

バックボーン刷新で見た BGP PICの落とし穴

～NextHopとStatic経路が高速切替を止めた話～



合同会社DMM.com
ITインフラ本部インフラ部ネットワークグループ NREチーム
大橋 幸輝
2026/01/20

CONFIDENTIAL

自己紹介

名前(年齢): 大橋 幸輝(25)

趣味:

- 自宅でのサーバ・ネットワーク運用とアプリ・システム開発
- 電気工事
- アマチュア無線
- フルルートを眺めること

経歴:

- 2022/8 ~ 23新卒としてDMM.comに入社(3月末までアルバイト)

業務範囲:

- NW設計/検証/構築/運用(バックボーン~拠点~サーバサイド)
- 自動化/ツール開発/クラウド・オンプレ自動化基盤運用
(Ansible/Python/Go/Container/k8s/Proxmox/DB/Terraform/MSA)

JANOG歴

- JANOG53 Hakataで初参加し継続中
- JANOG55 登壇



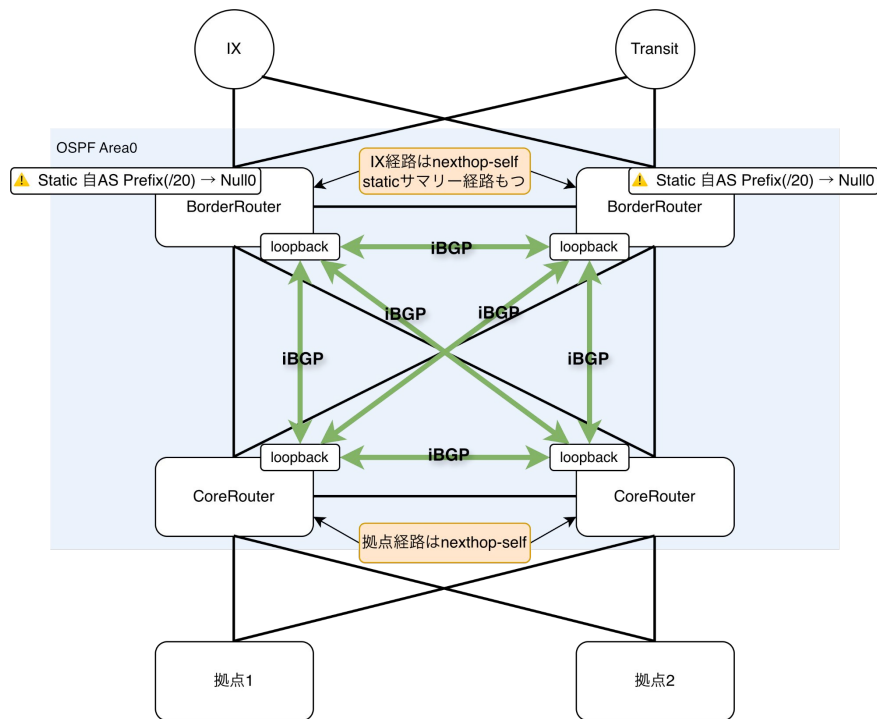
Agenda

1. 背景
2. 想定していたPICの動作(理論上)
3. 起きたこと(現実)
4. 原因
5. 解決策と教訓
6. 疑問

目的と背景 ~なぜPICを入れたのか~

- 機材のEOLに伴うバックボーンのフルリプレイスを実施
 - 構成・機材を見直し
 - 新規構成で一から障害試験を実施
- バックボーン要件
 - 可用性
 - 障害時は瞬断レベル (1秒程度)で切替
 - BGP PICを採用
 - その他
 - フルルート以外はNextHop-Selfする

ネットワーク構成



● 構成の要点

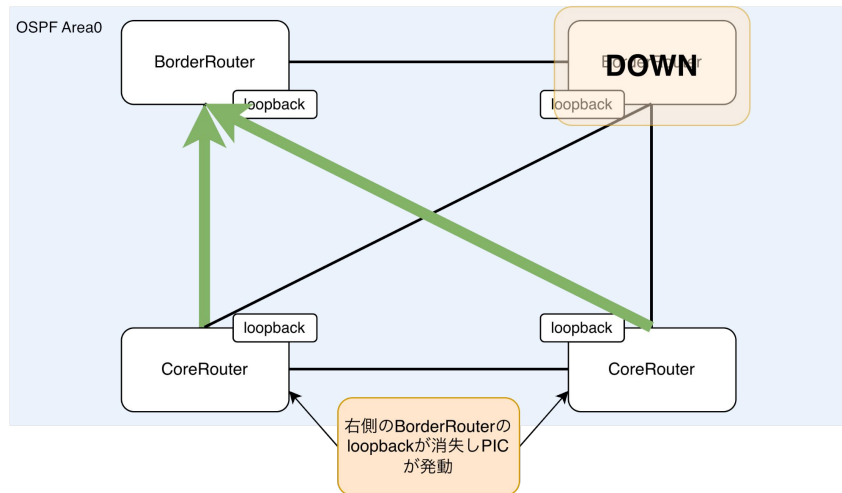
- OSPFをIGPとして利用
- LoopbackはOSPFで解決
- Loopbackに設定するアドレスは自ASのグローバルIPアドレス
- iBGPはLoopback同士でフルメッシュ

