s•terra®

Ваш ориентир в мире безопасности

Отказоустойчивые решения на базе продуктов С-Терра

Шпаков Андрей
Ведущий инженер
Отдел технического консалтинга

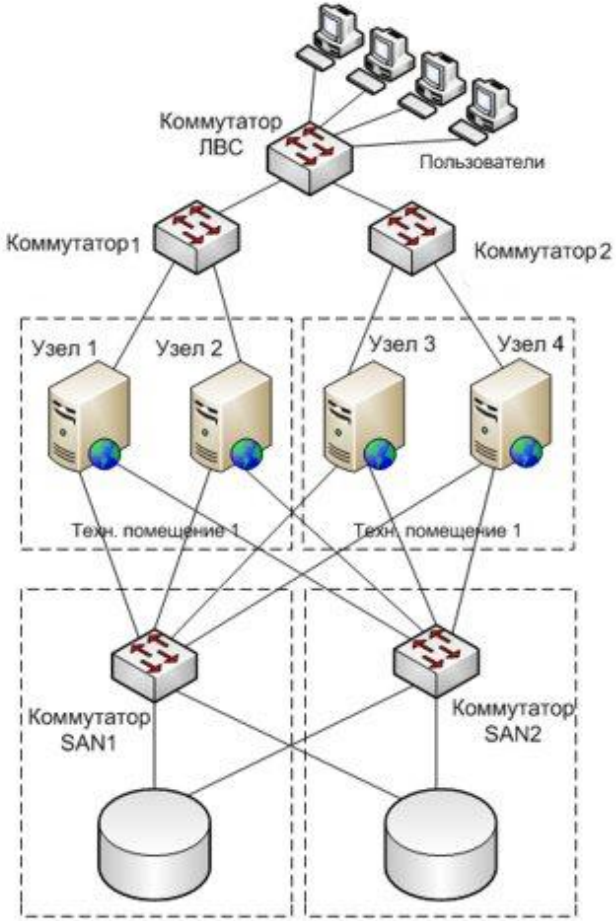
О чем пойдет речь

- Какие виды отказоустойчивости есть в продуктах С-Терра?
- Резервирование шлюзов: технологии VRRP, RRI
- Резервирование интерфейсов: Link-aggregation
- Резервирование каналов связи: GRE-over-IPsec, Object tracking
- Сценарии балансировки нагрузки с резервированием

Отказоустойчивость

Что должно резервироваться?

- Отдельные компоненты АП
- Шлюзы
- интерфейсы на устройствах
- Каналы связи



Резервирование компонентов АП

G-7000-D-5067-6-RED-ST-KC1	3 460	300 000	Linux	Kraftway Express Lite EL20, Redundant, 6xLAN 1Gb, Rack mount 1U, 3yOS_5D
G-7000-D-5067-4-RED10-ST-KC1	3 520	300 000	Linux	Kraftway Express Lite EL20, Redundant, 4xLAN 1Gb, Rack mount 1U, 3yOS_5D
G-7000-D-5067-6-RED10-ST-KC1	3 760	300 000	Linux	Kraftway Express Lite EL20, Redundant, 6xLAN 1Gb, Rack mount 1U, 3yOS_5D
G-7000-D-5067-6-ST-KC2	2 620	320 000	Linux	Kraftway Express Lite EL20, 6xLAN 1Gb, Rack mount 1U, 3yOS_5D, KB-SOBOL 3.0 K1
G-7000-D-5067-4-RED-ST-KC2	3 350	320 000	Linux	Kraftway Express Lite EL20, Redundant, 4xLAN 1Gb, Rack mount 1U, 3yOS_5D, KB-S K12 V1
G-7000-D-5067-4-RED10-ST-KC2	3 660	320 000	Linux	Kraftway Express Lite EL20, Redundant, 4xLAN 1Gb, Rack mount 1U, 3yOS_5D, KB-S K12 V1
G-7000-D-7123-6-2-RP-ST-KC2	6 380	320 000	Linux	Huawei RH1288 V3, 6xLAN 1Gb, 2x10GBase-SR, Rack mount 1U, Redundant PS, 3yO 01
G-7000-D-7123-4-4-RP-ST-KC2	7 510	320 000	Linux	Huawei RH1288 V3, 4xLAN 1Gb, 4x10GBase-SR, Rack mount 1U, Redundant PS, 3yO 01
G-7000-D-7124-4-ST-KC2	5 280	320 000	Linux	Huawei RH1288 V3, 4xLAN 1Gb, Rack mount 1U, 3yOS, KB-SOBOL 3.0 K12 V1
G-7000-D-7124-6-ST-KC2	5 510	320 000	Linux	Huawei RH1288 V3, 6xLAN 1Gb, Rack mount 1U, 3yOS, KB-SOBOL 3.0 K12 V1

