

# Принципы маршрутизации на уровне автономных систем

Степанов Евгений Павлович

# Программа курса

## Подходы:

- 1. Балансировка**
  - Балансировка нагрузки и управление трафиком
- 2. Управление перегрузкой**
  - Современные протоколы управления перегрузкой TCP
- 3. Демультимплексирование/мультиплексирование**
  - Многопоточные транспортные протоколы
  - Маршрутизация на уровне интернет провайдеров
  - Network Coding
- 4. Сегментация**
  - TCP Proxy

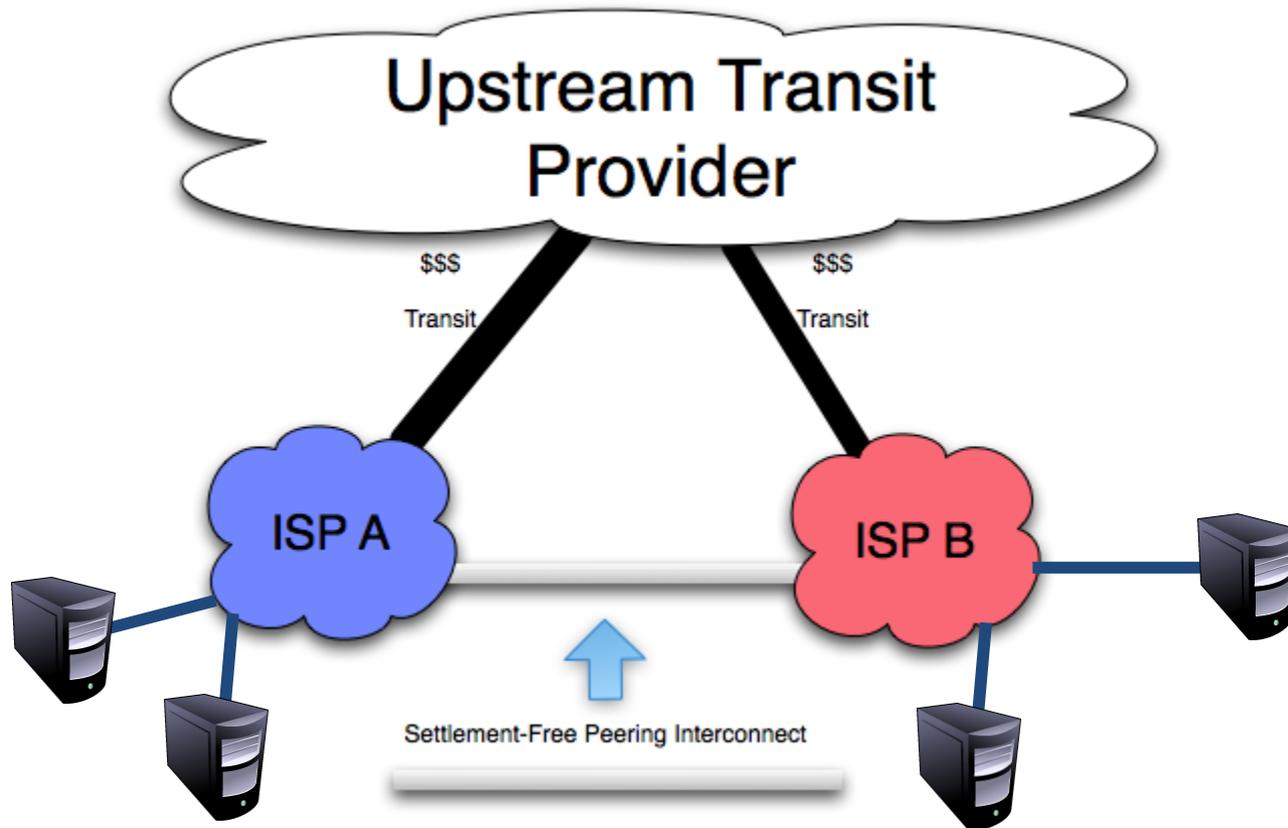
## Модели:

- Сетевое исчисление: математический подход к качеству сервиса
- NS3: моделирование поведения сети с высокой точностью

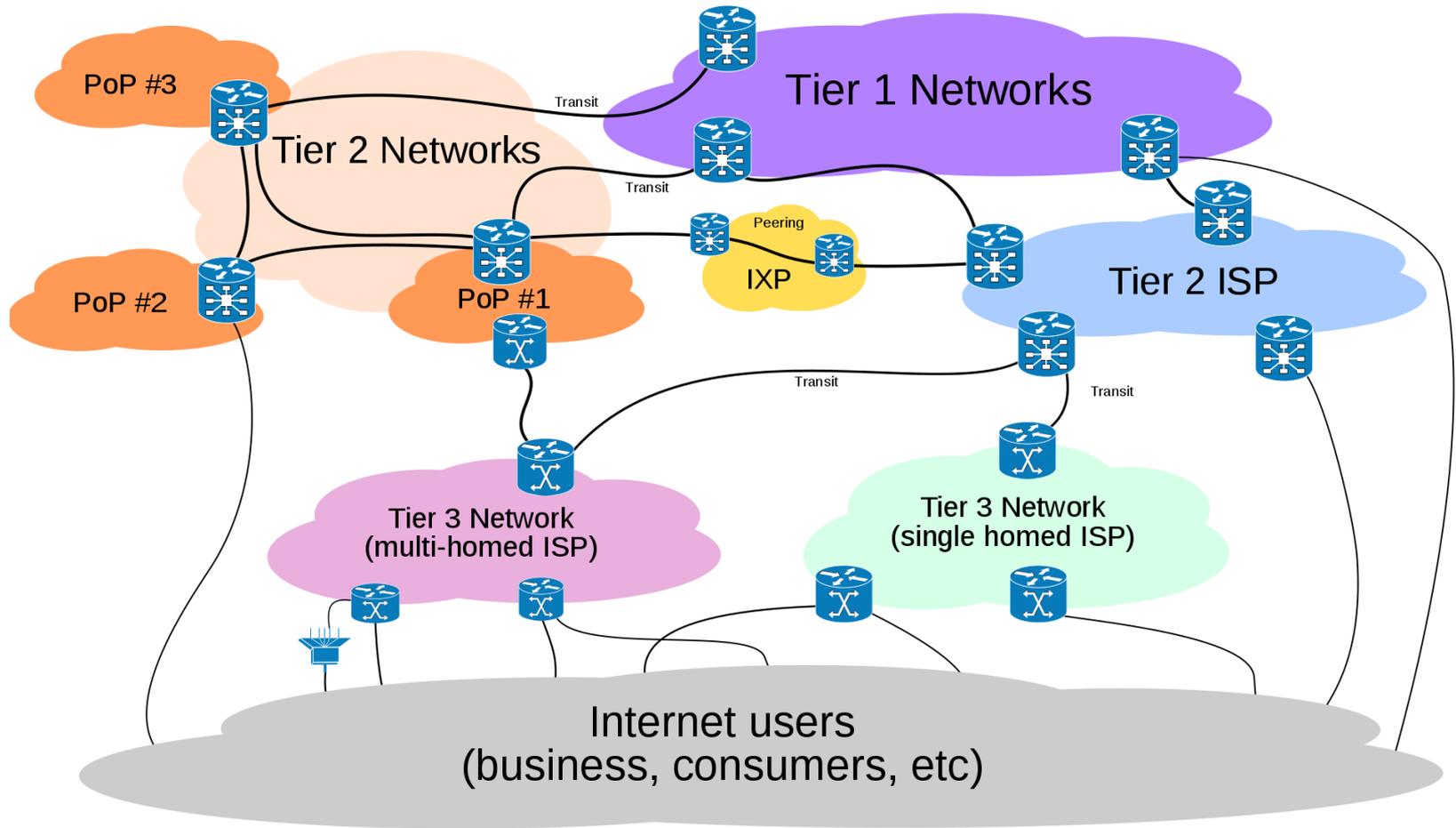
## Примеры:

- Управление сетевыми ресурсами в Центрах Обработки Данных
- Обеспечение качества сервиса в сетях доставки контента
- Пропускная способность по требованию

# Transit & peering



# Структура современного интернета



# Структура современного интернета

## Иерархия Автономных Систем (AS)

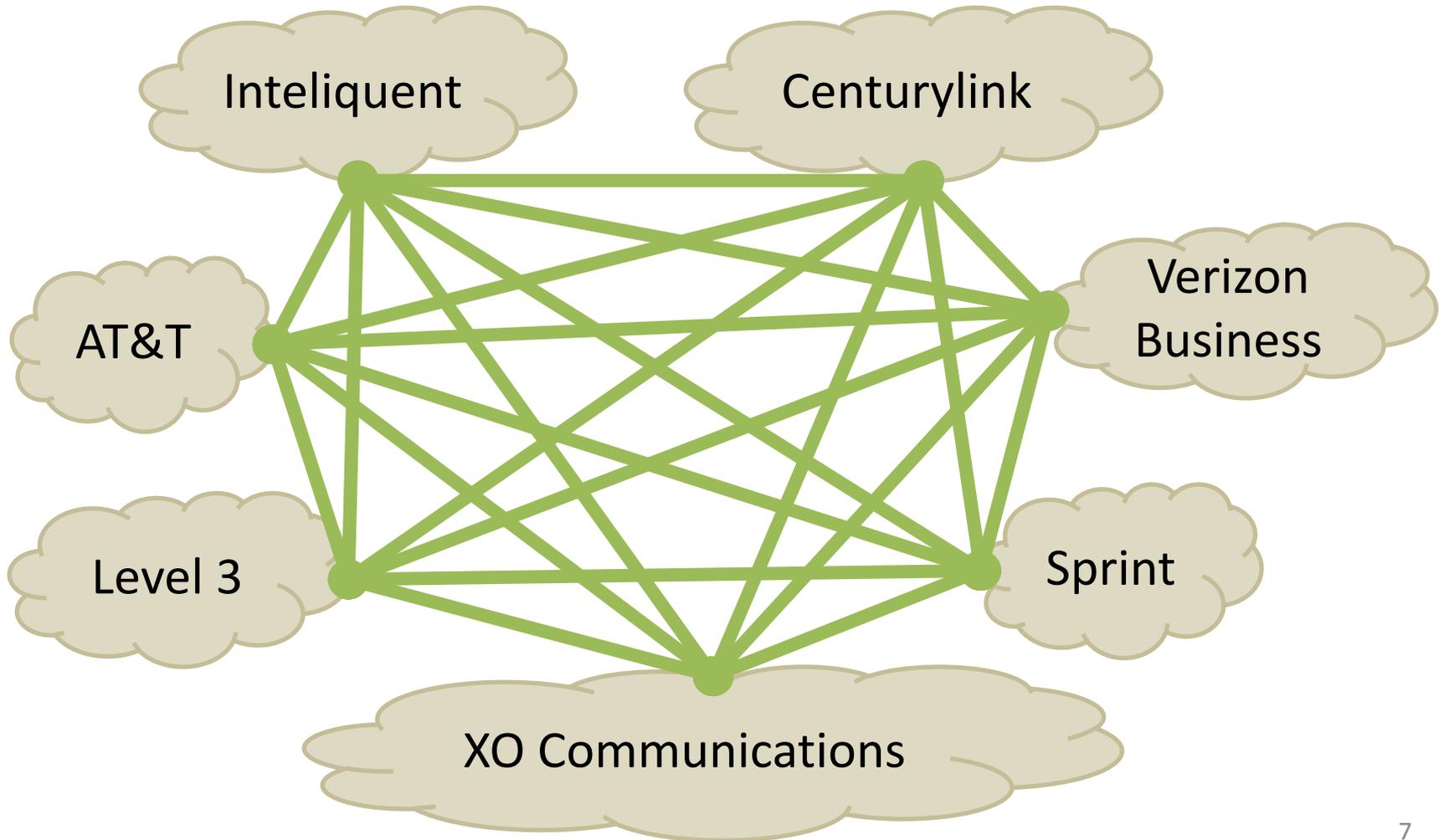
- Tier-1
  - Обмен трафиком с любой сетью в интернете без какой-либо платы за его транзит или пиринг
- Tier-2
  - Бесплатный обмен трафиком с некоторыми сетями
  - Обмен с остальными сетями оплачивается
- Tier-3
  - Любой обмен трафиком с сетями интернет требует оплаты за его транзит или пиринг

# Пиринг между сетями Tier-1

## *settlement-free peering*

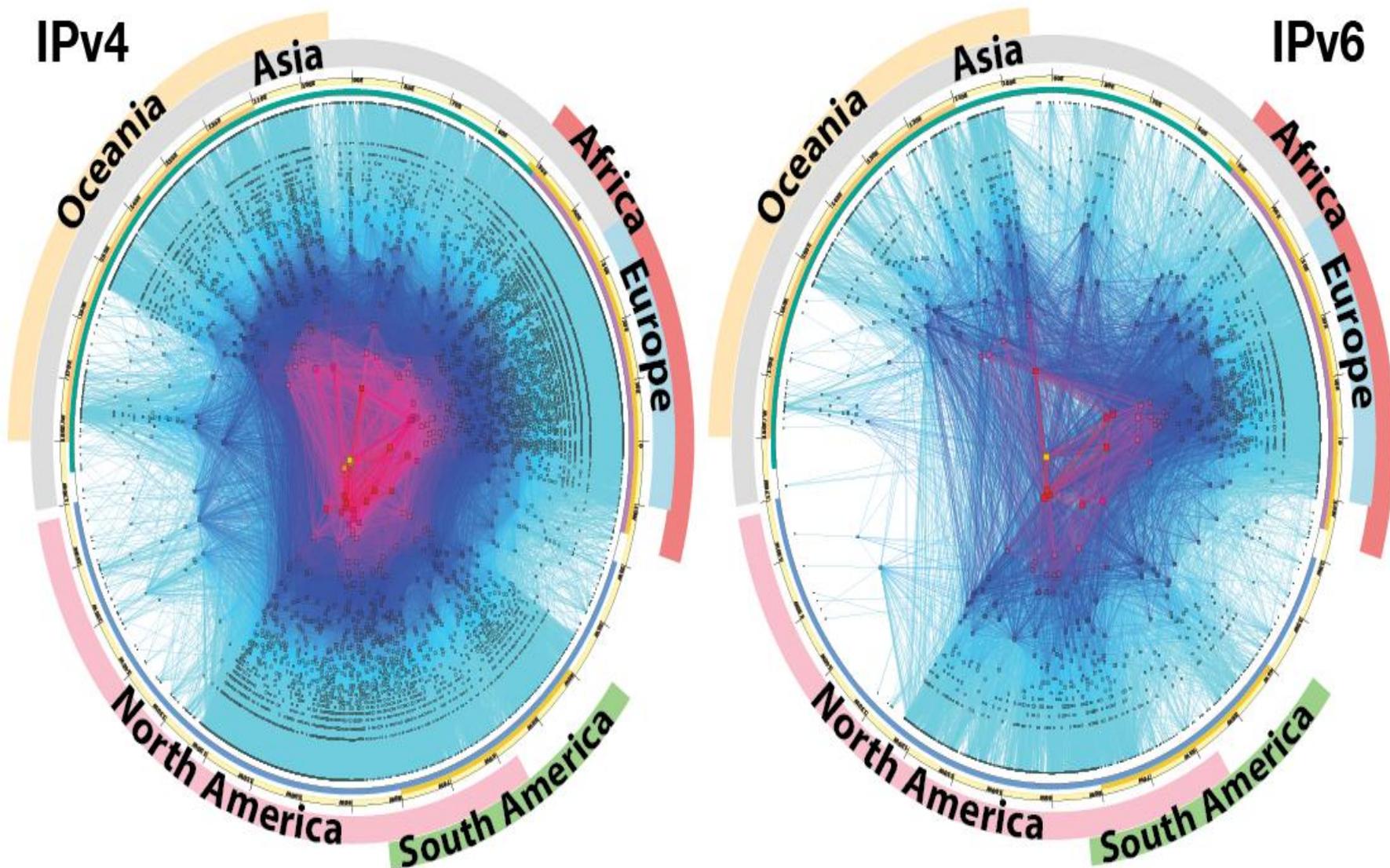
- Сети Tier-1 имеют прямое соединение, и оба партнёра считают это соединение выгодным
  - Если один из партнёров принимает больше, чем передаёт – он может потребовать дополнительную плату за соединение
- Передача трафика в сети, подключённые к партнёрской сети Tier-1, бесплатна
- Стоимость услуг, который провайдер А выполняет от имени провайдера В, идентична стоимости услуг, которые провайдер В выполняет от имени провайдера А