● 重要設計

- フルルート以外はNextHop-Self
 - BorderRouter / CoreRouter
- BorderRouterにはIP広報のためのStaticサマリー経路が設定

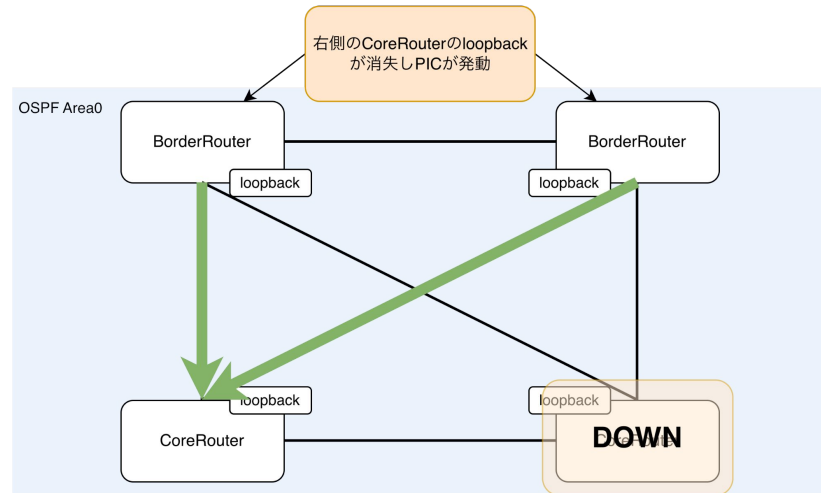
想定していたBGP PICの動作(理論上)

✓ ケース1: BorderRouterで筐体障害



CoreRouterのFIB/RIBからBorderRouterのLoopback IPが消失しBGP PICが発動

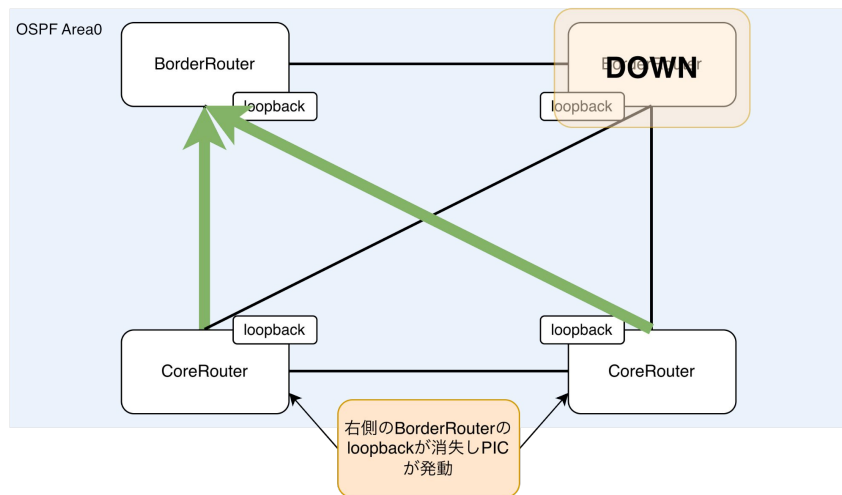
✓ ケース2: CoreRouterで筐体障害



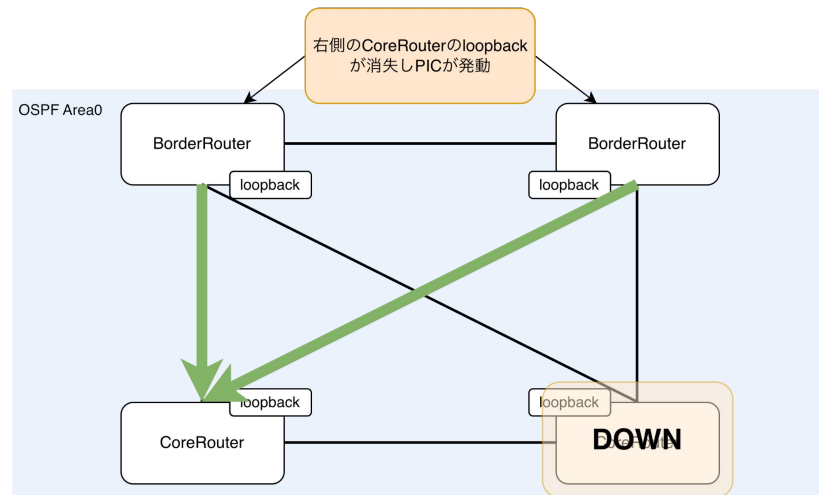
BorderRouterのFIB/RIBからCoreRouterのLoopback IPが消失しBGP PICが発動

想定していたBGP PICの動作(理論上)