Прайс - https://www.s-terra.ru/upload/medialibrary/06d/price_list_4.1_ST.pdf

- Жесткие диске в массиве RAID1, RAID 10.
- Резервирование жестких дисков
- Гарантия на все АП – 3 года
- Параметр MBTF – время наработки на отказ

Резервирование шлюзов безопасности

VRRP (*англ*. Virtual Router Redundancy Protocol, RFC 3768) — сетевой протокол, предназначенный для увеличения доступности сетевых устройств, в частности шлюзов безопасности.

Особенности протокола

- Несколько шлюзов объединяются в один виртуальный
- Используется единый виртуальный IP и маршрут
- Масштабируемость шлюзов, любая топология
- Кластер Active/Passive
- Среднее время перестроения туннеля ~ 90 секунд для версии 4.1
  ~ 30 секунд для версии 4.2.
- Доступен как для L2, так и для L3

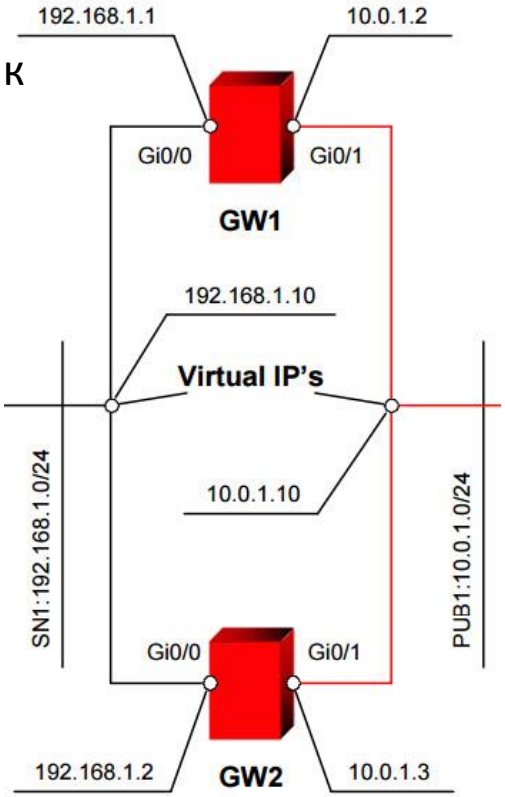
VRRP. Логика работы

1. Настраиваем состояние MASTER (GW1), BACKUP(GW2), а так же виртуальный IP и маршрут.

2. MASTER-шлюз отправляет VRRP-пакеты на multicast 224.0.0.18.

3. Если BACKUP-шлюз не получает VRRP-пакетов, он становится MASTER и отправляет GARP со своим MAC-адресом.

4. Когда бывший MASTER-шлюз(GW1) восстанавливает работу, он отправляет VRRP-пакет и восстанавливает свой статус.