# Пиринг между сетями Tier-1



# CAIDA's IPv4 vs IPv6 AS Core AS-level Internet Graph

Archipelago July 2015

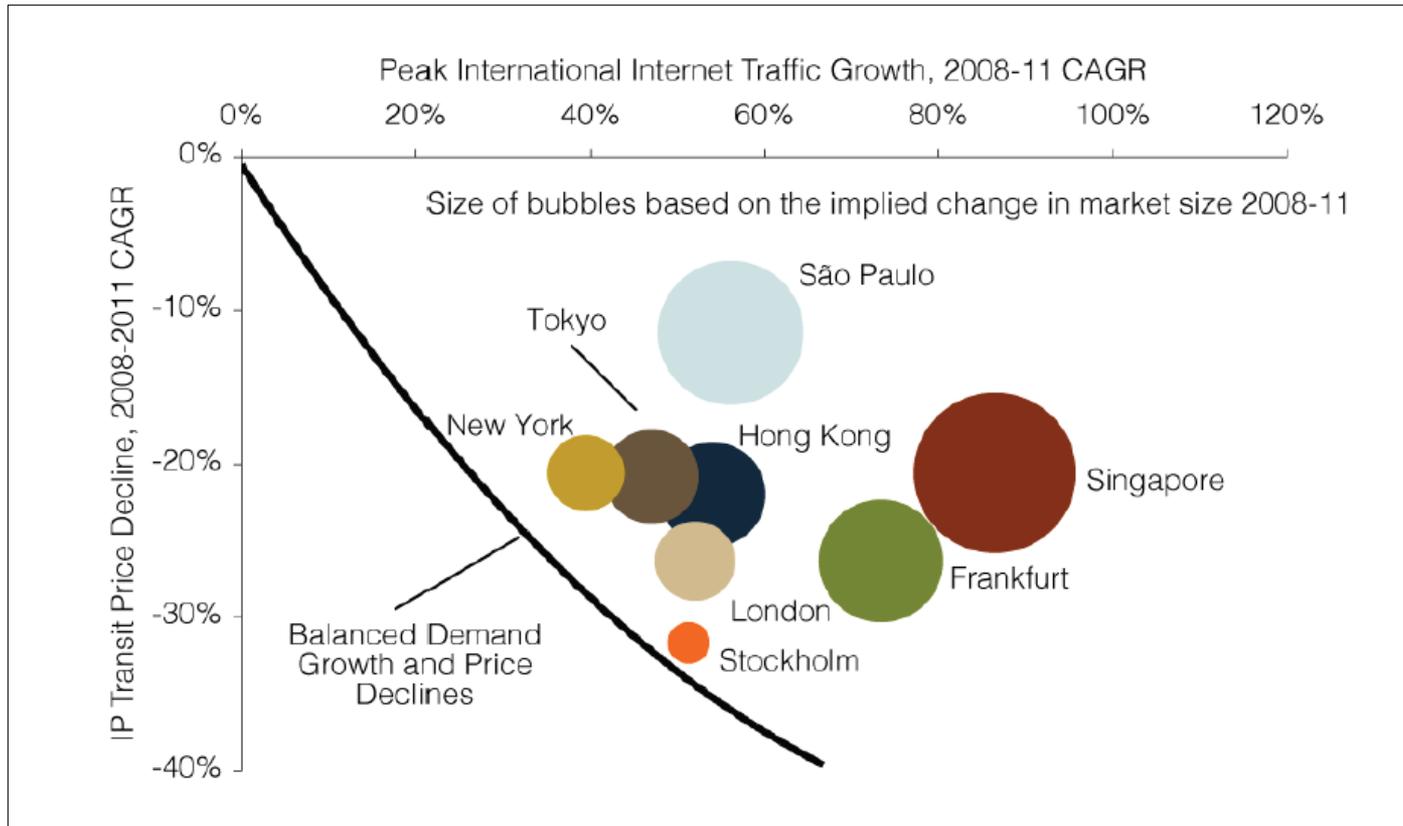


# Транзит на уровне Tier-1

- Поскольку на уровне Tier-1
  - Каждая сеть соединена с каждой другой сетью, и
  - Соглашения пиринга позволяют каждой Tier-1 сети подключиться к каждой сети в интернет,
- То каждая Tier-1 сеть включает в себя всех абонентов сети интернет
- Оплата транзита на уровне Tier-1 означала бы, что сеть оплачивает собственные услуги
- Tier-1 сети продают услуги по транзиту трафика сетям нижних уровней

# Price decay vs. traffic growth

Per-Mbps transit price decline vs. interdomain traffic growth



# Internet Exchange Points (IXPs)

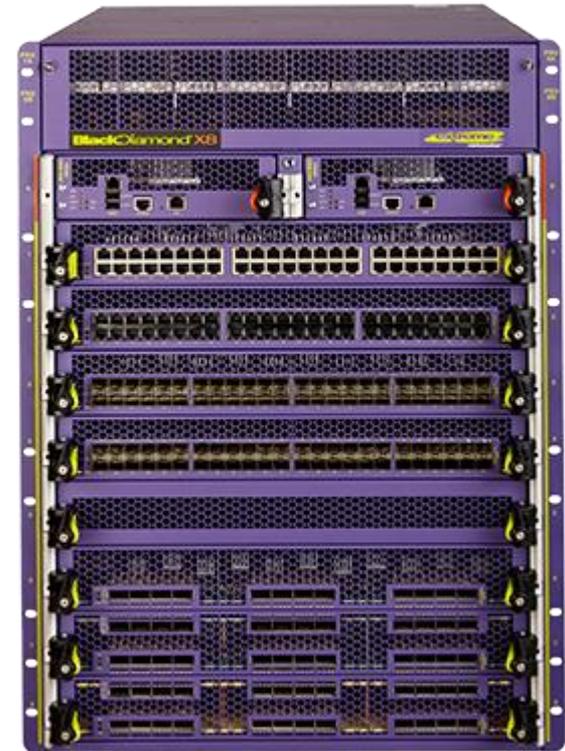
- Покупка услуг по транзиту трафика у провайдеров Tier-1 не всегда осмыслена
- IXP – нейтральная территория, на которой ISP нижних уровней могут осуществлять обмен трафиком на более выгодных условиях
- Каждому ISP достаточно наладить соединение с одним IXP, чтобы получить возможность связаться со всеми его клиентами
  - Network effect