✓ ケース1: BorderRouterで筐体障害



✓ ケース2: CoreRouterで筐体障害



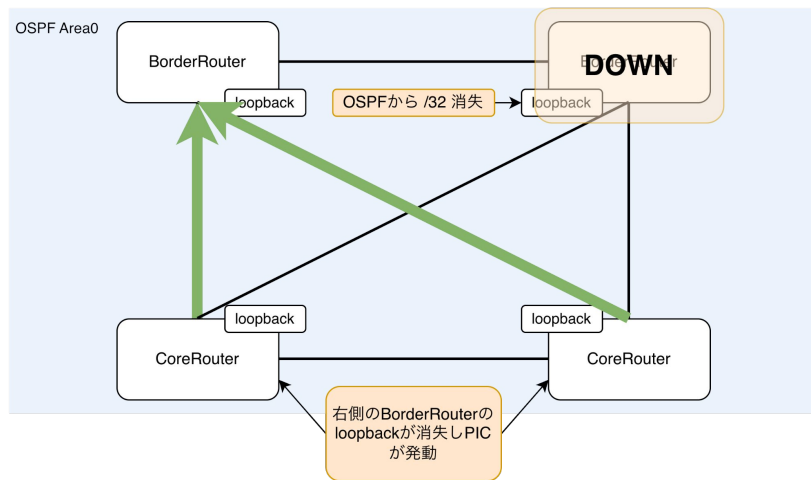
CoreRouterの
Loopback IPが

完璧な設計だ！と思っていた時期が私にもありました

CoreRouterの
Loopback IPが消失しPICが発動

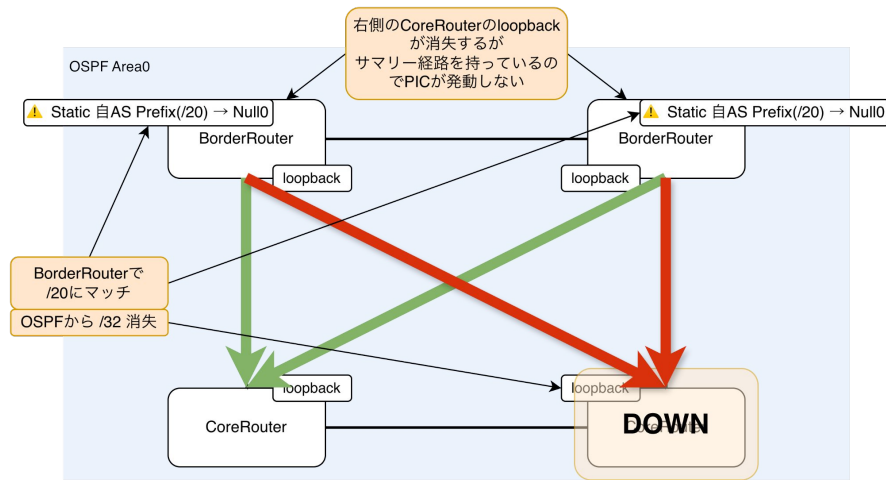
起きたこと(現実)

✓ ケース1: BorderRouterで筐体障害



CoreRouterのFIB/RIBからBorderRouterのLoopback IPが消失しBGP PICが発動

✗ ケース2: CoreRouterで筐体障害



OSPFからCoreRouterのLoopback IPが消失するが、Staticサマリー経路を持っているためBGP PICが発動しない

原因 ~なぜPICは動かなかったのか~

起きていたこと

- CoreRouterがDown
- OSPFからLoopback IP(/32)が消失

ルータの判断

- NextHopはルートテーブル上に存在
- 到達性あり

BorderRouterのFIB/RIBでは

1. /32が消えた
2. 次にマッチする経路を探索
3. Staticのサマリー経路(/20)にヒット

結果

- PICのトリガーである、「nexthop down」にならない
- 条件を満たしていないのでPICが動作しない

解決策と結果

NextHopを本当に消すために

- Loopback設計の変更
 - バックボーン内でStaticで定義していないPrivate IPへ
- iBGP / NextHop-selfもPrivate IPに統一
- IPv6ネイバーも同様に変更

結果

1. OSPFからLoopbackが消失
2. すぐにnexthop down
3. BGP PICが発動
4. iBGPネイバーもすぐにdown

1秒以内の切替を達成！

教訓

教訓

BGP PICを使うときは
NextHopが他のStaticや集約経路に
含まれていないか必ず確認する

障害試験は大事

他に試したこと

- iBGPネイバーにBFDを入れる
 - ネイバーは即落ちるが
 - リンクフラップ時に頻繁に BGPがフラップ
 - **PICの代替にはならない**
- Loopback範囲だけ広報しない
 - バックボーン以外にも利用しているので難しい
 - 乗っ取りが怖い
- Aggregate でサマリー経路を生成
 - Static ではなく Aggregation で /20 を生成
 - FIB上では /20 → Null0 に変わりはない
 - **PICの条件は改善されなかった**

今も残っている疑問

① 機器固有の機能を使うべきか？

- リプレイス前機器には「RIBに入れない」機能があった

② PICを利用する際は Loopback を専用アドレス帯にすべきか？

- 明確なルールは無い認識
- なにか特殊な設定があるのか
- ベストプラクティスは何なのか

PIC設計の知見をお持ちの方、後ほどお声がけいただけると嬉しいです