Конфигурация VRRP (версия 4.1)

Конфигурация /etc/keepalived.conf

MASTER-шлюз

Отслеживание работы VPN-демона

```
vrrp_script chk_dead {script "pgrep -x vpnsvc" interval 1}
```

Синхронизация VRRP-процессов:

```
vrrp_sync_group G1 { group { VI_1 VI_2 }
```

Параметры первого VRRP-процесса:

```
vrrp_instance VI_1 {
      state MASTER
      interface eth1
      garp_master_delay 10
      virtual_router_id 51
      priority 100
      advert_int 1
      virtual_ipaddress {10.0.1.10/24}
      authentication {auth_type PASS auth_pass password}
      track_script { chk_dead }
      virtual_routes {src 10.0.1.10 10.0.2.10 via
10.0.1.1}}
```

Параметры второго VRRP-процесса :

```
vrrp_instance VI_2 {…
      interface eth0
      virtual_router_id 52
      virtual_ipaddress { 192.168.1.10/24 }
      …}
```

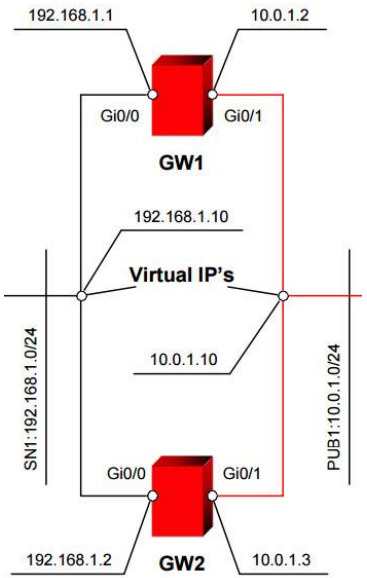
BACKUP-шлюз

вместо state MASTER будет state BACKUP:

```
vrrp_instance VI_1 {
      state BACKUP …}
```

Кроме того параметр priority должен быть изменен с 100 на 50:

```
vrrp_instance VI_1 {
      priority 50 …}
```



Конфигурация VRRP (версии 4.2)

Вся настройка из CLI

MASTER-шлюз

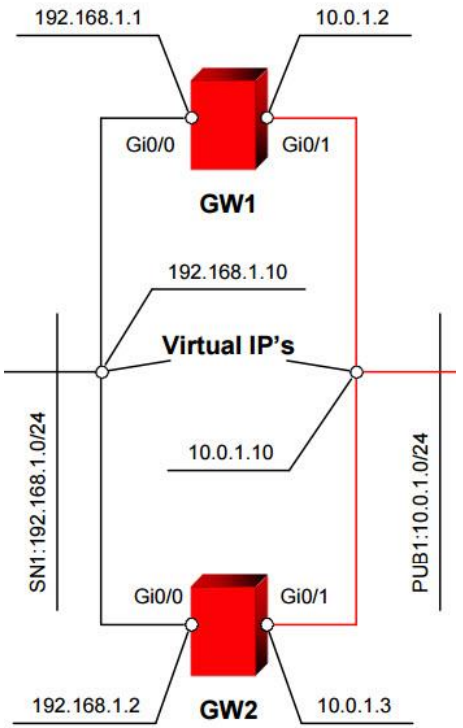
Настройка в CLI:

```
GW1(config)#interface GigabitEthernet 0/0
GW1(config-if)#vrrp 51 ip 192.168.1.10
GW1(config-if)#vrrp 51 priority 100
GW1(config-if)#vrrp 51 authentication pass123
GW1(config-if)#exit
GW1(config)#interface GigabitEthernet 0/1
GW1(config-if)#vrrp 52 ip 10.0.1.10
GW1(config-if)#vrrp 52 priority 100
GW1(config-if)#vrrp 52 authentication pass123
GW1(config-if)#exit
GW1(config)#vrrp ip route 10.0.2.0 255.255.255.0
10.0.1.1 src 10.0.1.10
GW1(config)#vrrp notify master
GW1(config)#vrrp notify backup
GW1(config)#vrrp notify fault
GW1(config)#exit
```