# Принципы работы IXP

- Современные IXP работают на уровне L2
  - один или несколько коммутаторов, к которым могут подключаться взаимодействующие между собой ISP
  - ISP оплачивают доступ к портам коммутаторов IXP
- ISP подключают к IXP свои маршрутизаторы
- Маршрутизаторы взаимодействуют по BGP
  - Взаимодействие может осуществляться как напрямую, так и через специальный route-server
- IXP взимают плату за порт
  - привязка тарифа к трафику препятствует увеличению масштабов использования IXP

# Коммутационное оборудование IXR

- Extreme Black Diamond
  - up to 2,352 1Gbps ports,  
or up to 582 10Gbps  
Ethernet ports in a single  
rack

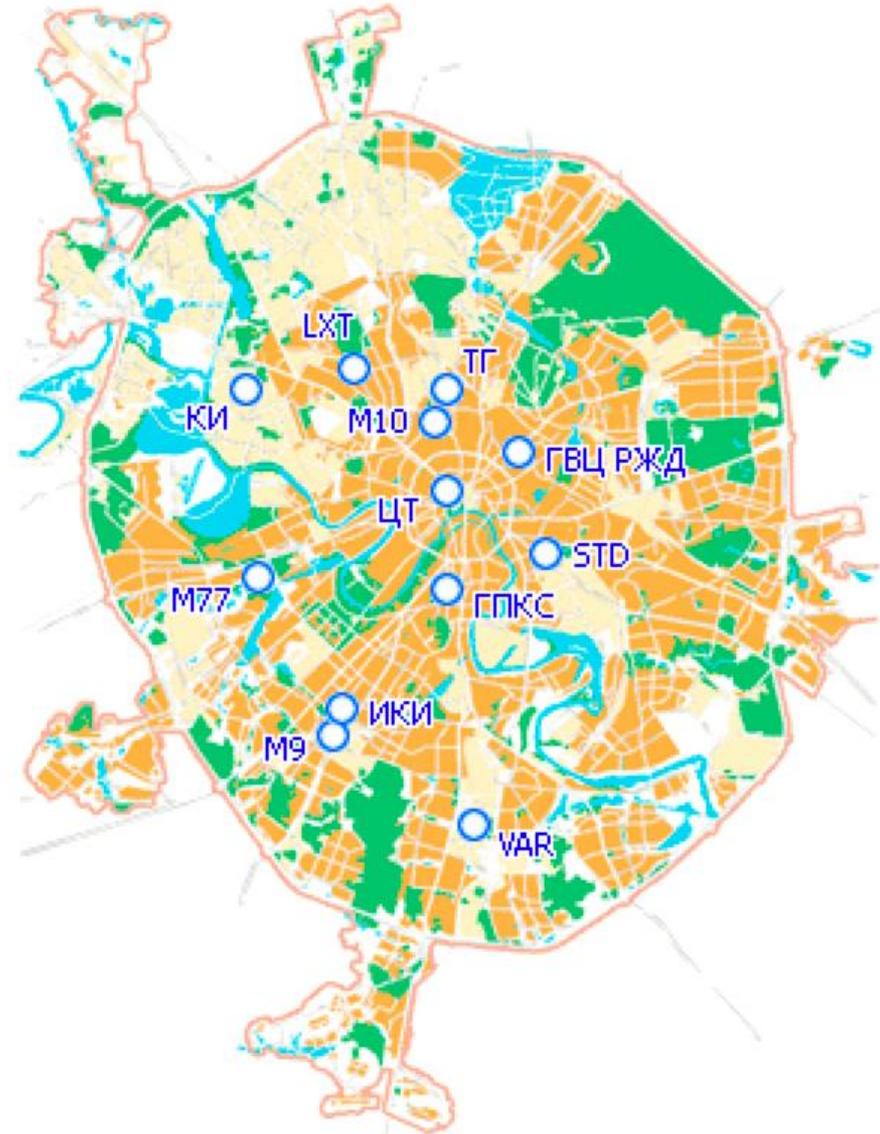
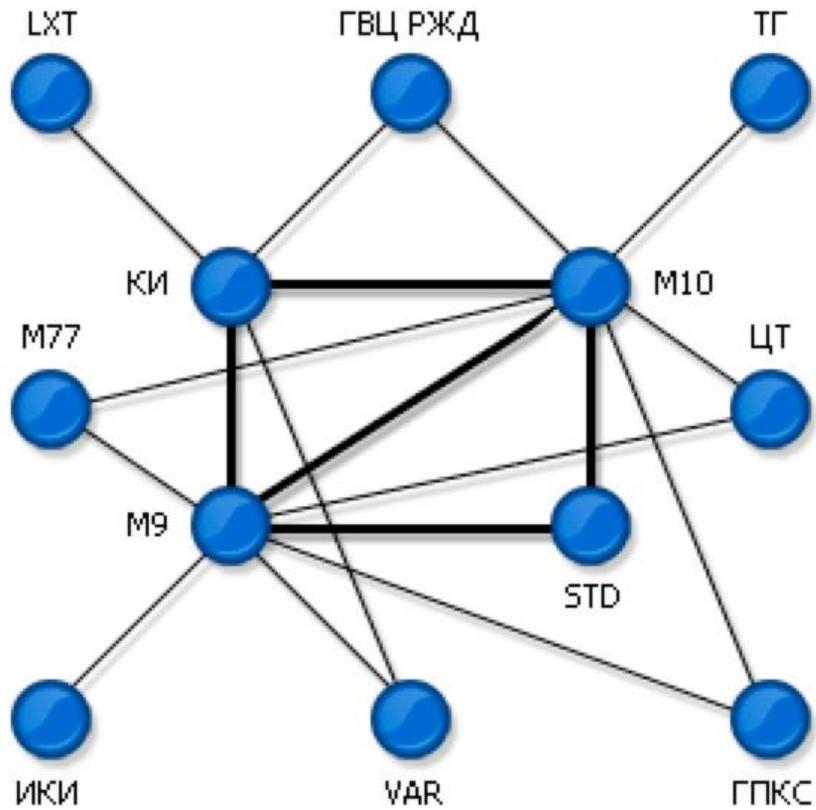


