



Networking
For everyone

BGP Fast Convergence



В этом разделе

- Next-Hop tracking
- BGP Prefix Independent Convergence
 - Edge
 - Core



BGP PIC

- BGP PIC = BGP Fast Convergence
- PIC Core
 - Уменьшает время сходимости сети при выходе из строя Core маршрутизатора
 - вернее сказать, при изменении выходного IGP интерфейса до PE
 - IGP должен найти новый путь до PE
- PIC Edge
 - Уменьшает время сходимости сети при выходе из строя PE маршрутизатора
 - BGP должен переключиться на новый PE



BGP конвергенция

- После установления сессии BGP входит в режим read-only на время, определяемое как *bgp update-delay*
- Определение того, пришли ли все префиксы от соседа:
 - keepalive
 - end-of-rib (если поддерживается)
- По выходу из read-only запускается процесс обчёта лучших путей
- Лучший путь считается для каждого префикса
- Обновление RIB влечёт за собой обновление FIB
 - обе эти операции могут быть довольно времязатратными



BGP конвергенция

- После обновления RIB, лучшие пути отправляются всем соседям
- Для оптимизация – dynamic update groups
- Обновления отправляются только по истечении advertisement interval для конкретного соседа
 - это механизм подавления для нестабильных соседей
 - по-умолчанию 30 сек



BGP и IGP

- Адрес Next-Hop чаще всего известен посредством IGP
- Доступность NH проверяется процессом **BGP scanner** на периодической основе
 - по-умолчанию 60 сек
- К сожалению, возможен очень медленный вариант поиска недоступного NH
- Решение – **Next-Hop Tracking (NHT)**



Next-Hop Tracking

- Процесс **RIB watcher** связывает RIB и BGP
- Обычно кол-во NH = кол-ву PE (не самое огромное число)
- Чаще всего случаются два IGP события:
 - IGP префикс становится недоступен
 - изменяется IGP метрика
- IGP передаёт изменения в сторону BGP через таймер **bgp nexthop trigger delay**
- В качестве реакции IGP NH, BGP пересчитывает лучшие пути
 - затрагивает каждый IP префикс, для которого изменился NH
 - время перестроения не предсказуемо
 - упал NH = пересчитали Full View 😊



Плоский FIB

- В случае организации плоского FIB, даже если с т.з. IGP не поменялся NH, а только выходной интерфейс – BGP будет конвергировать!

```
PE1#  
RT: updating ospf 6.6.6.6/32 (0x0) :  
    via 192.168.25.5 Gi0/2  0 1048578  
  
RT: closer admin distance for 6.6.6.6, flushing 1 routes  
RT: add 6.6.6.6/32 via 192.168.25.5, ospf metric [110/4]  
RT: updating bgp 8.8.8.8/32 (0x0) :  
    via 6.6.6.6  0 1048577
```



