MikroTik RouterOS Семинар

QoS Лучшая практика

(перевод на русский язык white_crow 2010 г. от P.X)

Прага,

MUM Чехия2009

Вопросы и ответы

- *Вопрос:* Это возможно – приоритезация трафика по типу (1) и строгие ограничения на каждого отдельного клиента (2) на том же маршрутизаторе?
- *Ответ:* Да!

- *Вопрос:* Что для этого нужно?
- *Ответ:* Вам понадобяться:

  1)Packet Flow Diagram

  2)HTB (queue tree)

  3)Mangle

  4)PCQ

  5)Address List

Mangle

- Mangle позволяет вам маркировать IP пакеты специальными метками.

- Эти метки используются другими средствами маршрутизатора, такими как маршрутизация и управление полосой пропускания для идентификации пакетов.

- В добавок, средствами mangle можно модифицировать некоторые поля в IP заголовке, такие как TOS (DSCP) и TTL поля.

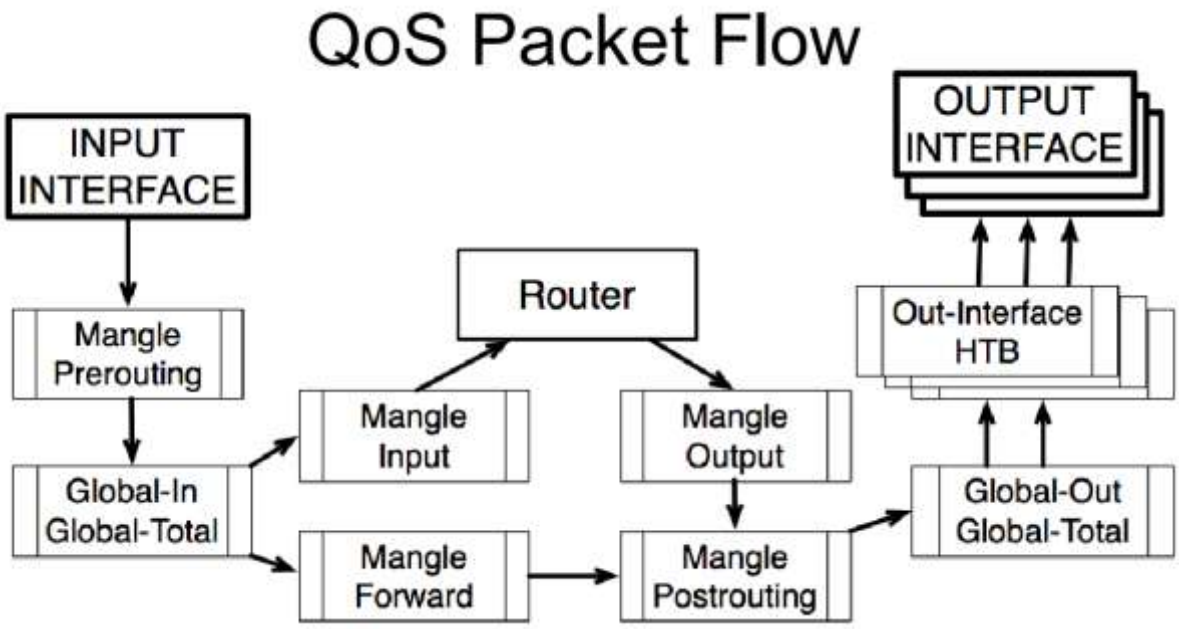
Hierarchical Token Bucket

- Вся реализация управления пропускной способностью в RouterOS основана на иерархии - Hierarchical Token Bucket (HTB)

- HTB позволяет вам создавать иерархические структуры очередей и определять отношения между очередями

- RouterOS поддерживает 3 виртуальных HTB (global-in, global-total, global-out) и еще один прямо перед каждым выходным интерфейсом.

Движение пакетов



http://wiki.mikrotik.com/wiki/Packet_Flow

Двойной QoS

- Возможно маркировать и шейпить трафик дважды на одном роутере:

  1)

  - Prerouting в цепочке Mangle – для первой маркировки
  - Global-in HTB – для первого шейпинга

  (*прим. Переводчика – это шейпинг по типу трафика)*

  2)

  - Forward или Postrouting в цепочке Mangle - для второй маркировки.
  - Global-out или Out-interface HTB для второго шейпинга

  (*прим. Переводчика – это шейпинг по пользователям)*

  (*прим. Переводчика – можно использовать только один из двух этапов – в зависимости от задач)*

- Двойной QoS возможен только с помощью Queue Tree.

Почему не Simple Queues?

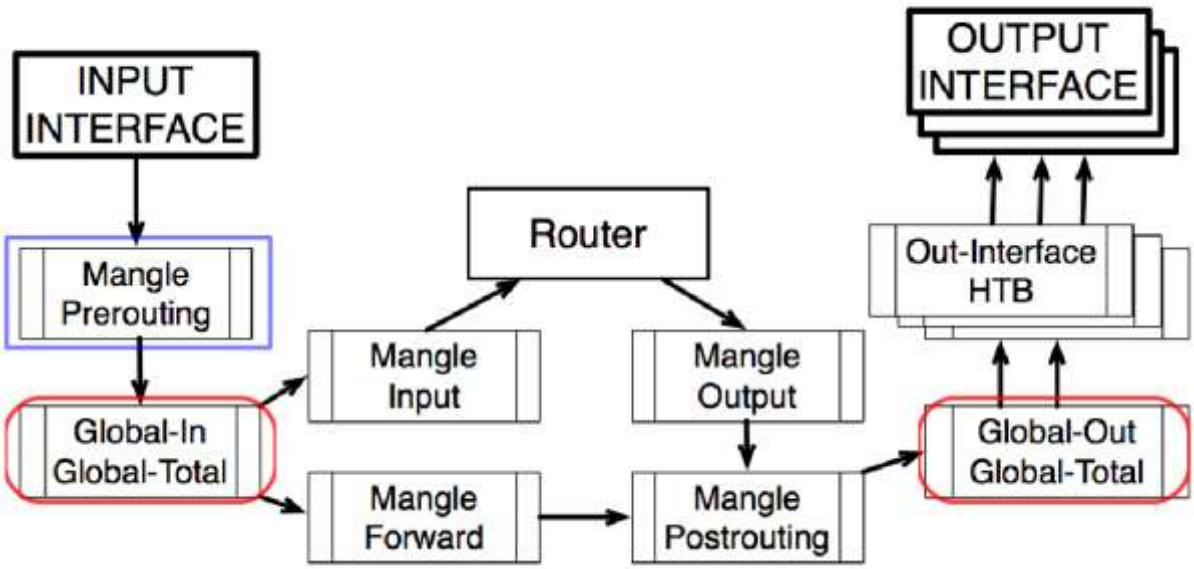
- Simple queues (простые очереди) отсортированы – по аналогии с правилами фаервола
  - Для того, чтобы добраться до 999-ой очереди, пакет должны быть проверен на соответствие всем 998-ми предыдущим очередям.

- Каждая простая очередь **может** стоять в 3 раздельных очередях:
  - One in Global-in (“direct” part)
  - One in Global-out (“reverse” part)
  - One in Global-total (“total” part)

(*прим. Переводчика:*
Вывод - Simple Queues использовать при большом количестве пользователей – не эффективно – существенно возрастает нагрузка на процессор, погрешность шейпинга также возрастает. И, как уже было сказано, также не возможен двойной QoS при использовании Simple Queues*)*

Simple Queues and Mangle

(простые очереди и Mangle)



Queue Tree

(дерево очередей)

- Дерево очередей является только однонаправленным и может быть расположено в любом из доступных HTB.
- Очереди **Queue Tree** обрабатываются не по порядку – весь трафик обрабатывается одновременно.
- Все дочерние очереди должны иметь пакеты, маркированные с помощью средств “/ip firewall mangle” , назначенных для них.
- Простая очередь (Simple queue) помещенная в тот же HTB, будет «принимать» весь трафик от Queue Tree очереди.

Global-Out или Interface HTB?

Существует два основных отличия

- В случае SRC-NAT (masquerade) - Global-Out будет знать о частных адресах клиента, но Интерфейс HTB не будет - интерфейс HTB находится после SRC-NAT.