# Типы пиринга на IXP

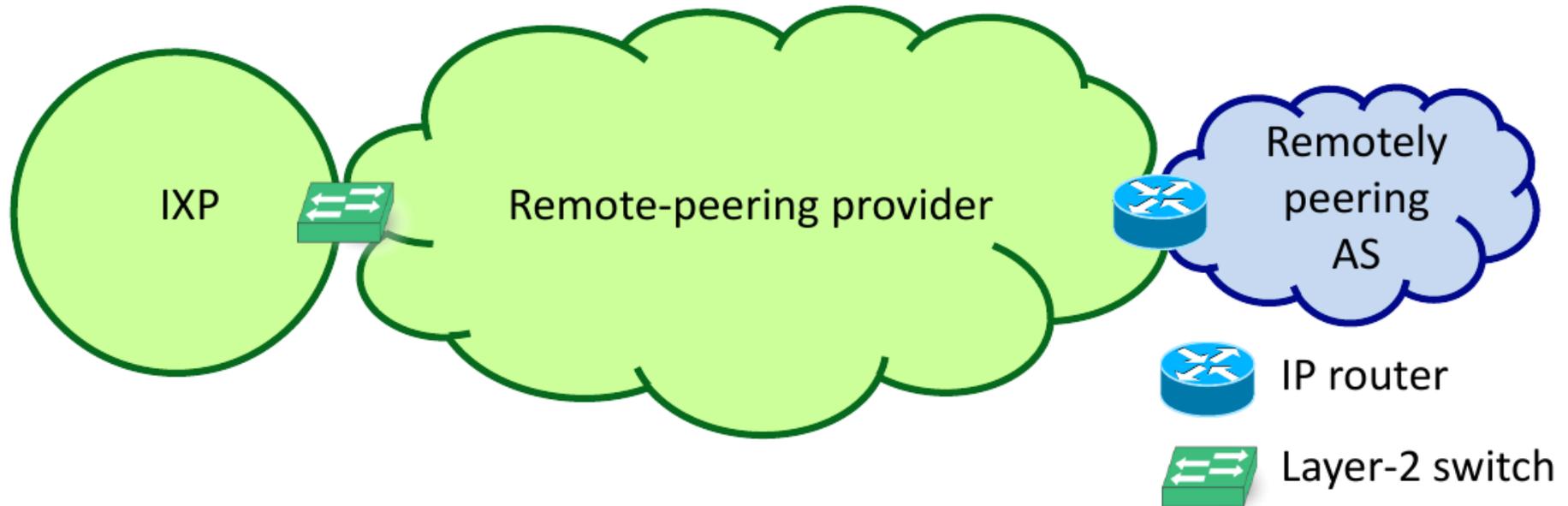
- Mandatory Multi-Lateral Peering (MMLP)
  - Каждый ISP обязан участвовать в пиринге с каждым из остальных ISP
  - Не используется на практике
- Multi-Lateral Peering (MPL)
  - Каждый ISP заключает общее соглашение, по которому он может осуществлять пиринг с любым из других ISP на IXP
- Bilateral Peering
  - ISP заключают индивидуальные соглашения друг с другом, IXP используется как площадка
  - Самый распространённый тип пиринга

