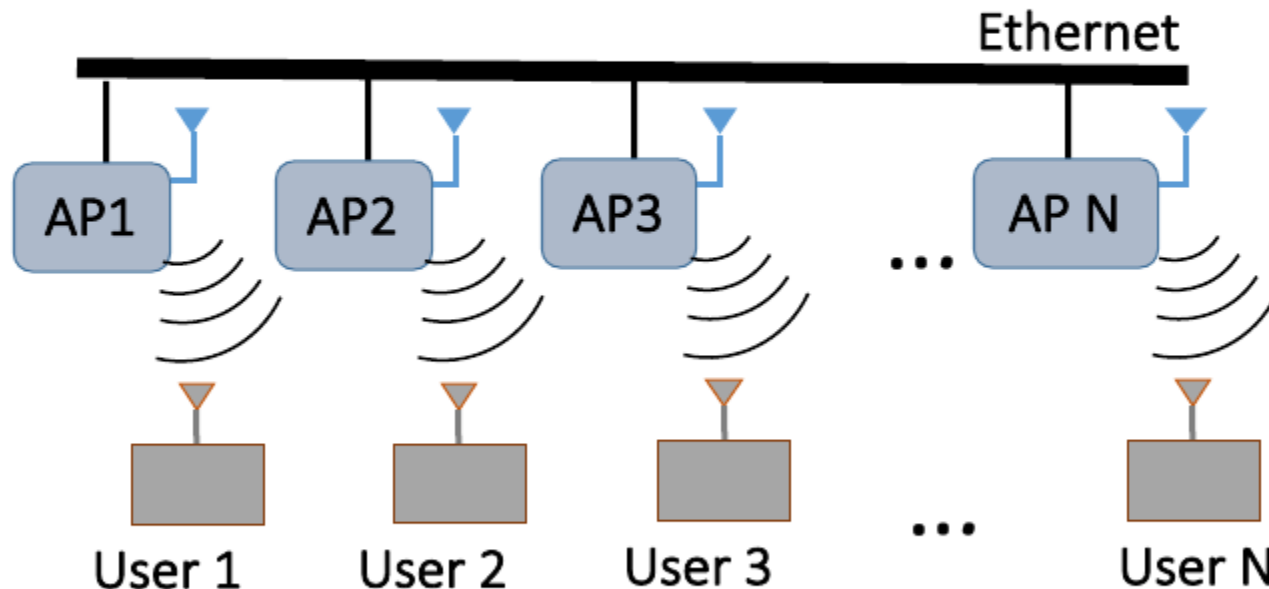
Transmitter

>300x improvement
when fitness band
transmits to laptop

Real-time Distributed MIMO Systems

Distributed MIMO is the Holy Grail

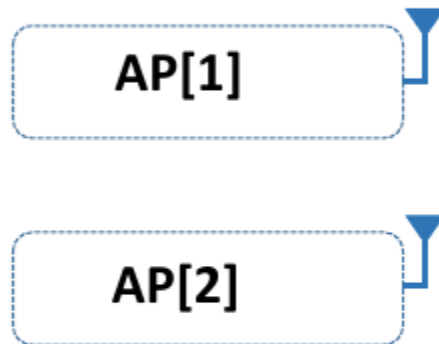
Distributed protocol for APs to act as a huge MIMO transmitter with sum of antennas



N APs $\rightarrow$ N times higher throughput

Real-time Distributed MIMO Systems

What happens with Distributed MIMO?