BACKUP-шлюз

Аналогично кроме параметра priority

```
GW2(config)#interface GigabitEthernet
0/0
GW2(config-if)#vrrp 51 priority 50
GW2(config-if)#exit
GW2(config)#interface GigabitEthernet
0/1
GW2(config-if)#vrrp 52 priority 50
GW2(config-if)#exit
```



Резервирование шлюзов безопасности

RRI (*англ*. Reverse Route Injection)— механизм, предназначенный для отказоустойчивости Remote-Access VPN.

Особенности механизма

- Отказоустойчивость и балансировка нагрузки.
- Необходимость внешнего баласировщика (маршрутизатора)
- Соединения инициируют удаленные узлы
- Работает как на статической, так и динамической cryptomap
- В отдельных случаях резервирует и провайдеров

Конфигурация RRI

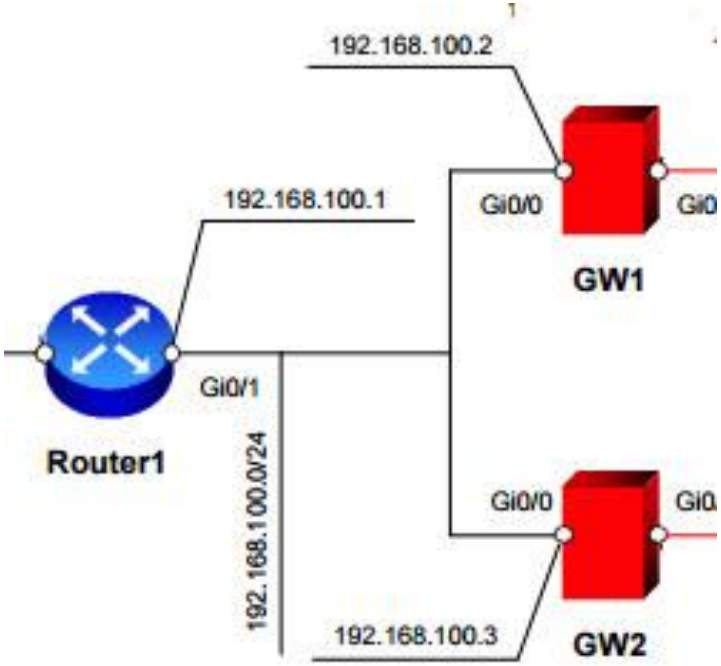
s•terra

1. Настройка RRI в криптокарте

```
GW1(config)#crypto dynamic-map DMAP 1
GW1(config-crypto-map)#match address LIST
GW1(config-crypto-map)#set transform-set TSET
GW1(config-crypto-map)#set pfs vko
GW1(config-crypto-map)#reverse-route
GW1(config-crypto-map)#exit
```

2. Настройка протокола RIP через /etc/quagga/ripd.conf

```
hostname ripd
password zebra
debug rip events
debug rip packet
router rip version 2
redistribute kernel
network eth0
distribute-list acl-in in
distribute-list acl-out out
access-list acl-in deny any
access-list acl-out permit 192.168.11.0/24
access-list acl-out deny any
log file /var/log/quagga/ripd.log
log stdout3.
```



Резервирование в IPsec – DPD

DPD (*англ*. Dead Peer Detection)— механизм обнаружения неработающего пира в рамках IKE и IPsec