- Каждый интерфейс HTB получает только трафик, который будет отправится через определенный интерфейс - нет необходимости в разделении upload и download в mangle

Выводы

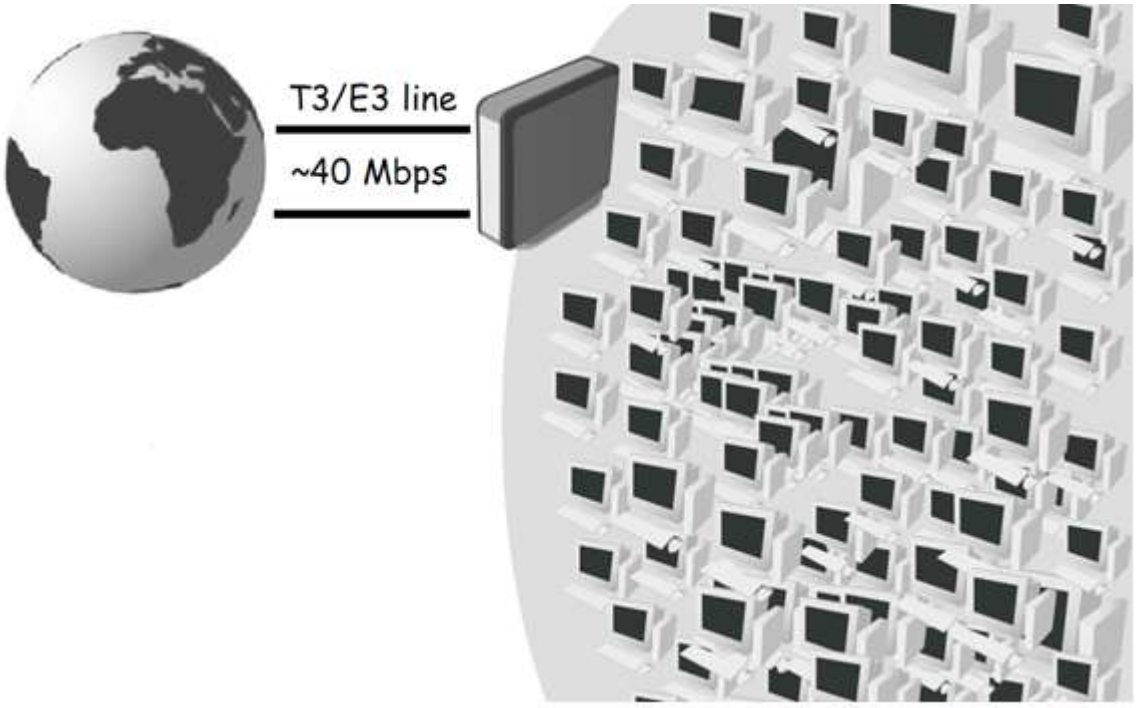
- Мы будем использовать mangle и queue tree:
  - Маркировать трафик по типу - в mangle цепочке Prerouting
  - Приоритезация и ограничение трафика по типу - в Global-in HTB

    (и/или) (*прим. Переводчика)*

  - Перемаркировка трафика по клиентам в mangle цепочке Forward
  - Ограничение трафика по клиентам в Interface HTB

- Необходимо свести количество правил mangle и очередей к минимуму, чтобы увеличить производительность этой конфигурации.

Ограничение клиентов

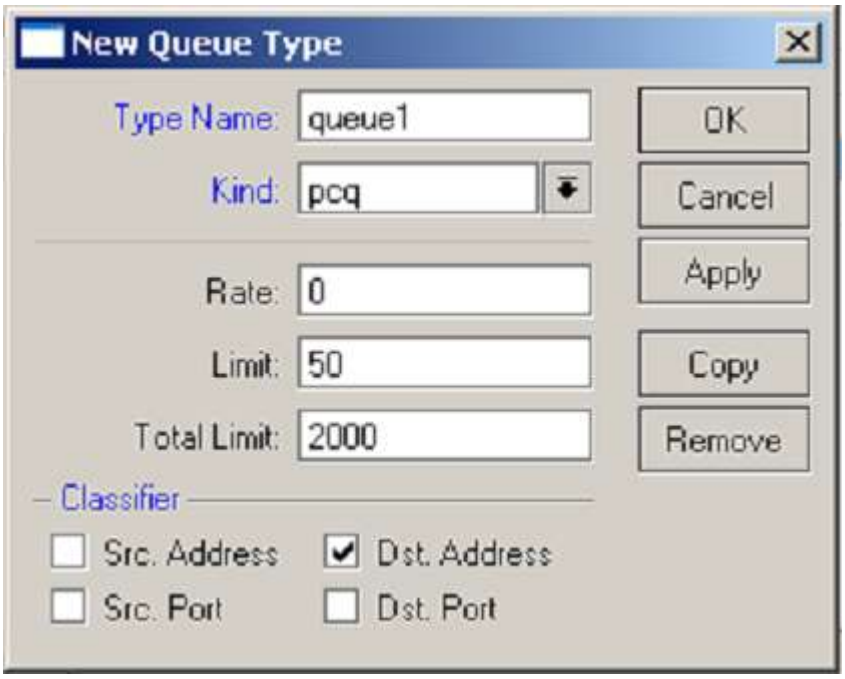


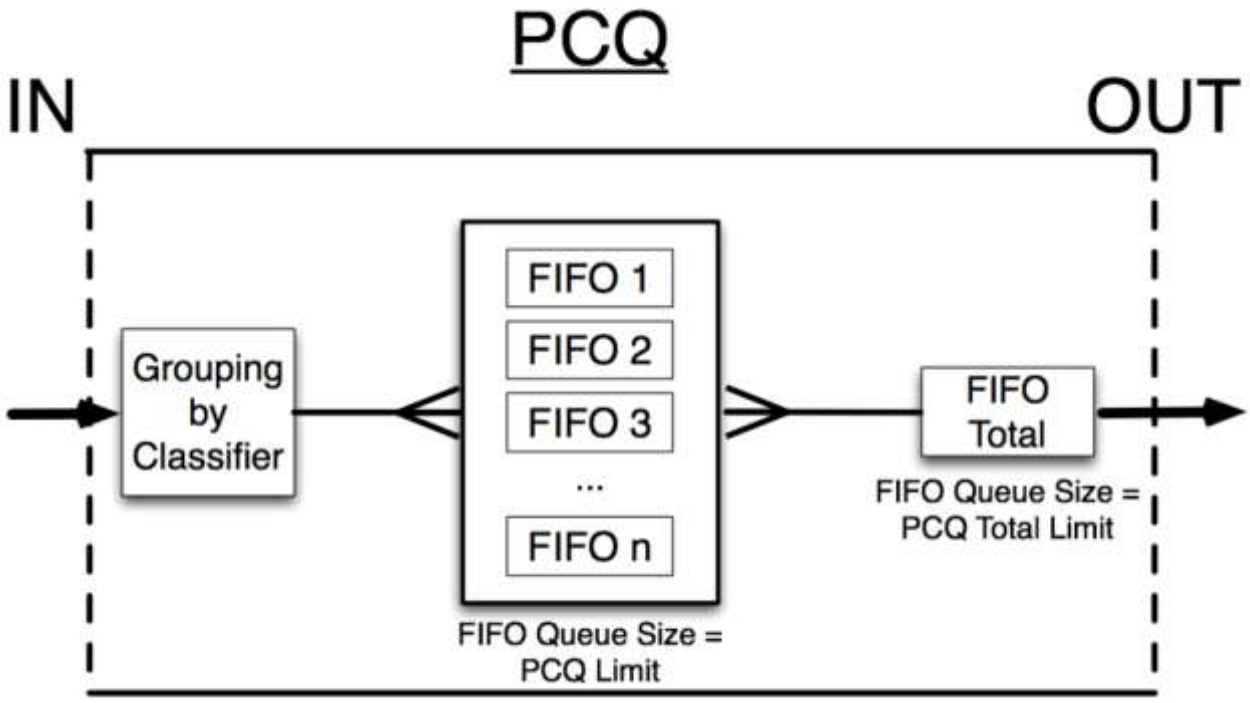
- Вы имеете более чем 400 клиентов и 3 разных тарифа для подключений:
  - «Бизнес» (4Mbps/1Mbps)
  - «Стандарт» (750kbps/250kbps)
  - «Базовый» (375kbps/125kbps)

PCQ

- Очередь по подключению (Per Connection Queue – PCQ) – это тип очереди, способная делить трафик на под-потоки (sub-streams), на основе выбранных классификаторов.

- Каждый под-поток будет проходить FIFO очередь, с размером очереди, указанным в опции “pcq-limit” и с максимальной скоростью, указанной в опции “pcq-rate”.