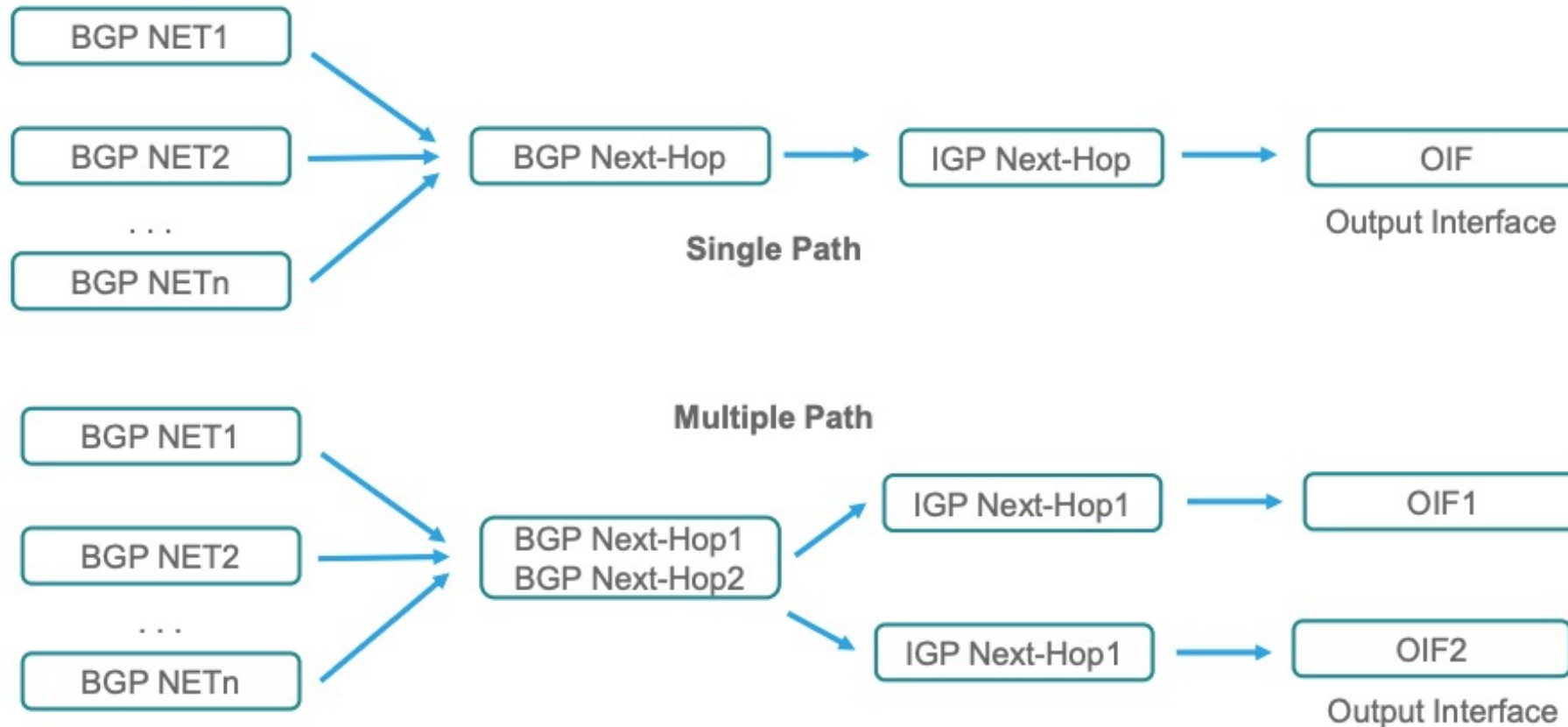
- Решение – иерархический FIB



Иерархический FIB

- Используется в IOS-XR и NX-OS

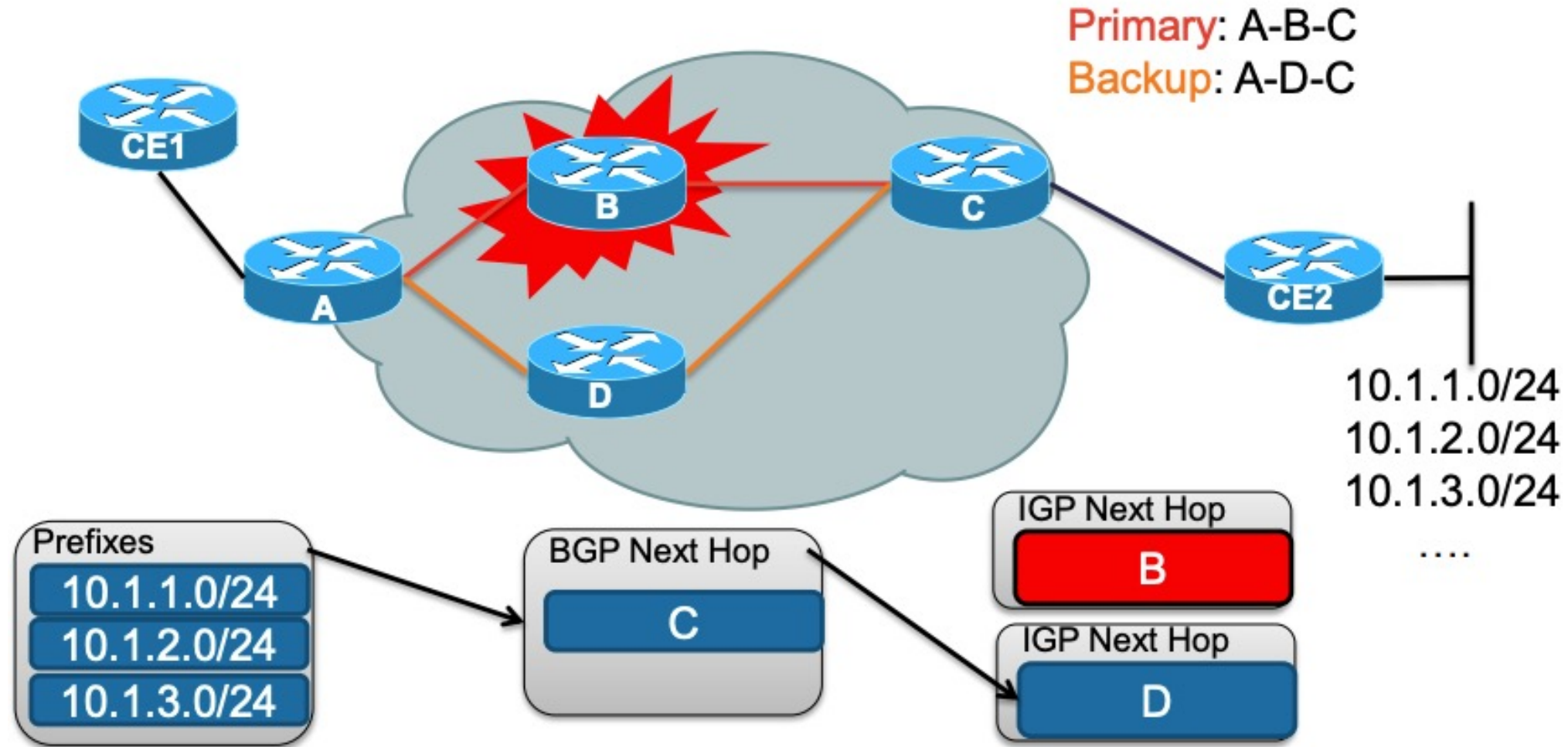
```
(config)#cef table output-chain build favor convergence-speed
```



BGP PIC CORE



Networking
For everyone





Идея PIC Edge

- Основная идея – чтобы BGP знал о резервном пути
- PE и RR должны договориться о передаче резервных путей
 - функционал additional path
- Резервный путь не только должен быть в BGP таблице, но и установлен в FIB

```
ASR-1K# sh ip bgp vpnv4 vrf site-111111 2.0.0.0
BGP routing table entry for 65300:111111:2.0.0.0/24, version 150035
Paths: (3 available, best #1, table sie-111111)
Additional-path-install
<snip>
10.200.1.2 from 10.200.1.2 (10.200.1.2)
Origin incomplete, localpref 100, valid, external, best
Extended Community: RT:64300:111111 , recursive-via-connected
rx pathid: 0, tx pathid: 0x0
<snip>
Refresh Epoch 1
20570 20570
10.10.10.2 from 10.10.10.2 (5.5.5.5)
Origin incomplete, localpref 100, valid, internal, backup/repair
Extended Community: RT:64300:111111 , recursive-via-host
rx pathid: 0, tx pathid: 0
```



Networking
For everyone