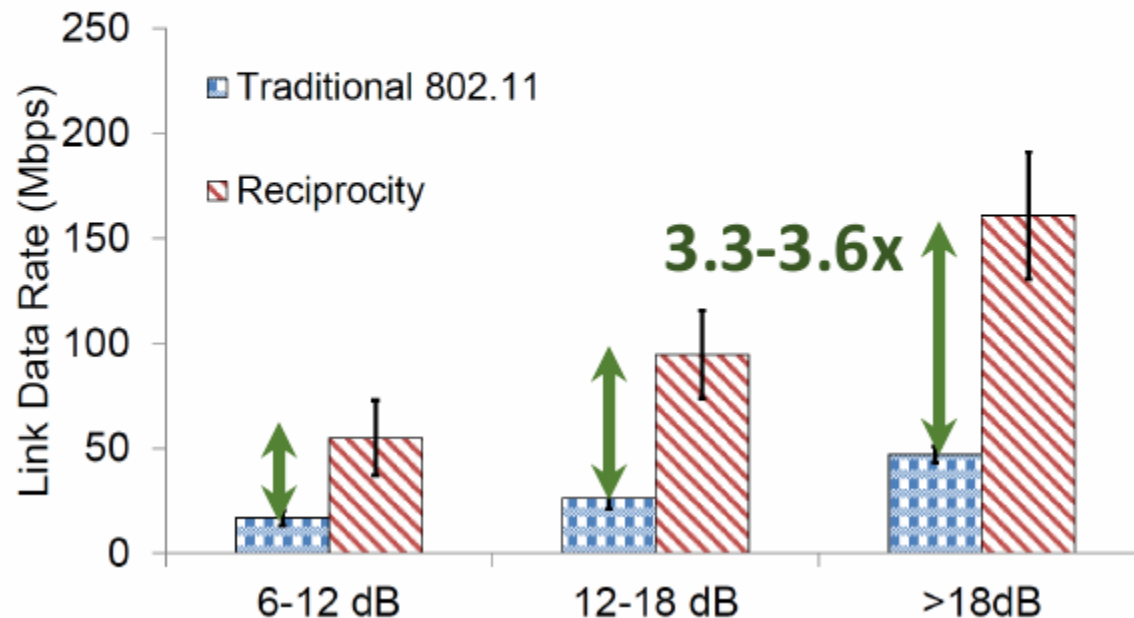


Separate devices → Different Crystals

→ RF chains have oscillator offset relative to each other

Real-time Distributed MIMO Systems

MegaMIMO 2.0 vs. Traditional 802.11



We need a protocol for ensuring that the multipliers are the same despite being applied on different boxes

Преимущества использования множества путей передачи данных

Принцип *Resource Pooling*

- Телефонные сети
 - Коммутация каналов позволяет соединять одновременно лишь двух абонентов
- Многопоточная маршрутизация – коммутация пакетов 2.0
 - Коммутация по единственному пути приводит к возникновению перегрузок и снижает скорость передачи

Преимущества использования множества путей передачи данных

- Повышение надёжности и устойчивости к возникновению неисправностей
- Повышение конфиденциальности и анонимности в сети
- Усреднение характеристик каналов, повышение устойчивости приложений к возникновению перегрузок

Как использовать маршруты?

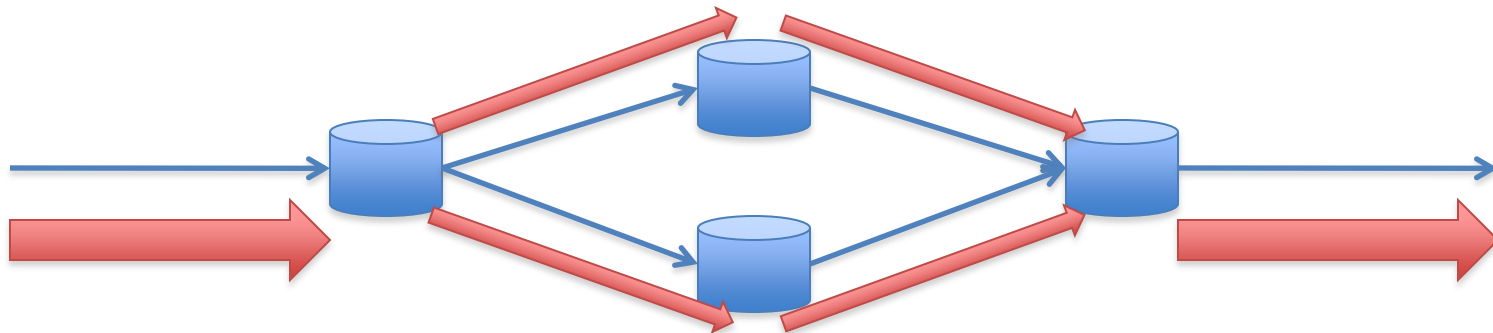
- Количество маршрутов
- Основные и запасные маршруты

Построение множества маршрутов

Source-based vs Hop-Based vs SDN

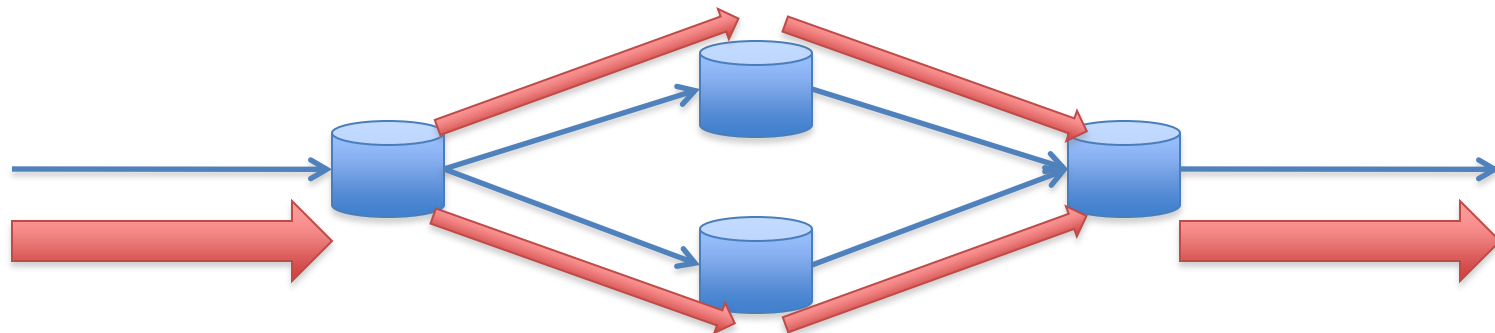
- Маршрутизация от источника
 - При выборе можно использовать дополнительную информацию о хостах
 - Как получить информацию о маршрутах?
 - Как указать, какой маршрут использовать?
- Маршрутизация по хопам
 - Как строить множество маршрутов?
- Централизованная маршрутизация
 - Как уменьшить накладные расходы на использование альтернативных путей?