# MSK-IX

Топология сети MSK-IX

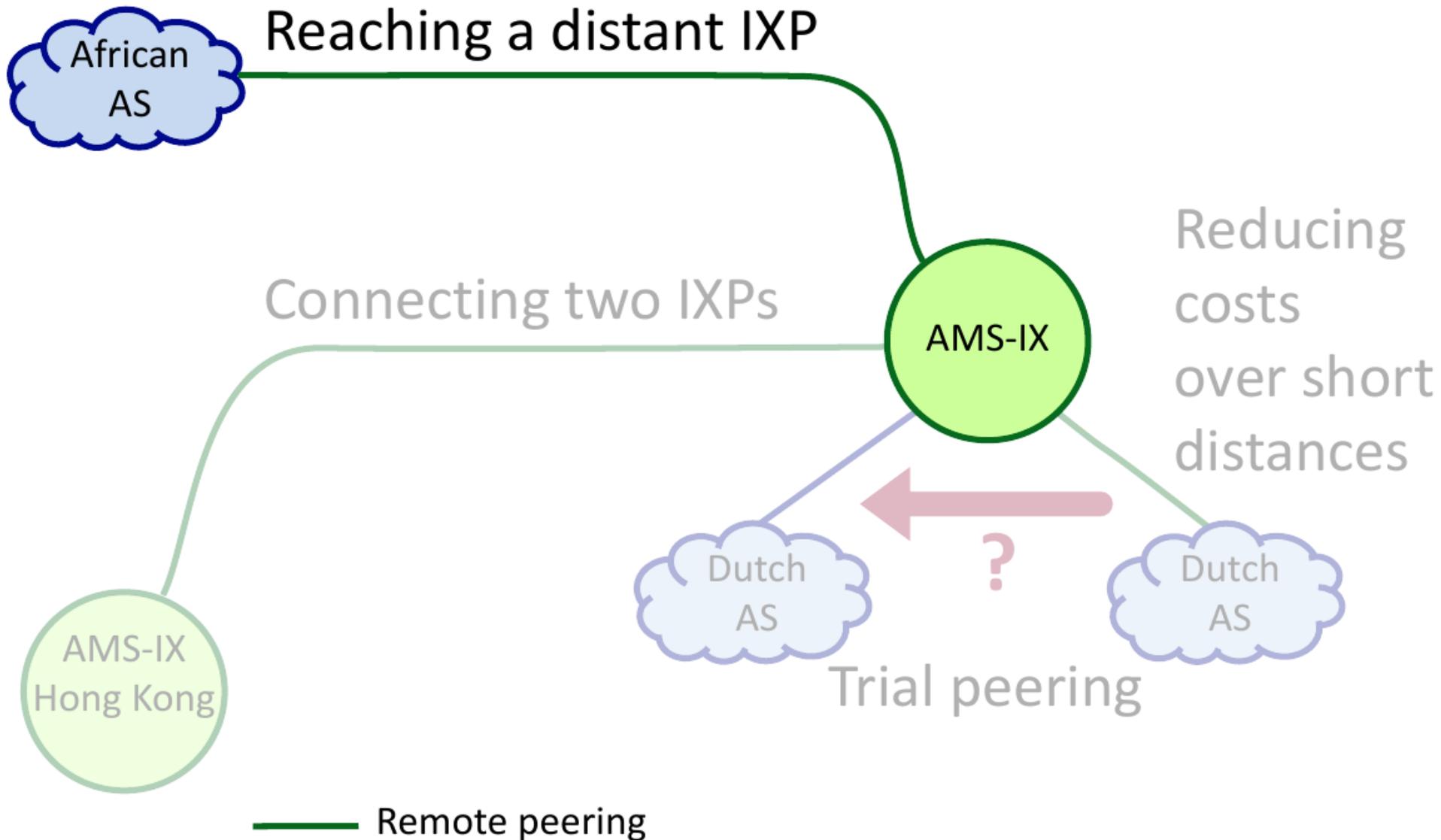


# Remote Peering

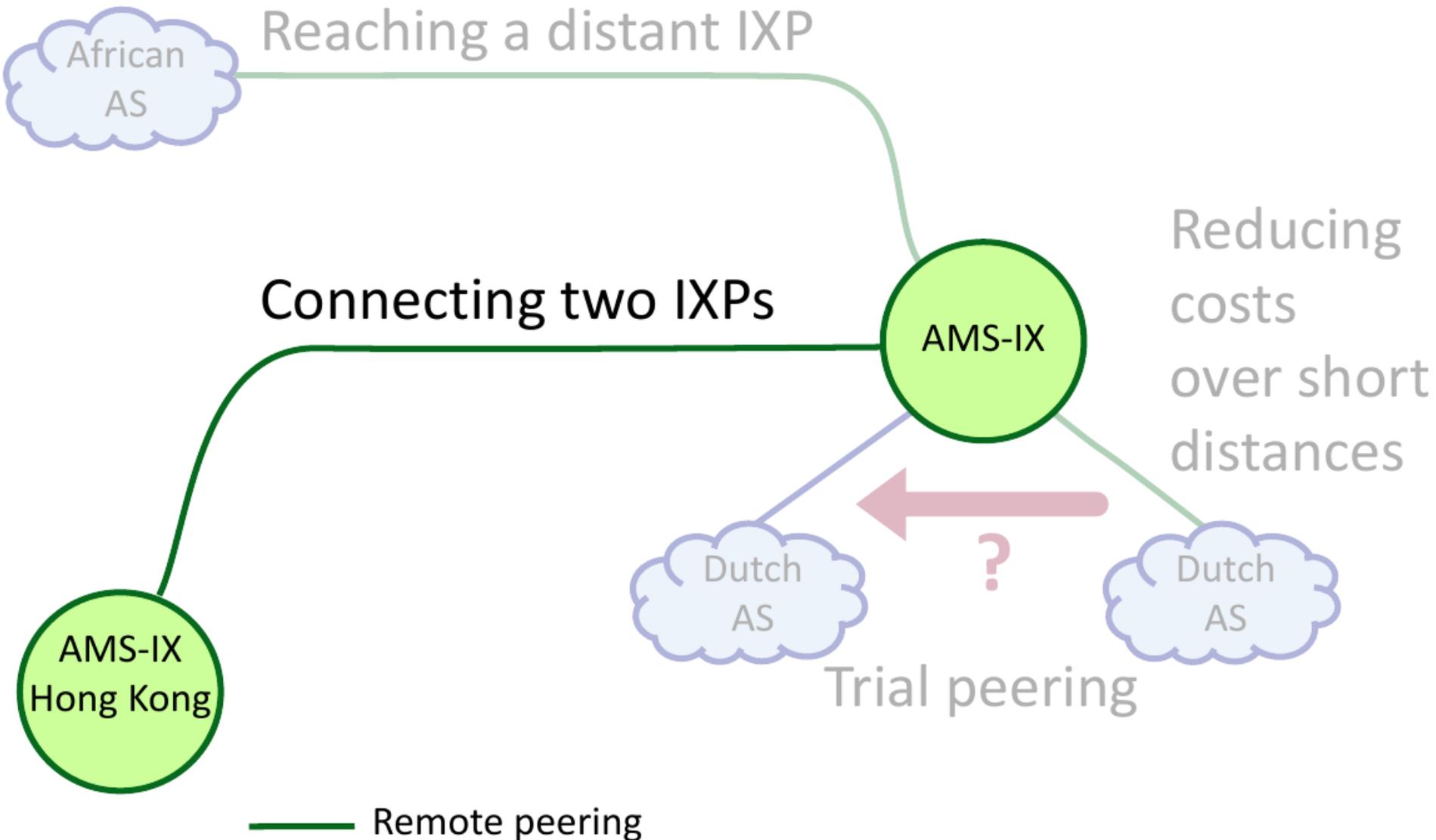


- Подключение ISP к IXP с использованием транзита трафика по L2
- Оборудование для пиринга на IXP