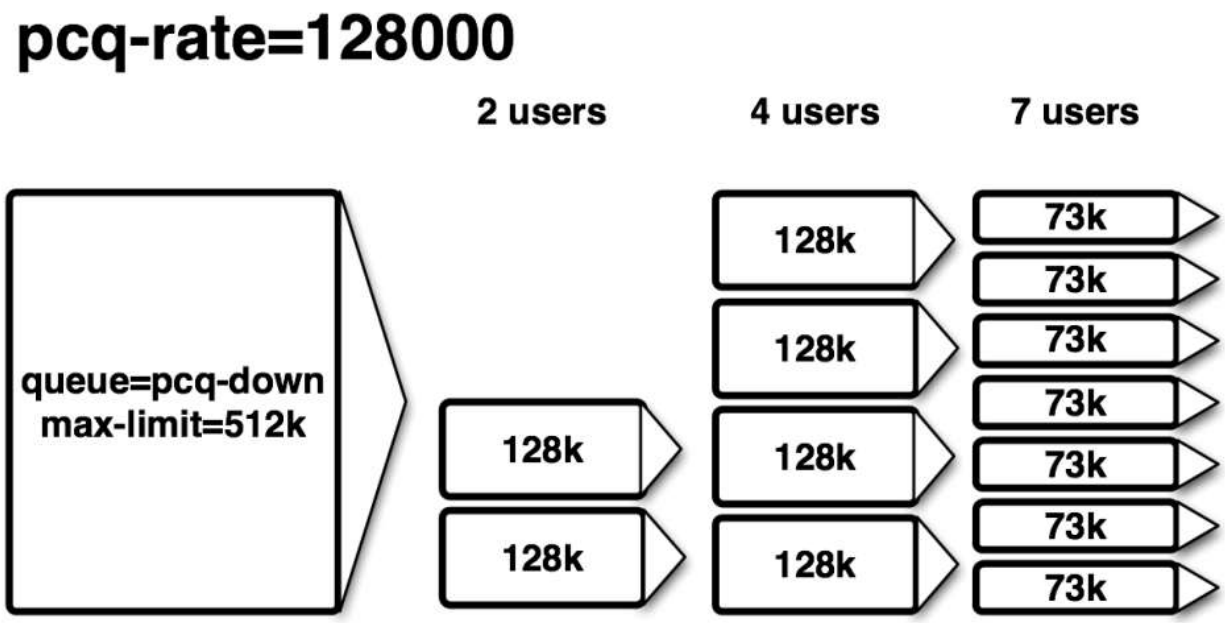


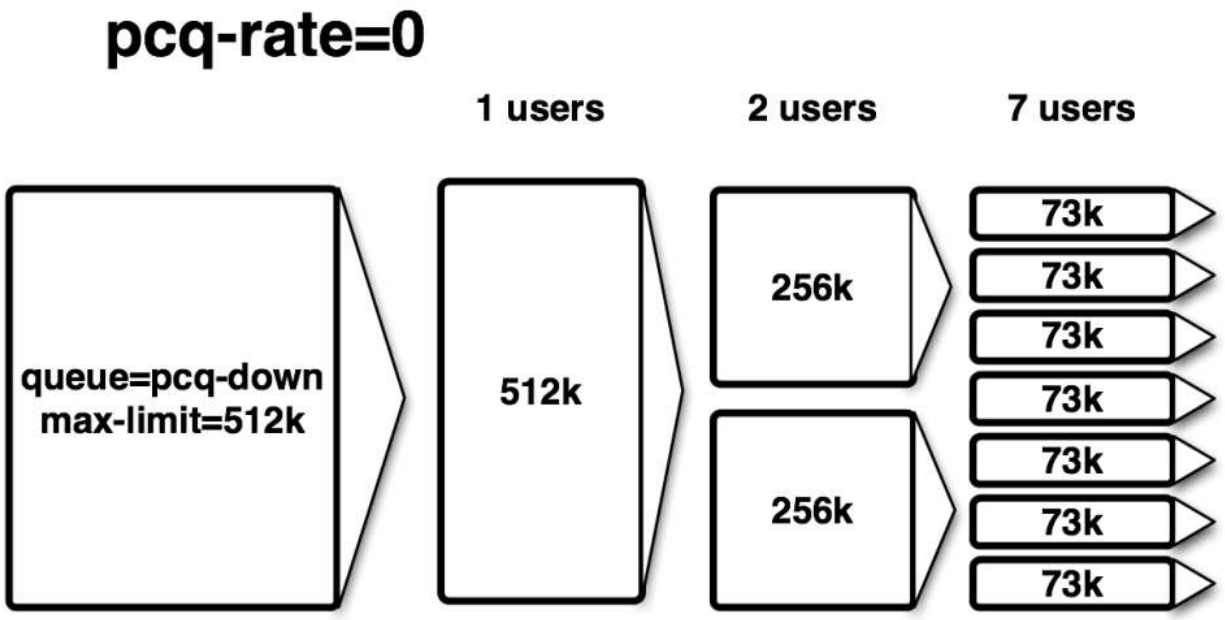


PCQ Part 2

- Для того, чтобы гарантировать, что каждый PCQ под-поток представляет собой одного конкретного клиента , мы должны создать 2 разных PCQ типа:
  - PCQ_upload – в качестве классификатора – адрес источника (source address)
  - PCQ_download - в качестве классификатора – адрес назначения (destination address)

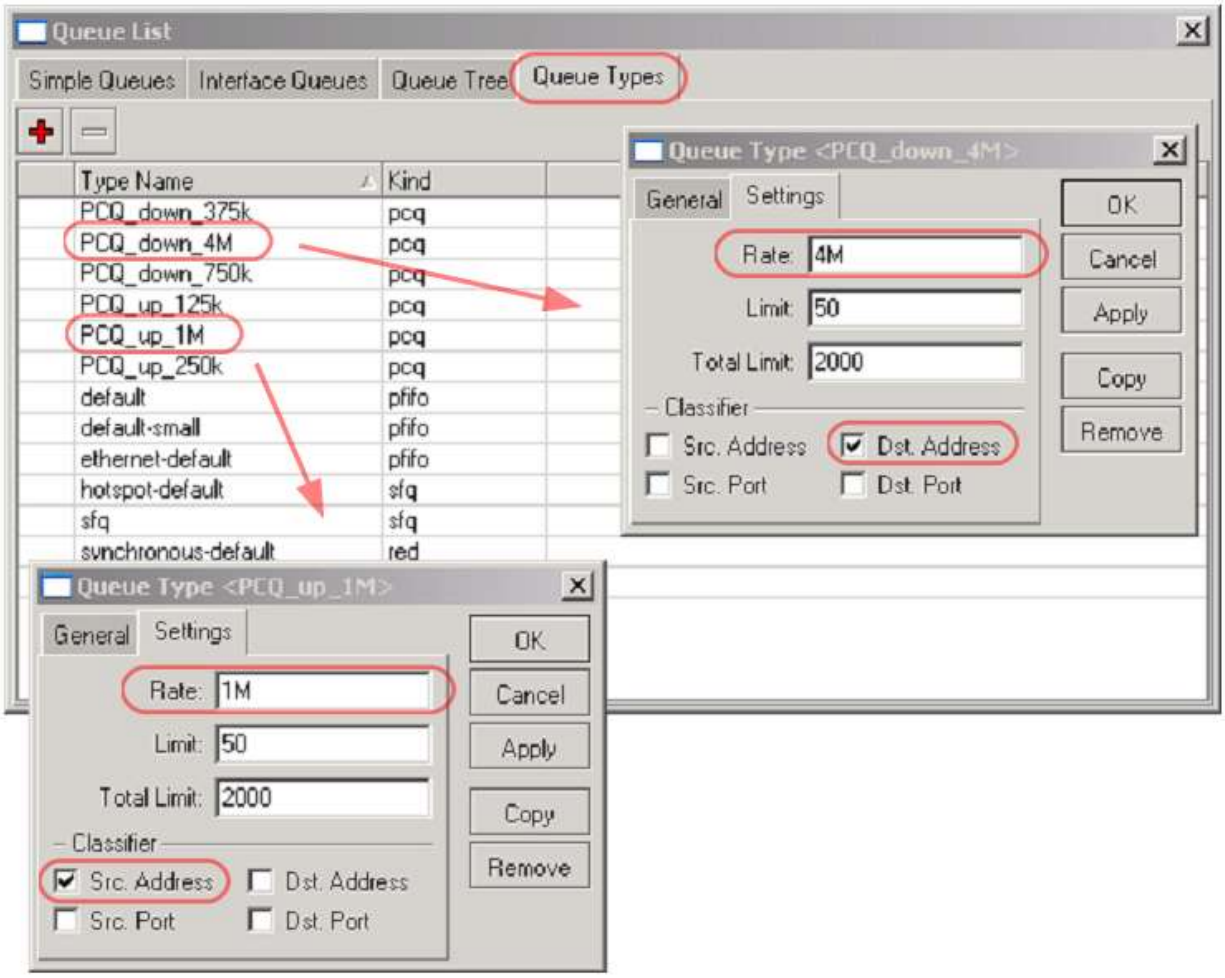
- PCQ будет распределять имеющийся трафик равномерно между под-очередями, пока скорость pcq-rate доступна (если она указана)