Equal-cost multipath routing (ECMP)



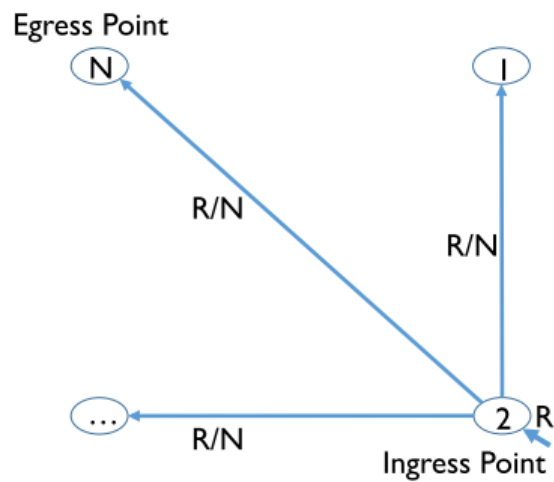
- ECMP
 - Распределяет нагрузку между множеством путей с одинаковыми метриками маршрутизации
- Round-Robin использовать нельзя:
 - Переупорядочивание пакетов (TCP такого не любит)
 - Разные пути могут иметь разные RTT, PLR...
 - На разных путях может быть разный MTU

Equal-cost multipath routing (ECMP)

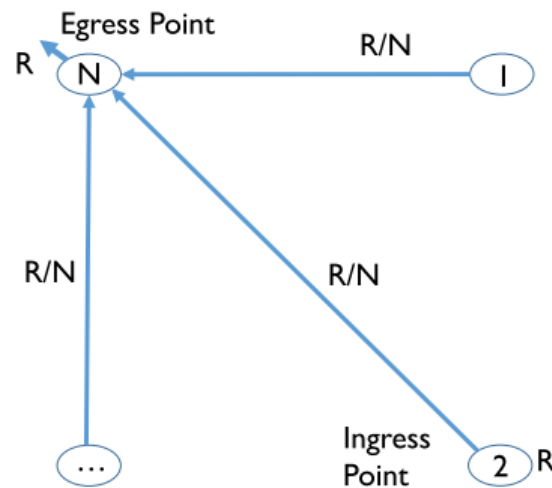


- Для балансировки используется хэширование
 - количество бакетов = количеству выходов
 - Хэш на основе адресов пакетов сохраняет привязку потоков по выходам

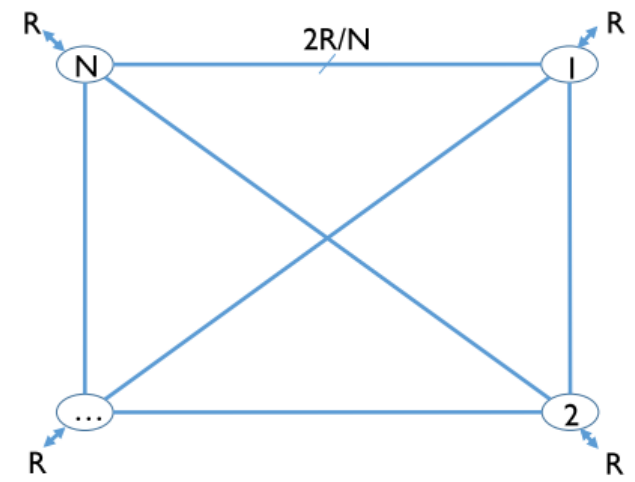
Valiant Load Balancing



(a) VLB Stage 1



(b) VLB Stage 2



(c) VLB Stage 1+2

Динамическая балансировка

- Проблема возникновения коллизий
- Проблема переупорядочивания пакетов
- FLARE
- HEDERA

На каком уровне проводить балансировку нагрузки?

- На уровне каналов
 - TRILL, Wireless MIMO
- На уровне маршрутизаторов
 - ECMP, VL2, HEDERA, MARA, etc
- На уровне транспорта
 - SCTP
 - MPTCP
- На уровне приложений
 - BitTorrent