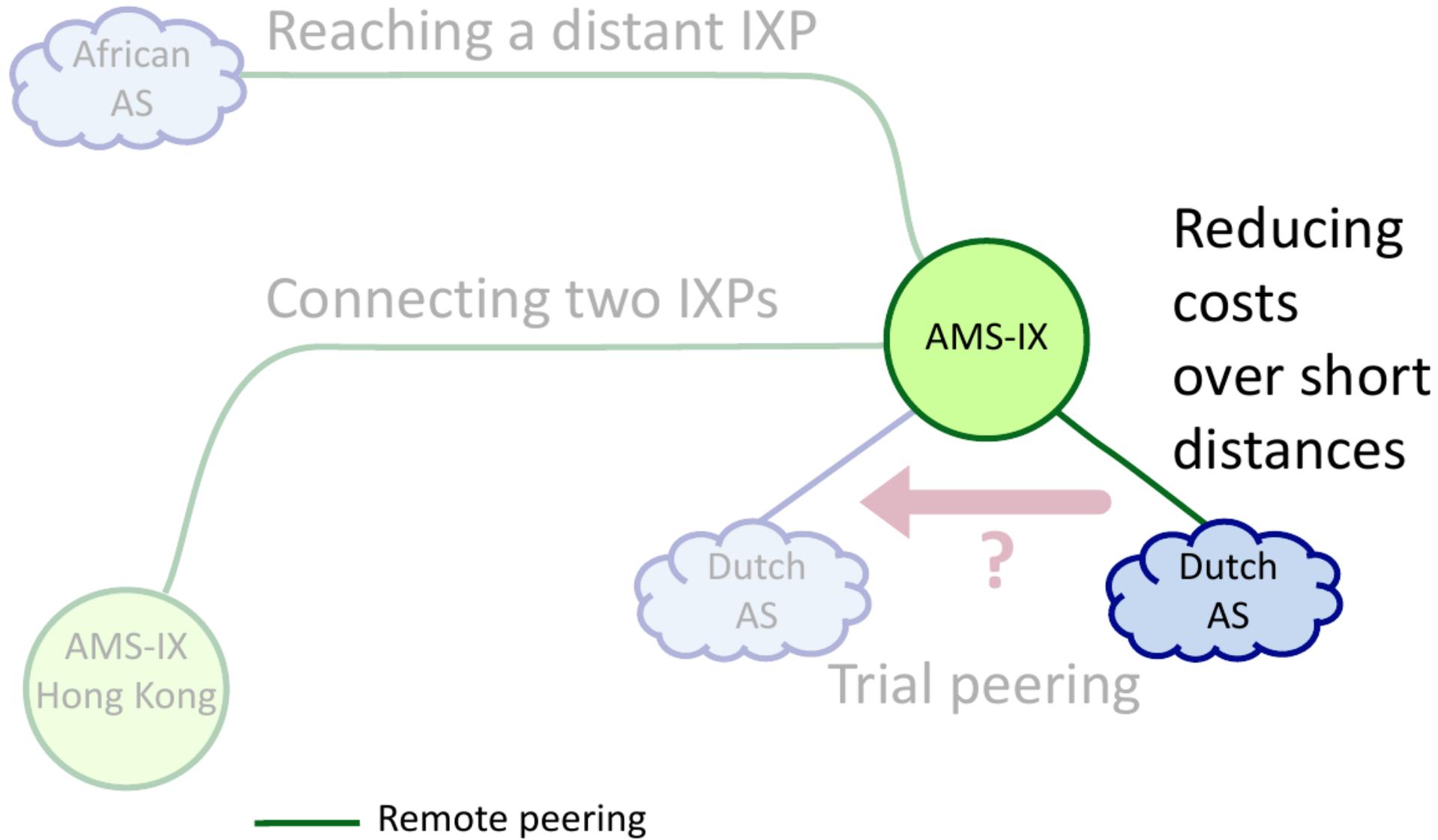
# Remote peering: примеры ИСПОЛЬЗОВАНИЯ



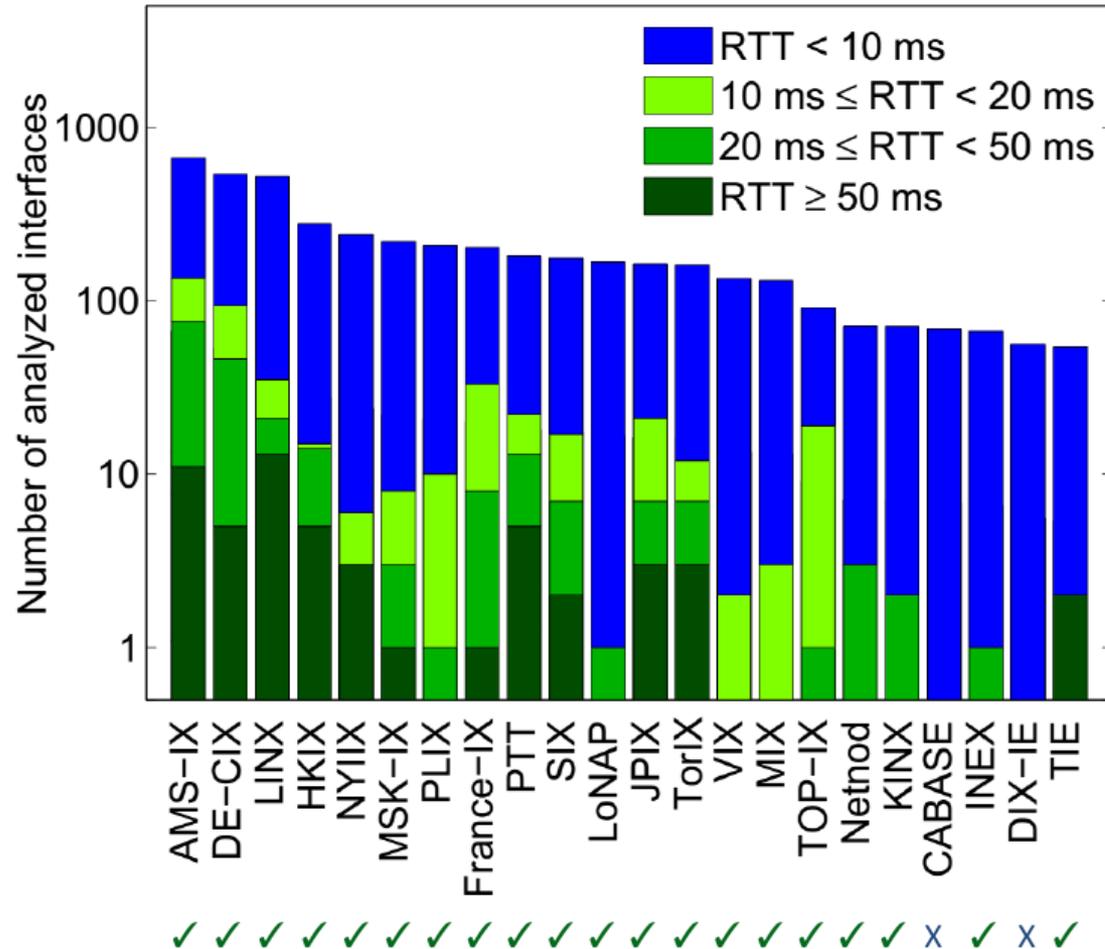
# Remote peering: примеры ИСПОЛЬЗОВАНИЯ



# Remote peering: примеры ИСПОЛЬЗОВАНИЯ



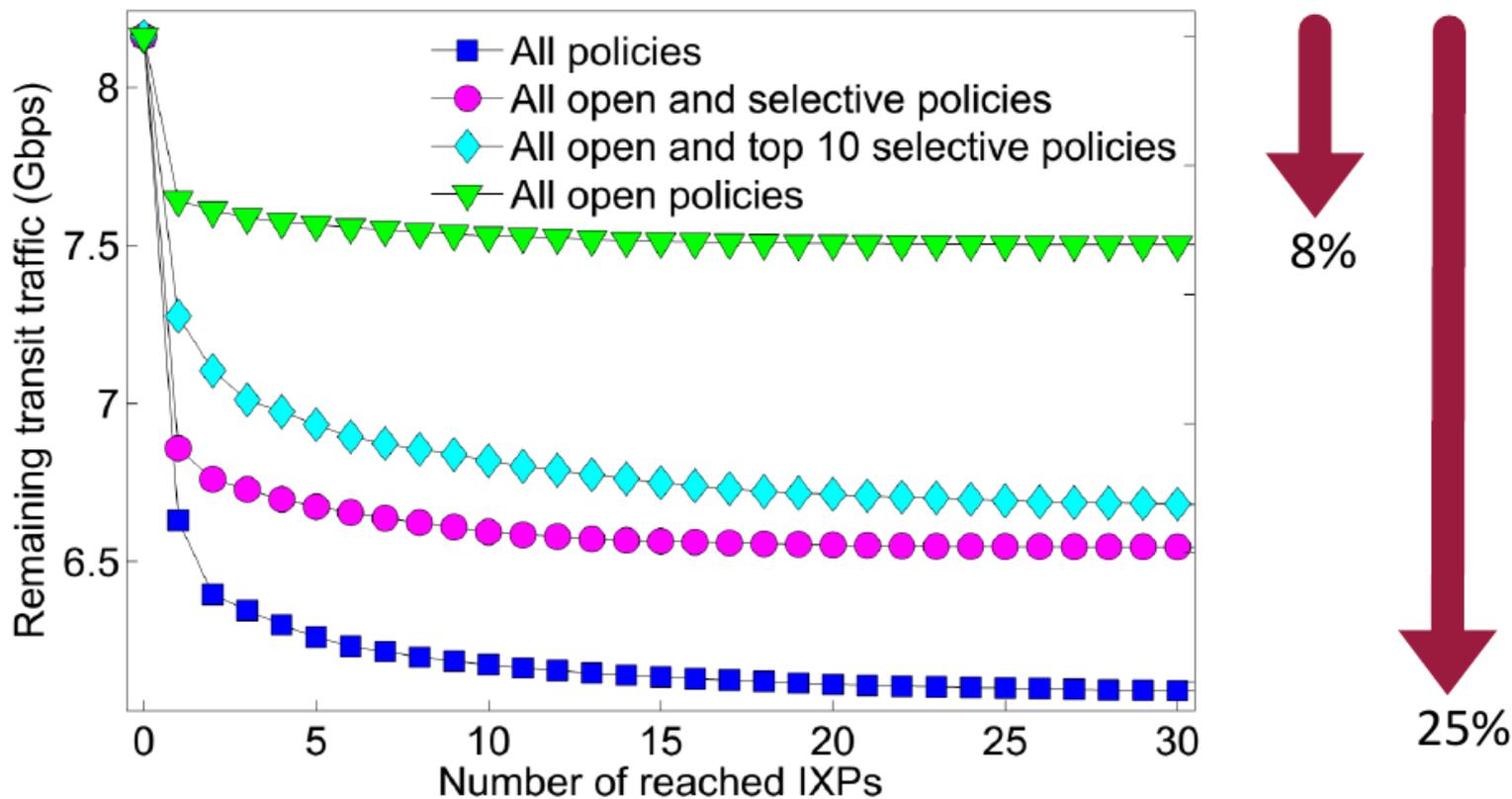
# Статистика по IXP



91% of the IXPs have remote peering



# Сколько трафика можно передать без использования транзита?



Between 8% and 25% of reduction in transit traffic

# Мифы Internet Exchange

- Tier-1 ISP не используют IXP
  - Tier-1 провайдеры являются участниками большинства IXP, и используют их для пиринга
  - Используются ограничения политики
- Заключение договорённости о пиринге на IXP сложнее из-за лишнего посредника
  - Multi-Lateral Peering
- IXP используются в основном для backup'а
  - Большинство провайдеров используют пиринг даже при наличии выделенных маршрутов

# Мифы Internet Exchange

- IXP сильно отличаются от AS
  - На самом деле IXP похожи на крупные AS:
    - Договоры IXP учитывают SLA
    - IXP образуют собственные протяжённые сети  
AMS-IX имеет точку в Гонконге
    - IXP предоставляет расширенные возможности для управления и мониторинга трафика  
Route-Server и политики BGP

# Как сделать транзит дешевле?

Оптимизировать использование каналов

Committed Data Rate (Mbps)	Стоимость за Mbps в месяц
10	25
50	15
100	10
1000	5
10000	4

Стоимость транзита трафика voxel.net

- Economic of scale – массовая продажа дешевле
- Стоимость – субаддитивная функция