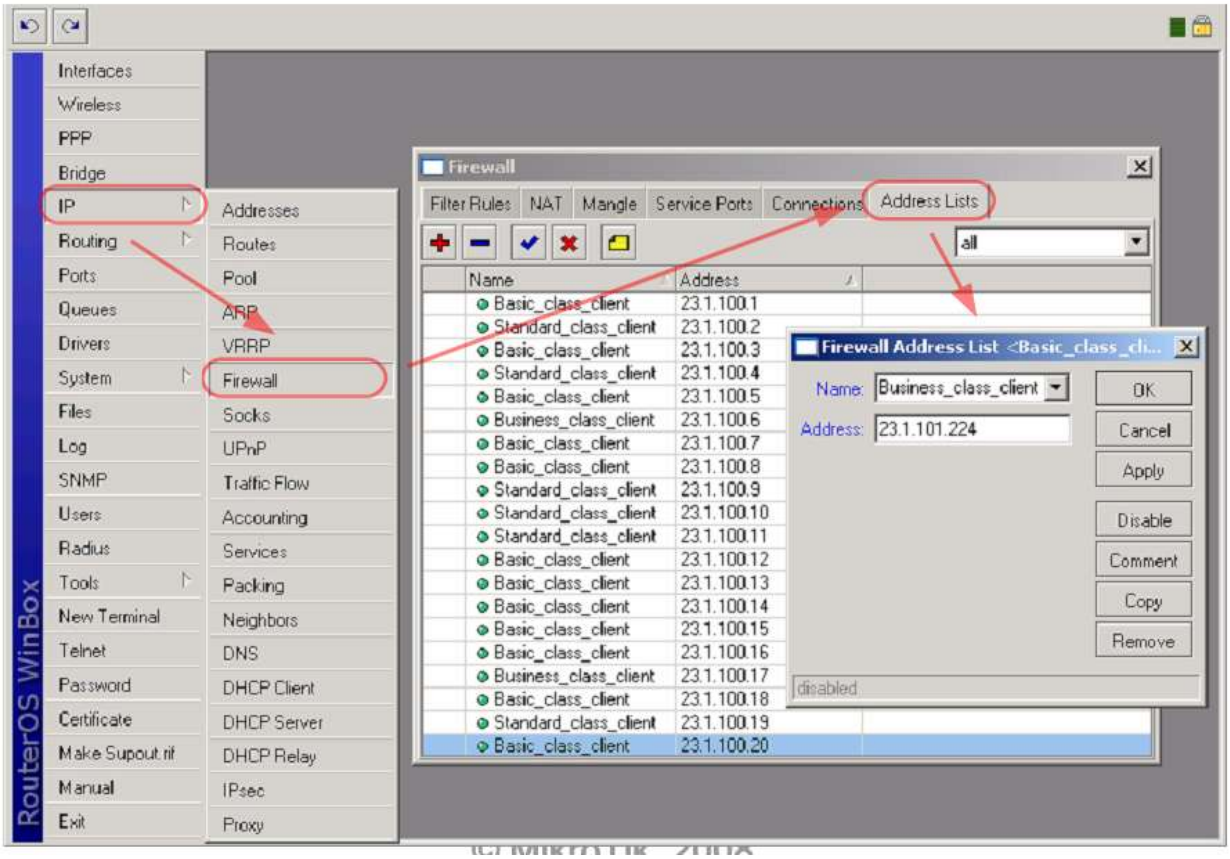


(Примечание переводчика – как видите, если оставить PCQ-rate = 0, то на каждого отдельного юзера в группе нет ограничения, но можно выставить общее ограничение на эту группу в очереди, и тогда в пределах этого общего ограничения юзеры будут делить конкурентно выделенную пропускную способность канала – т.е. если все юзеры из этой группы ушли спать , а один остался – он прокачает всю скорость группы)

Обзор PCQ Тип: Winbox



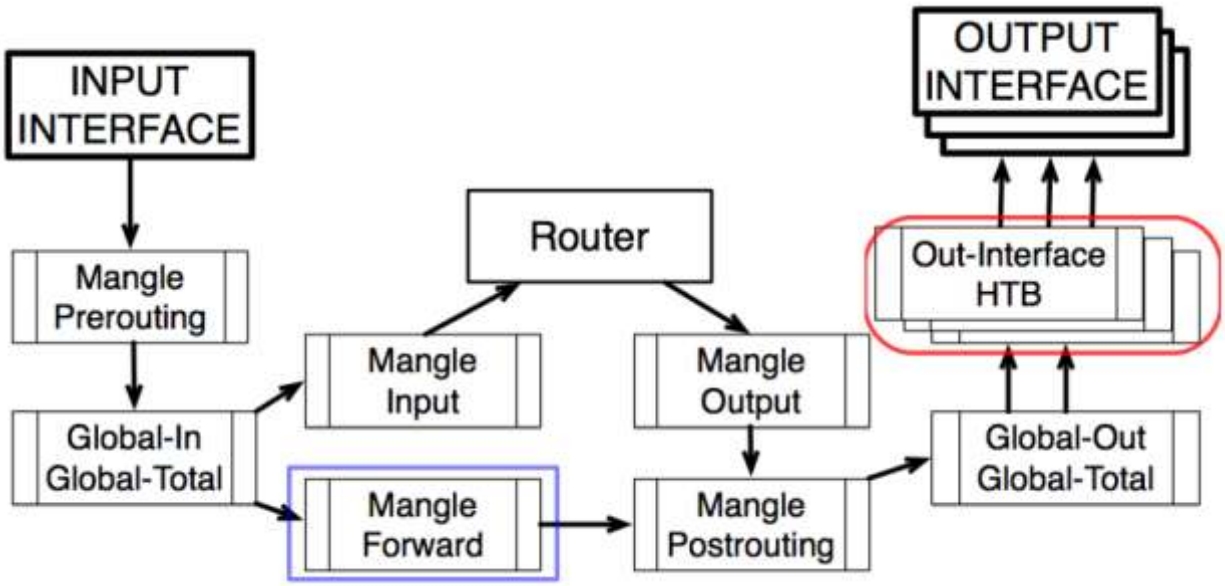
Address Lists



- Адрес листы (Address lists) были введены для того, чтобы определить множество IP адресов в одно и тоже правило фаервола, таким образом сокращается общее число правил фаервола и повышается производительность роутера.

- Адресные листы могут быть созданы:
  - Вручную
  - Автоматически из PPP профиля – просто укажите опцию address-list и как только клиент соединиться, он будет добавлен в надлежащий адрес-лист.
  - Автоматически через RADIUS-атрибут: “Mikrotik:19” (Mikrotik-Address-List).

Где ?