Особенности механизма

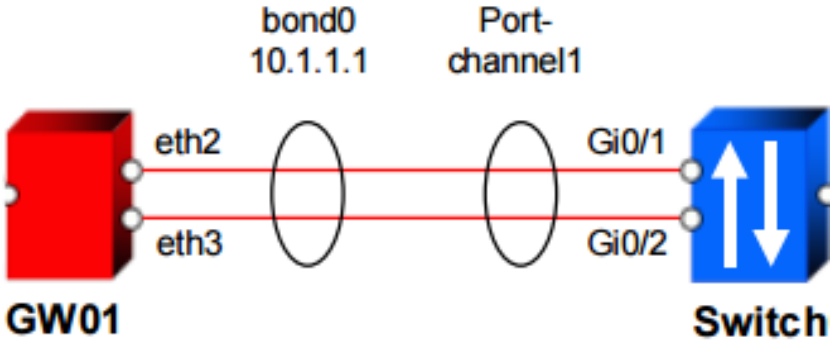
- Составная часть IPsec, не требует дополнительных портов
- Нет ответа – идем на next peer в cryptomap
- Настраивается командой **crypto isakmp keepalive** secs [retries]
- Требует настройки в случае нестабильного канала между устройствами
- Более гибкий в версии 4.2: **crypto isakmp keepalive** retry-count {retry-count}

Резервирование портов. Link-aggregation

Агрегирование каналов (*англ*. Link-aggregation)— технологии объединения нескольких параллельных линков в один логический, позволяющие увеличить пропускную способность и повысить надёжность.

Особенности:

- Описана в IEEE 802.3ad.
- Реализуется средствами ОС шлюза
- Контроль канала – протокол LACP



Настройка Link-aggregation

s•terra

Конфигурация /etc/network/interfaces

1. Внесите изменение в /etc/network/interfaces:

```
auto bond0
   iface bond0 inet static
   address 10.1.1.1
   netmask 255.255.255.0
   slaves eth2 eth3
   bond_mode 802.3ad
   bond_miimon 100
   bond_xmit_hash_policy layer2+3
```

2. Далее, перезапустите сервис:

```
   root@sterragate:~# service networking restart
```

3. Проверьте наличие интерфейса bond0 в системе:

```
   root@sterragate:~# ip address show | grep bond0
   4: eth2: mtu 1500 qdisc mq master bond0 state UP qlen
1000
   5: eth3: mtu 1500 qdisc mq master bond0 state UP qlen
1000
   10: bond0: mtu 1500 qdisc noqueue state UP inet
10.1.1.1/24 brd 10.1.1.255 scope global bond0
```

4. Отредактируйте ifaliaces.cf, удалив eth2 и eth3:

```
   root@sterragate:~# vim.tiny
/etc/ifaliases.cf
   interface (name="GigabitEthernet0/0"
pattern="eth0")
   interface (name="GigabitEthernet0/1"
pattern="eth1")
   interface (name="GigabitEthernet1/0"
pattern="bond0")
```

5. Затем пересчитайте контрольную сумму и перезапустите VPN-демона

```
root@sterragate:~# integr_mgr calc -f
/etc/ifaliases.cf
root@sterragate:~# service vpngate restart
```

Резервирование провайдеров GRE-over-IPsec

GRE over IPsec— технология для обеспечения отказоустойчивости оборудования и каналов связи.

Составные компоненты:

- GRE - протокол туннелирования сетевых пакетов, разработанный компанией Cisco Systems.
- OSPF – протокол динамической маршрутизации для обеспечения доступности площадок, отказоустойчивости и балансировки нагрузки
- IPsec – стек протоколов шифрования

Настройка GRE-over-IPsec

1. Настроим первый туннель:

```
root@GW1:~# ip tunnel add gre1 mode gre remote 10.0.2.2 local 10.0.1.2 ttl 255
root@GW1:~# ip link set gre1 up
root@GW1:~# ip link set gre1 mtu 1400
root@GW1:~# ip addr add 1.1.1.1/30 dev gre1
```

2. Настроим второй туннель:

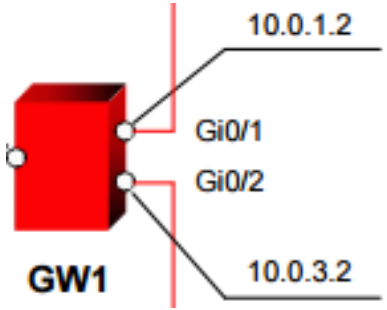
```
root@GW1:~# ip tunnel add gre2 mode gre remote 10.0.4.2 local 10.0.3.2 ttl 255
root@GW1:~# ip link set gre2 up
root@GW1:~# ip link set gre2 mtu 1400
root@GW1:~# ip addr add 2.2.2.1/30 dev gre2
```

3. Опишите трафик, который планируется защищать. Для этого создайте расширенный список доступа:

```
GW1(config)#ip access-list extended LIST
GW1(config-ext-nacl)#permit gre host 10.0.1.2 host 10.0.2.2 GW1(config-ext-nacl)#exit
GW1(config)#ip access-list extended LIST2
GW1(config-ext-nacl)#permit gre host 10.0.3.2 host 10.0.4.2 GW1(config-ext-nacl)#exit
```

4. Настройке OSPF в файле /etc/quagga/ospf.conf

```
hostname ospfd
password zebra
enable password zebra
router ospf
 network 192.168.1.0/24 area 0
 network 1.1.1.0/30 area 0
 network 2.2.2.0/30 area 0
 log file /var/log/quagga/ospfd.log
 log stdout
```



Object Tracking

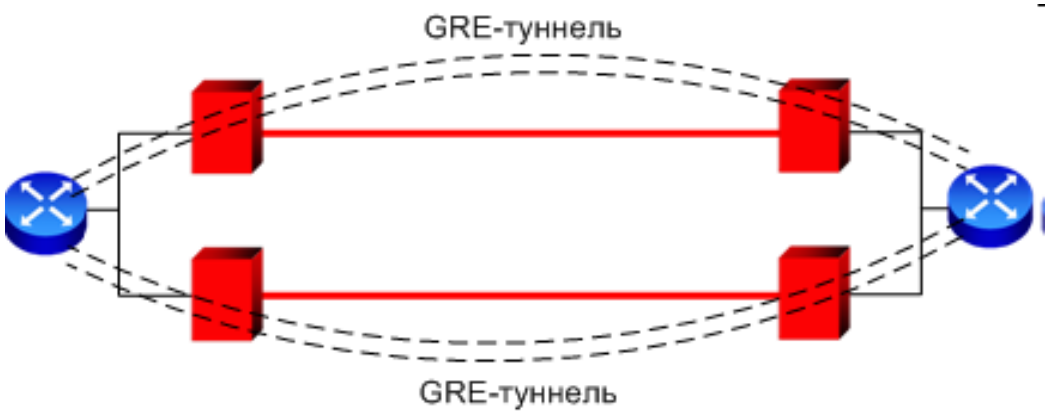
Object Tracking— механизм переключения каналов, основанный на периодической проверке доступности узла в сети.

Скрипт изменяет метрики основного и резервного маршрутов до узла.

Параметры скрипта /etc/rc.local.inc

```
#!/bin/bash
PING_IP=8.8.8.8                    #IP-арес для ICMP запросов
PING_COUNT=10                      #Кол-во пропущеных ICMP-запросов для переключения
PING_TIMEOUT=3                     #Интервал отправки ICMP-запросов для переключения
RESTORE_PING_COUNT=10              #Кол-во пропущеных ICMP-запросов для восстановления
RESTORE_TIMEOUT=1                  #Интервал отправки ICMP-запросов для восстановления
MAIN_INTERFACE=eth1                #Интерфейс, подключенный к основному провайдеру
BACKUP_INTERFACE=eth2              #Интерфейс, подключенный к резервному провайдеру
MAIN_GATEWAY=10.1.1.1              #Шлюз основого провайдера
BACKUP_GATEWAY=10.1.2.1            #Шлюз резервного провайдера
```

