Содержание

Повышение отказоустойчивости - сетевые карты с режимом bypass

Почему Вы рекомендуете покупать карты SILICOM?
Причины следующие:

- карты обладают функциональностью bypass
- в комплект поставки можно включить лицензии на требуемые драйвера для максимальной производительности - DPDK & Libzero в маркировке таких карт есть обозначение -SQ1 (пример для 10GbE карты)

Работает ли байпасс в сетевых картах SILICOM при отключении питания?

- Оптический байпасс - работает при отключенном питании (проверено на карте PE210G2BPI9-SR-SQ1 short range/fiber).
- Медный байпасс - при отключении питания работает (проверено на карте PEG6BPi6)

Есть ли ручное управление байпасс в сетевых картах SILICOM?
DPI управляет bypass самостоятельно.
В случае необходимости ручное управление bypass возможно с помощью утилиты bpctl_util:

- bpctl_util all get_bypass - получить состояние bypass
- bpctl_util all set_bypass on - активировать bypass
- bpctl_util all set_bypass off - деактивировать bypass

Проблема: приобрели использованную карту и не работает bypass, что делать?
Проблема связана с картой перенастроенной в качестве стандартной, то есть с отключенной байпасс функциональностью.
Диагностика:

```
bpctl_util all get_std_nic
07:00.0 standard
07:00.1 slave
07:00.2 standard
07:00.3 slave
```

Должно быть non-standard.

Для установки карты в режим bypass требуется выполнить следующее:

```
bpctl_util all set_std_nic off
```

Команда переключает режим в не стандартный, то есть с bypass режимом.

Пояснение относительно времени переключения bypass?
Включение bypass длится короткий временной интервал около 0.5 с (по умолчанию), но по причинам передоговаривания интерфейсов в итоге может занять больший временной интервал, ниже пояснение относительно переключения bypass от производителя.
Такая длительность переключения может влиять на BGP, OSPF и другие механизмы, за счет

кратковременного разрыва связи (продолжительность может варьироваться см. описание ниже) или нескольких разрывов связи как в случае перезагрузки сервера или сервиса (*сервис →разрыв→bypass→разрыв→ сервис*), в этом случае время восстановления сессий (BGP, OSPF) зависит от их настроек и может длиться до нескольких десятков секунд. Для уменьшения данного интервала необходимо провести самостоятельно конфигурирование для уменьшения времени восстановления сессий после разрыва связи. **Например**, на оборудовании Juniper настраивается hold-timer down 500ms, что позволяет избежать разрыва BGP сессии и перестройки таблиц маршрутизации:

```
set interfaces <ifname> hold-time up 500 down 500
```

где 500 мс таймаут ожидания, прежде чем изменится operational status интерфейса

```
Basically, the time for the bypass mechanism to switch from one mode to
another is 10mS.
The timing that you are seeing relates to re-establish the link and then re-
establish
the connection (with new routing tables in switches and devices).
This switch to bypass mode is done in our product by physically connecting
the pair of
the ports together (wire to wire). This means that when this happen our
product is actually out
of the picture and the start of the traffic with this new connection will
depend on
the two networking devices (router / switch / device) on how they link
together and how
they establish the connection again. You can try to  force fix mode (not
auto-neg,
change to force 1G FD or so) this might reduce the time needed for the
negotiation.
Not sure how much.
For the change from bypass mode to normal mode - all the above also stand as
well.
The networking devices (router / switch / device) loss the link with each
other and
starts establish the connection with the Silicom NIC . Here you have more
control as
the link is done between the two devices and your system (Check that all the
devices
are set to the same speed settings)
From our customer and our experience a 1-3sec is a reasonable time to get
the Copper 1G link
to be establish between 2 network devices.
```

все хорошо кроме того, что один порт на сетевой работает в режиме байпасс и не фильтрует трафик.
Если в конфиге прописано (in/out_dev), но не переключает, то попробуйте сброс bypass свитча в карте в начальное состояние:

```
bpctl_util all set_bypass off
bpctl_util all set_dis_bypass off
```

```
bpctl_util all set_bypass_pwoff on
bpctl_util all set_bypass_pwup on
bpctl_util all set_std_nic off
bpctl_util all get_bypass_change on
bpctl_util all get_tx on
bpctl_util all get_tpl off
bpctl_util all get_wait_at_pwup off
bpctl_util all get_hw_reset off
bpctl_util all get_disc off
bpctl_util all get_disc_change off
bpctl_util all get_dis_disc off
bpctl_util all get_disc_pwup off
bpctl_util all get_wd_exp_mode bypass
bpctl_util all get_wd_autoreset disable
```

Если не помогает, значит дефект карты, меняйте ее по гарантии.

Настройка Juniper, чтобы перевод в bypass и обратно не приводил к перестроению маршрутов.

```
set interfaces <ifname> hold-time up 500 down 500

show xe-5/2/0
 description "-= 20G UPLINK LAGG =-";
hold-time up 1000 down 1000;
gigether-options {
 802.3ad ae1;
}
```

Настройка Cisco, чтобы перевод в bypass и обратно не приводил к перестроению маршрутов.

```
int fa0/0
ip bgp fast-external-fallover deny
```

Примечание:

```
BGP Fast-external-fallover command terminates external BGP sessions of any
directly adjacent peer if the link used to reach the peer goes down; without
waiting for the hold-down timer to expire
```

Список всех dna интерфейсов и их MAC
grep ^ /sys/class/net/dna?/address