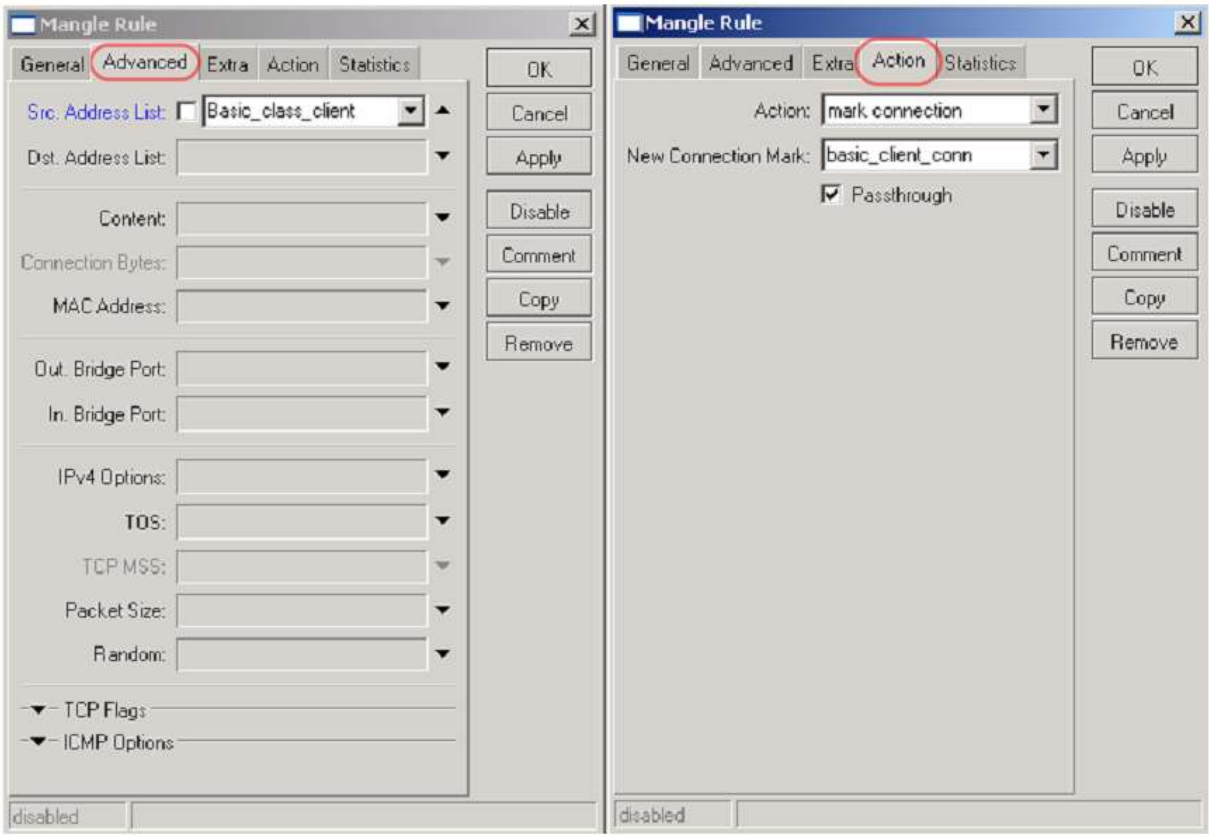
Маркировка пакетов

- Используйте действие (action) - “connection-mark”, чтобы классифицировать все соединения, основанные на клиентских адресных листах.

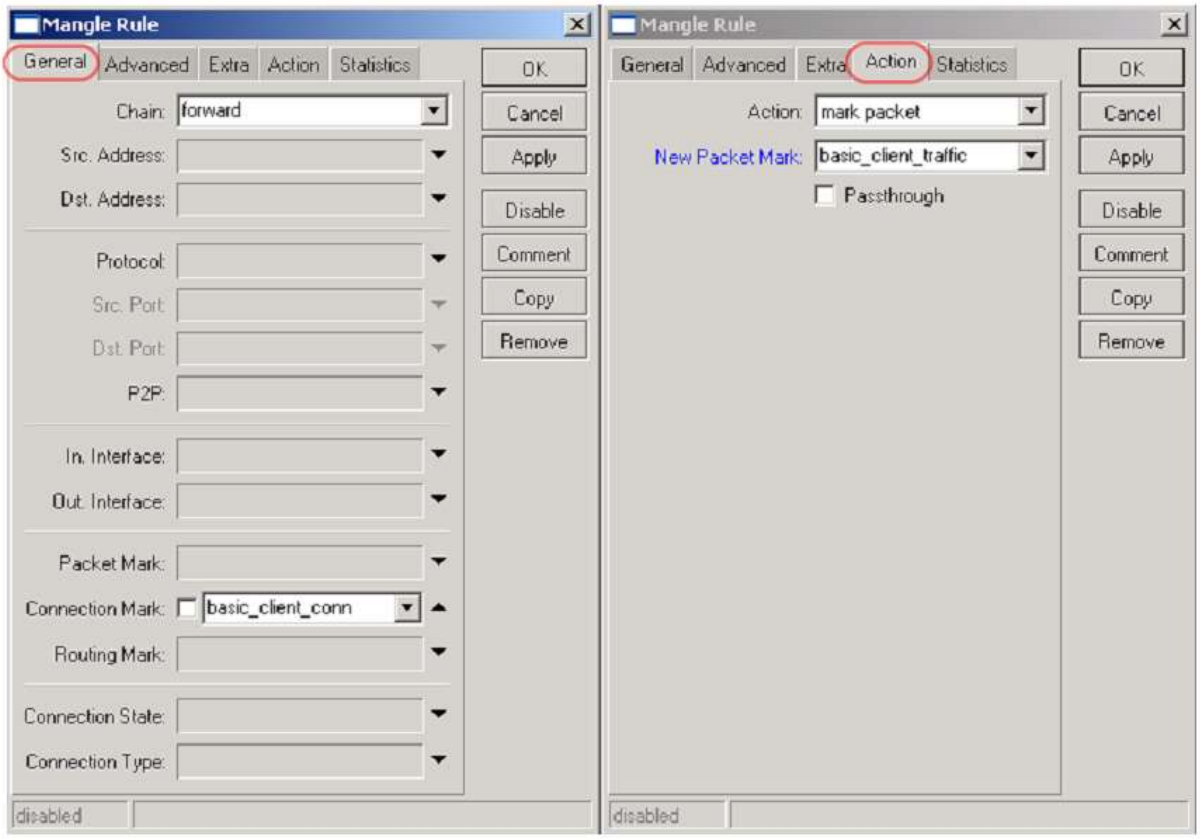
- Используйте действие - “packet-mark”, чтобы классифицировать весь трафик, основанный на маркировке соединений

- Вопросы для размышления:
  - Какая скорость должна быть доступна для клиентов, с тарифом «бизнес», если осуществляется загрузка (downloading) с клиентом тарифа «Базовый»?
  - У вас есть еще немаркированный трафик?

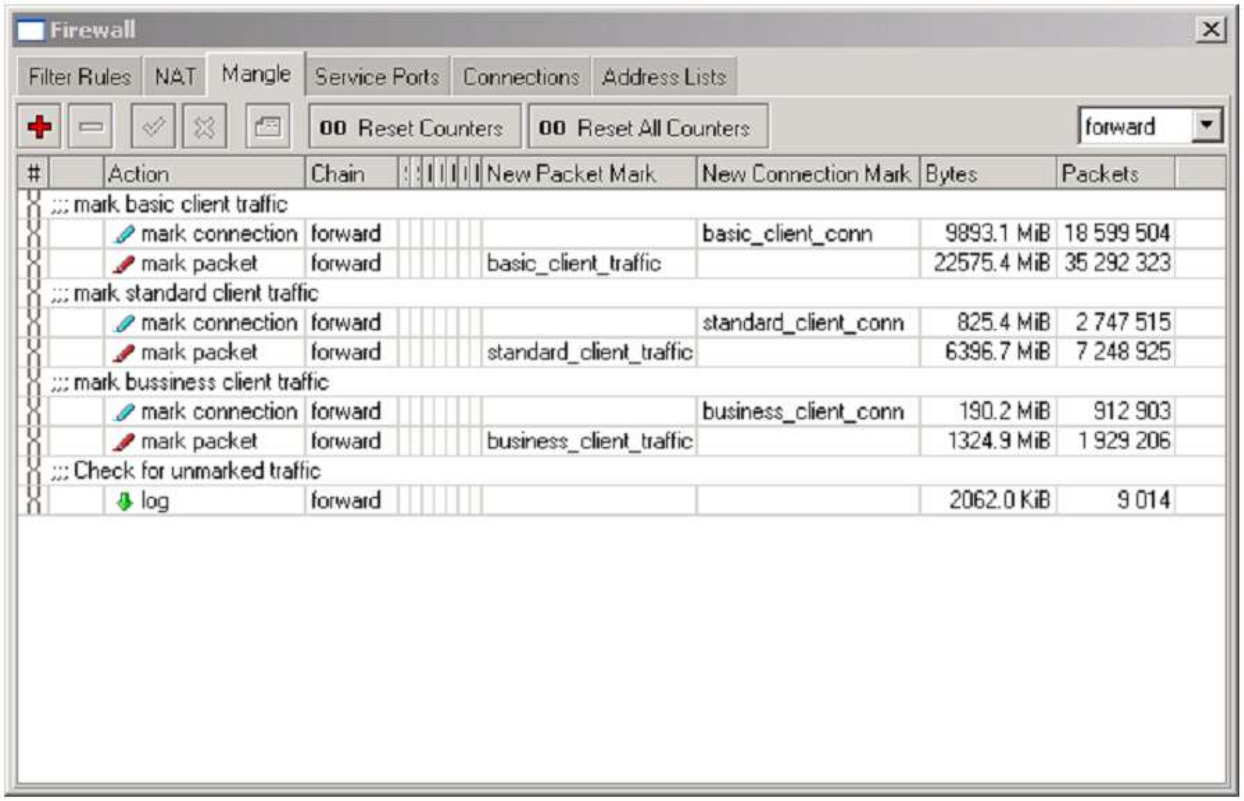
Правило «Connection-mark»