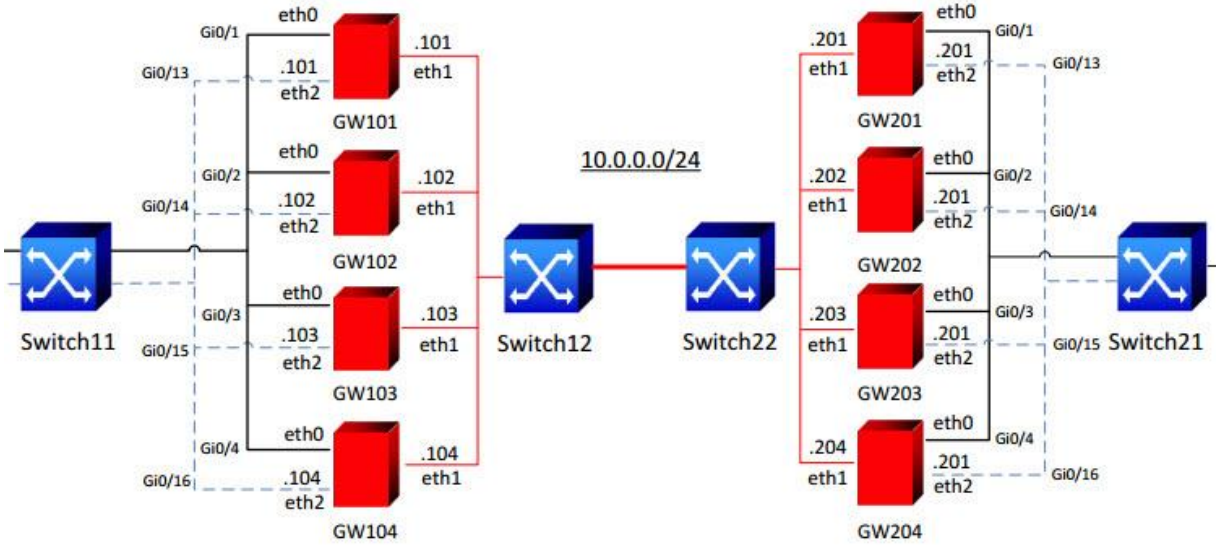
Масштабирование на L3, IPsec over GRE



Особенности решения

- GRE и OSPF/EIGRP-реализуется на маршрутизаторах
- Балансировка и отказоустойчивость на L3.
- Масштабируемость практически в любых объемах

Масштабирование на L2, Link-Aggregation



- Link-aggregation настраивается на коммутаторах
- На шлюзах используется пакет S-Terra L2
- Высокая масштабируемость решения

Резюме

s•terra

Решение	Основное назначение	Балан-сировка	Оптимальная топология	Время восстановления * (секунд)
VRRP	Резервирование шлюзов	Нет	Не важно	1-5
RRI	Резервирование шлюзов	Есть	Большое количество удаленных пользователей или филиалов.	7-15
Link Aggregation	Резервирование портов	Есть	Для локального сегмента	1-5
GRE over IPsec на шлюзе	Резервирование каналов	Есть	Site-to-Site или филиальная сеть.	7-15
Object Tracking	Резервирование каналов	Нет	Site-to-Site или филиальная сеть.	10-20
GRE over IPsec для балансировки нагрузки	Резервирование оборудования и каналов связи	Есть	Защита высокоскоростных каналов связи на уровне L3	7-15
Link Aggregation для балансировки нагрузки	Резервирование оборудования и каналов связи	Есть	Защита высокоскоростных каналов связи на уровне L2	1-5

* без учета времени перестроения туннелей. Время перестроения туннелей зависит от их количества и обычно лежит в диапазоне 0,5-1,5 минуты.

Шпаков Андрей

Ведущий инженер
Отдел технического консалтинга
Тел.: +7 (499) 940-90-61 (доб. 133)
Email: ashpakov@s-terra.ru, presale@s-terra.ru

s•terra®