Part 1: Tuángòu for IP transit  
*Why did we put our heads together?*

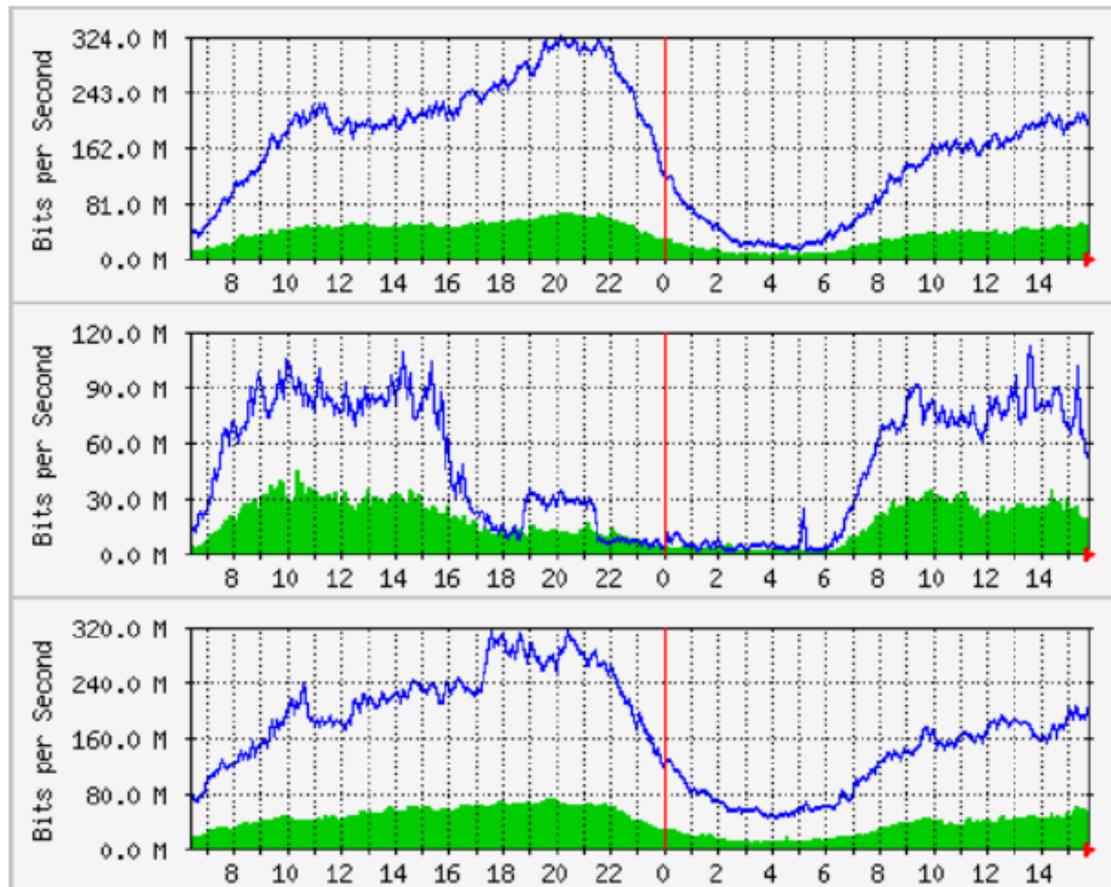
**TO SAVE MONEY!**



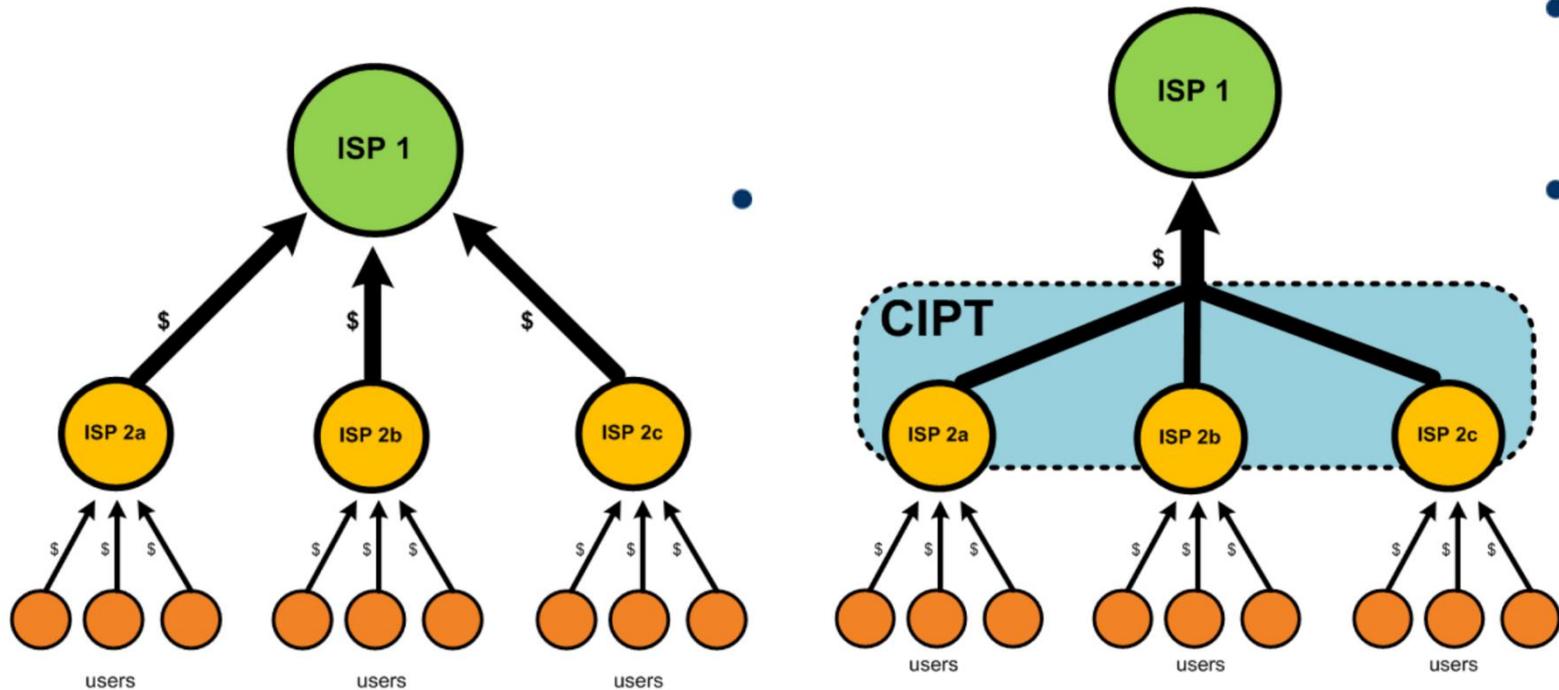
**BRADLEY GROUP SHOWERS**

# CIPT (Cooperative IP Transit)

Трафик разных ISP можно складывать



# CIPT можно рассматривать как отдельный вид ISP с федеративной структурой



# Как разделить стоимость аренды канала между участниками CIPT?

- Вектор Шейпли (Shapley value)

ISP  $i$ 's expected marginal contribution if the players join the coalition one at a time, in a uniformly random order

$$\phi_i(c) = \frac{1}{N!} \sum_{\pi \in S_N} \underbrace{(c(S(\pi, i)) - c(S(\pi, i) \setminus i))}_{i\text{'s marginal contribution}}$$

$N$  = number of players

$c(S)$  = cost of coalition  $S$

$S(\pi, i)$  = set of players arrived in the system not later than  $i$

$\pi$  = permutations of the set of players  $N$

# Как рассчитать значение Вектора Шейпли на практике?

## Использовать метод Монте-Карло

- Рассчитать значение стоимости аренды для  $K$  случайных порядков  $\pi_k$

$$\hat{\phi}_i(c) = \frac{1}{K} \sum_{\pi \in \Pi_K} (c(S(\pi, i)) - c(S(\pi, i) \setminus i))$$

- При использовании  $K = 1000$  ошибка  $< 1\%^*$

(\*) D. Liben-Nowell, A. Sharp, T. Wexler, K. Woods, “Computing Shapley Value in Cooperative Supermodular Games”, Preprint, 2010.

# Резюме

- Сеть состоит из AS, маршрутизация между которыми зачастую определяется не метриками качества сервиса
- Транзит трафика дорогой
- Internet Exchange Points (IXPs) позволяют снизить стоимость обмена данными между сетями нижних уровней иерархии AS
- Неравномерность трафика и специфика формулы расчёта стоимости трафика делают выгодными разные формы коллаборации ISP