Правило «Packet-mark»



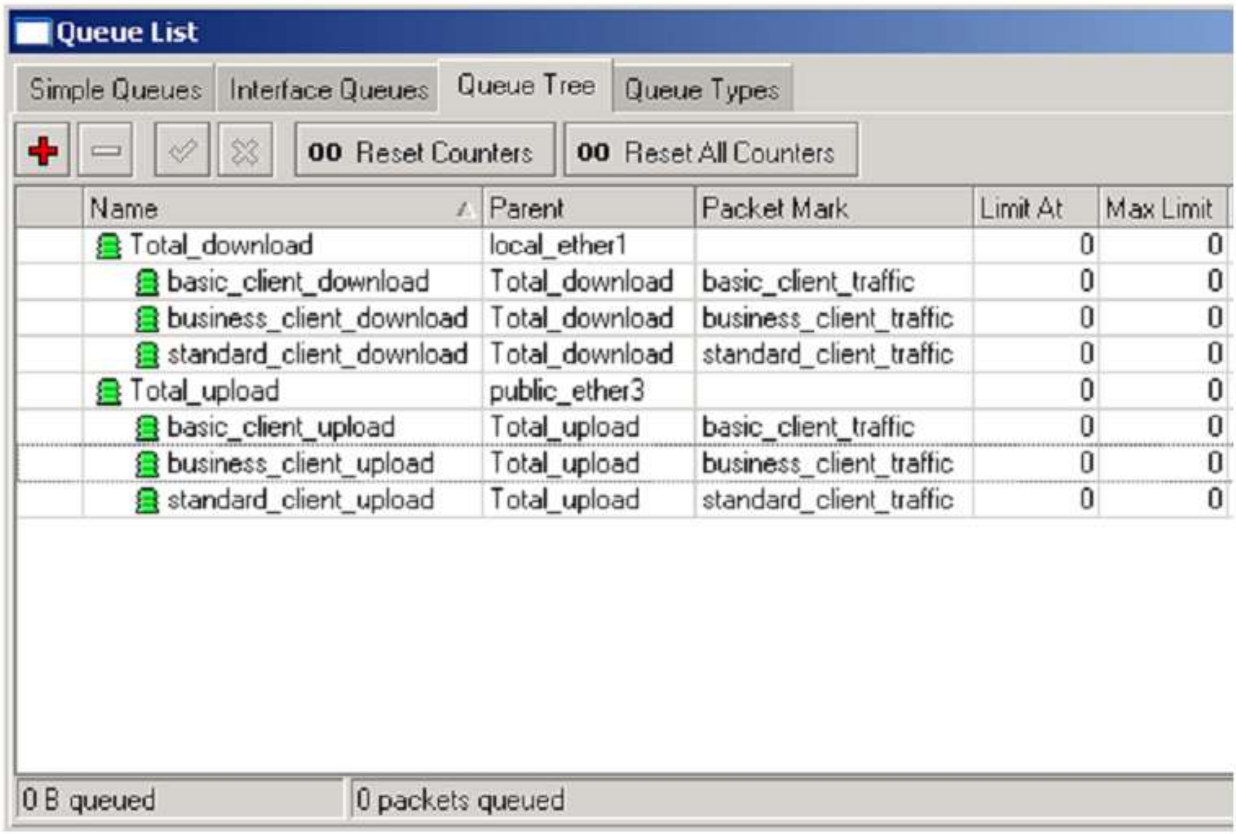
Обзор Mangle: Winbox



Обзор Mangle: Export

```
/ ip firewall mangle
add chain=forward src-address-list=Basic_class_client action=mark-connection \
    new-connection-mark=basic_client_conn passthrough=yes comment="mark basic \
    client traffic" disabled=no
add chain=forward connection-mark=basic_client_conn action=mark-packet \
    new-packet-mark=basic_client_traffic passthrough=no comment="" disabled=no
add chain=forward src-address-list=Standard_class_client \
    action=mark-connection new-connection-mark=standard_client_conn \
    passthrough=yes comment="mark standard client traffic" disabled=no
add chain=forward connection-mark=standard_client_conn action=mark-packet \
    new-packet-mark=standard_client_traffic passthrough=no comment="" \
    disabled=no
add chain=forward src-address-list=Business_class_client \
    action=mark-connection new-connection-mark=business_client_conn \
    passthrough=yes comment="mark bussiness client traffic" disabled=no
add chain=forward connection-mark=business_client_conn action=mark-packet \
    new-packet-mark=business_client_traffic passthrough=no comment="" \
    disabled=no
add chain=forward action=log log-prefix="" comment="Check for unmarked \
    traffic" disabled=no
```

Обзор Queue Tree: Winbox

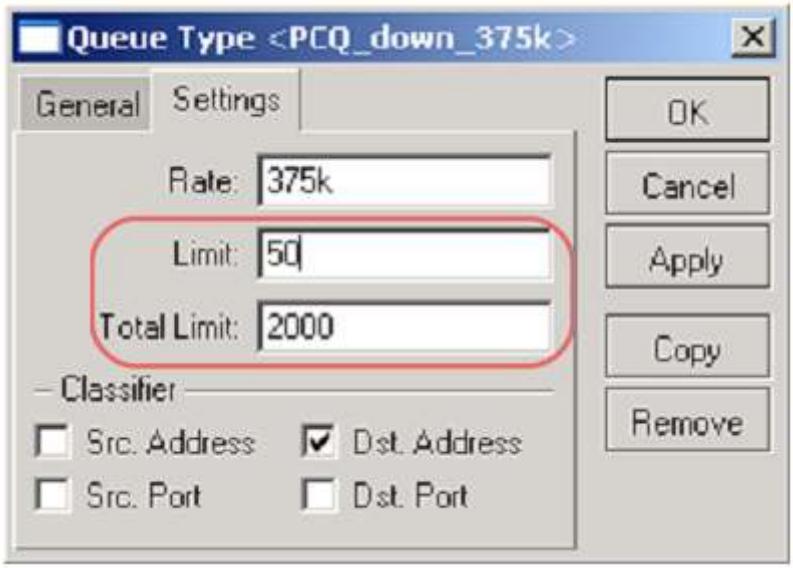


Обзор Queue Tree: Export

```
/ queue tree
add name="Total_download" parent=local_ether1 packet-mark="" limit-at=0 \
    queue=default priority=1 max-limit=0 burst-limit=0 burst-threshold=0 \
    burst-time=0s disabled=no
add name="basic_client_download" parent=Total_download \
    packet-mark=basic_client_traffic limit-at=0 queue=PCQ_down_375k priority=8 \
    max-limit=0 burst-limit=0 burst-threshold=0 burst-time=0s disabled=no
add name="standard_client_download" parent=Total_download \
    packet-mark=standard_client_traffic limit-at=0 queue=PCQ_down_750k \
    priority=4 max-limit=0 burst-limit=0 burst-threshold=0 burst-time=0s \
    disabled=no
add name="business_client_download" parent=Total_download \
    packet-mark=business_client_traffic limit-at=0 queue=default priority=1 \
    max-limit=0 burst-limit=0 burst-threshold=0 burst-time=0s disabled=no
add name="Total_upload" parent=public_ether3 packet-mark="" limit-at=0 \
    queue=default priority=8 max-limit=0 burst-limit=0 burst-threshold=0 \
    burst-time=0s disabled=no
add name="basic_client_upload" parent=Total_upload \
    packet-mark=basic_client_traffic limit-at=0 queue=PCQ_up_125k priority=8 \
    max-limit=0 burst-limit=0 burst-threshold=0 burst-time=0s disabled=no
add name="standard_client_upload" parent=Total_upload \
    packet-mark=standard_client_traffic limit-at=0 queue=PCQ_up_250k \
    priority=4 max-limit=0 burst-limit=0 burst-threshold=0 burst-time=0s \
    disabled=no
add name="business_client_upload" parent=Total_upload \
    packet-mark=business_client_traffic limit-at=0 queue=PCQ_up_1M priority=1 \
    max-limit=0 burst-limit=0 burst-threshold=0 burst-time=0s disabled=no
```

PCQ Queue Size

(Размер очереди PCQ)



Можно рассчитать, сколько очередь займет памяти:

Total_limit = X может занять до:

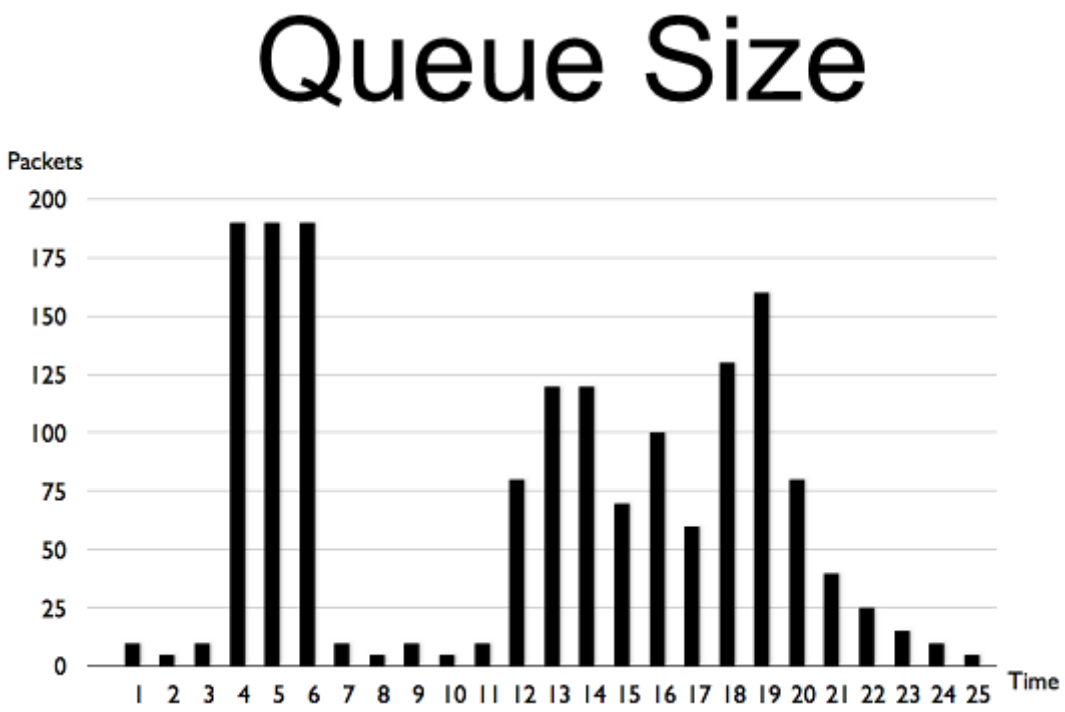
X*(2000 bytes + 200 bytes) of RAM

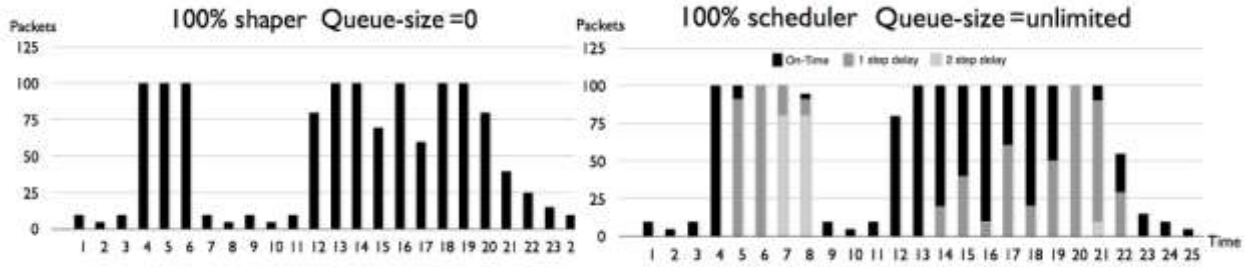
2000 bytes – buffer for 1 packet
200 bytes – service data for 1 packet

total_limit = 2000 =< 4,2MB RAM
total_limit = 5000 =< 10,5MB RAM

- Эти значения подходят для 40 пользователей - для заполнения очереди. (потому что total_limit/limit = 2000/50 = 40).
- Можно менять значения “limit”, "total_limit" в зависимости от к-ва пользователей и поставленных задач (смотрите следующую страницу про QueueSize (Размер очереди))
- Должно быть не менее 10-20 пакетов в очереди, доступных для каждого пользователя

Размер очереди





(*прим. Переводчика):* Смысл такой – если очередь короткая, то - когда не будет хватать пропускной способности внешнего канала - «лишний трафик» будет дропаться.

Чем длиннее очередь – тем меньше дропов, шейпер будет пытаться собирать «лишний трафик» в буфер и пытаться его пропустить. При этом - обратная сторона медали – возрастут задержки трафика.

Таким образом, исходя из ваших задач – вы решаете – что для вас существенно, потери или задержки трафика.

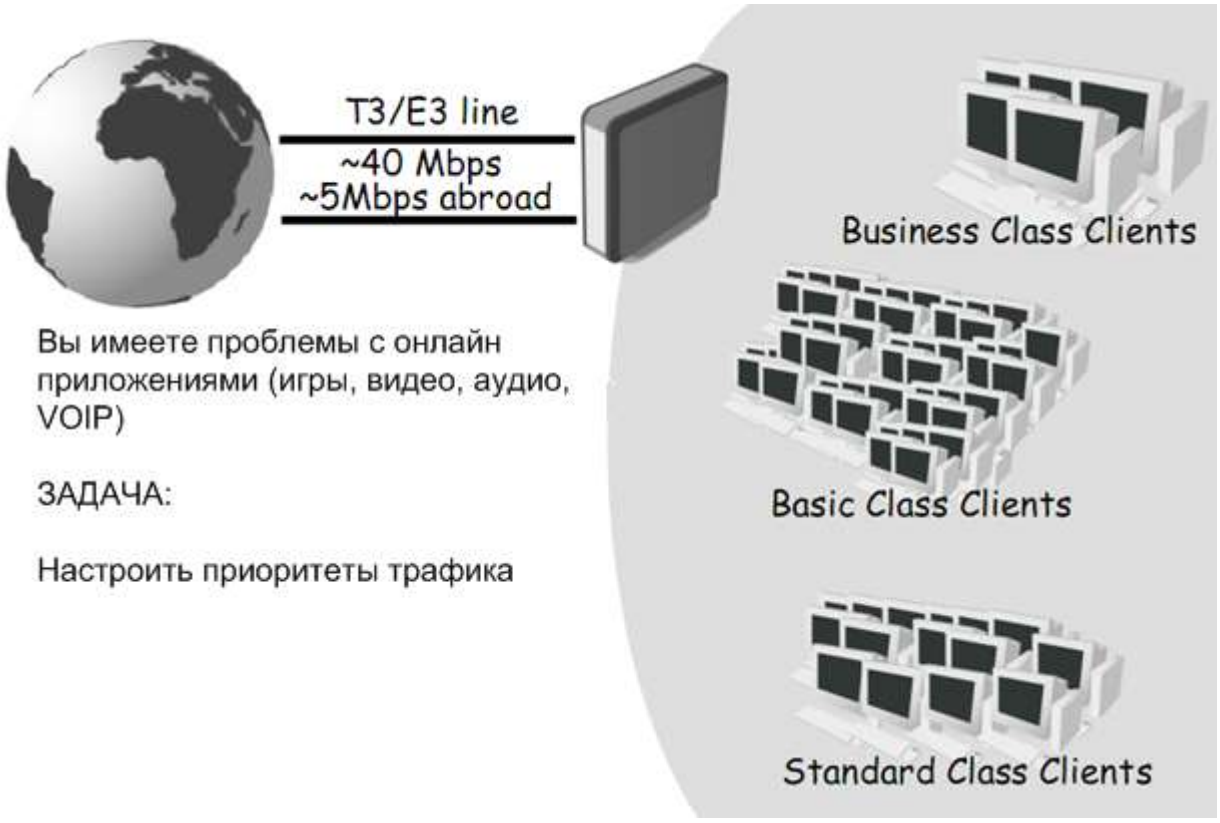
Настройка PCQ

- Есть ~340 клиентов, с тарифом «Базовый», то :
  - pcq_limit = 40
  - pcq_total_limit = 7000 (~20*340) (~15MB)

- Есть ~40 клиентов, с тарифом «Стандарт», то :
  - pcq_limit = 30
  - pcq_total_limit = 1000 (~20*40) (~2MB)

- Есть ~20 клиентов, с тарифом «Бизнес», то :
  - pcq_limit = 20 (!!!)(*прим. Переводчика – почему здесь восклицательные знаки ☺? – может потому что для бизнес клиентов важны низкие задержки трафика, поэтому очередь самая короткая, а приоритет самый высокий?)*
  - pcq_total_limit = 500 (~20*20) (~1MB)

Приоритезация трафика

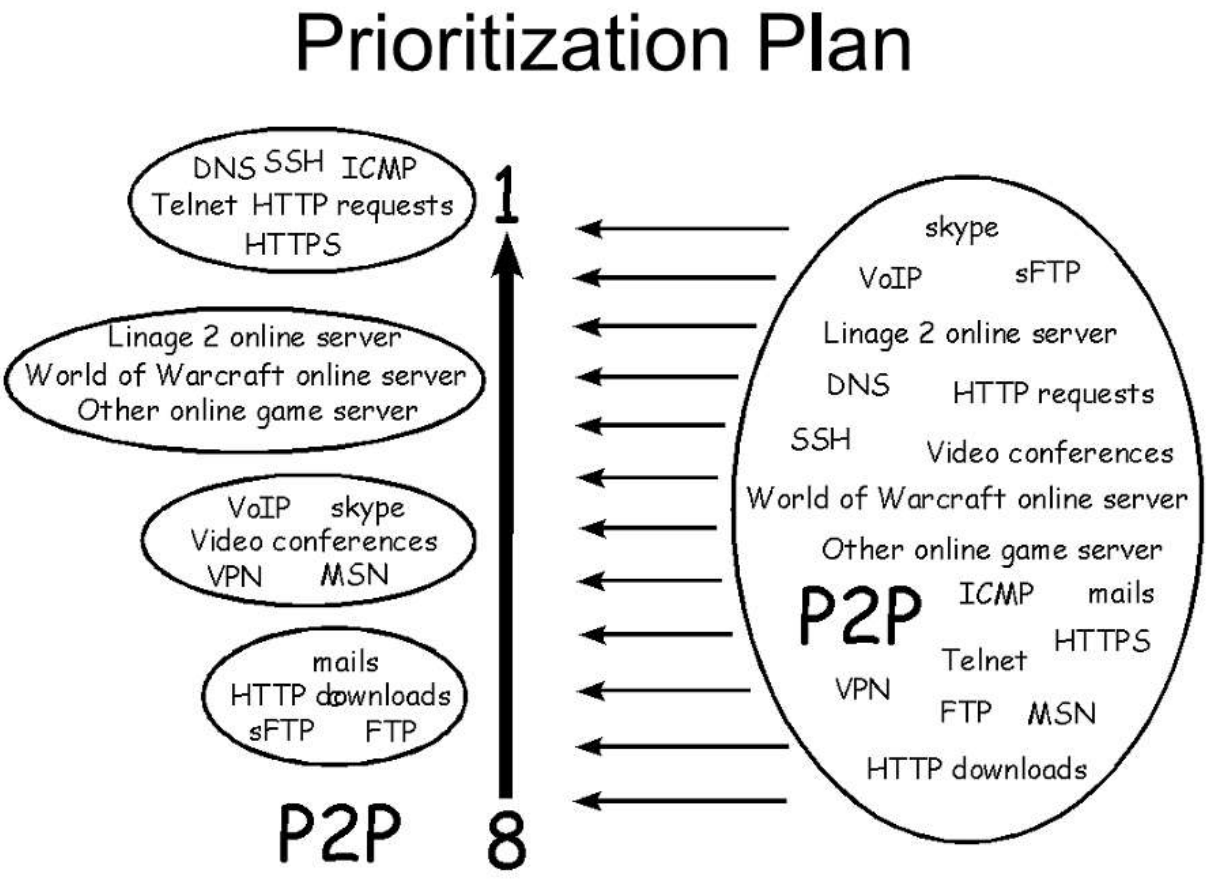


Вы имеете проблемы с онлайн приложениями (игры, видео, аудио, VOIP)

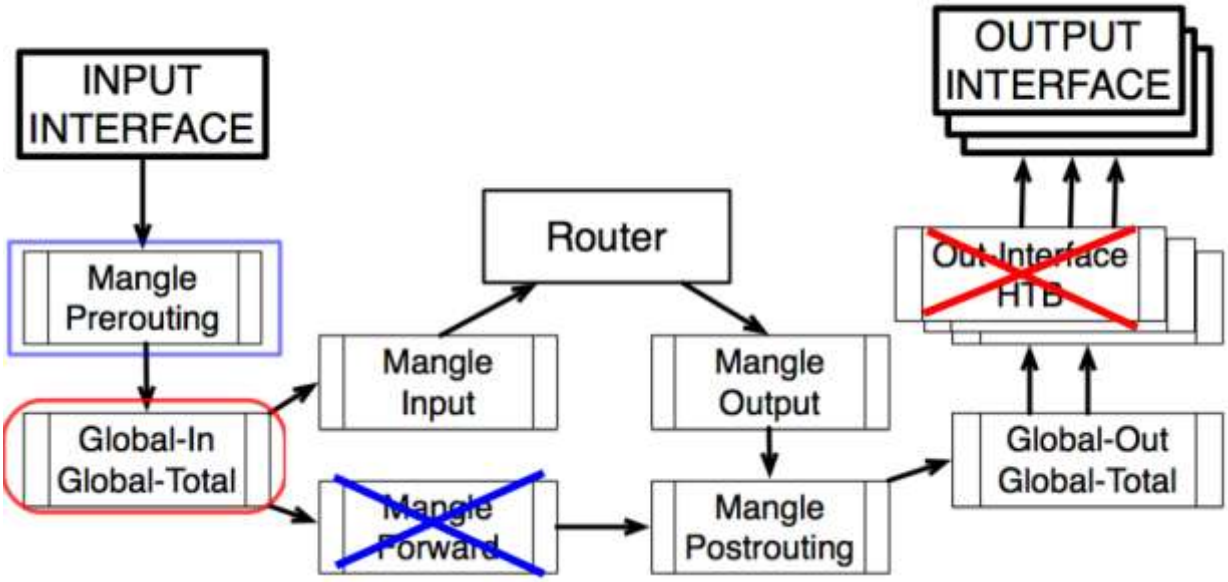
ЗАДАЧА:

Настроить приоритеты трафика

План приоритетов



Где?



Как?

Group	Service	Protocol	Dst-Port	Other conditions
P2P_services	P2P			p2p=all-p2p
Download_services	Mails	TCP	110	
		TCP	995	
		TCP	143	
		TCP	993	
		TCP	25	
	HTTP downloads	TCP	80	Connection-bytes=500000-0
	FTP	TCP	20	
		TCP	21	
	SFTP	TCP	22	Packet-size=1400-1500
Ensign_services	DNS	TCP	53	
		UDP	53	
	ICMP	ICMP	-	
	HTTPS	TCP	443	
	Telnet	TCP	23	
	SSH	TCP	22	Packet-size=0-1400
	HTTP requests	TCP	80	Connection-bytes=0-500000
User_requests	Online game servers			Dst-address-list=user_requests
Communication_services	VoIP			
	Skype			
	Video conferences			
	VPN			
	MSN			

Приоритеты

- Создайте маркировку пакетов в mangle цепочке “Prerouting” для приоритезации трафика в global-in очереди
  - Флагманские сервисы (Priority=1)
  - Пользовательские запросы (Priority=3)
  - Коммуникационные сервисы (Priority=5)
  - Сервисы закачек (Priority=7)
  - Сервисы P2P (Priority=8)

(*прим. Переводчика – тут не расписано подробно, как все же разметить пакеты по типу трафика. Потому что это тема отдельной книги – у каждого свои задачи и приоритеты. Особую сложность вызывает маркировка p2p трафика (сигнатуры все время меняются, анализ содержимого пакетов – это весьма ресурсоемкое занятие, а при шифровании трафика p2p клиентами – все усилия «словить и зашейпить» этот трафик - сводятся к нулю), да и тот же Skype вы просто так не промаркируете….)*