



MPLS-TE 運用の今

- トランジットプロバイダの例 -

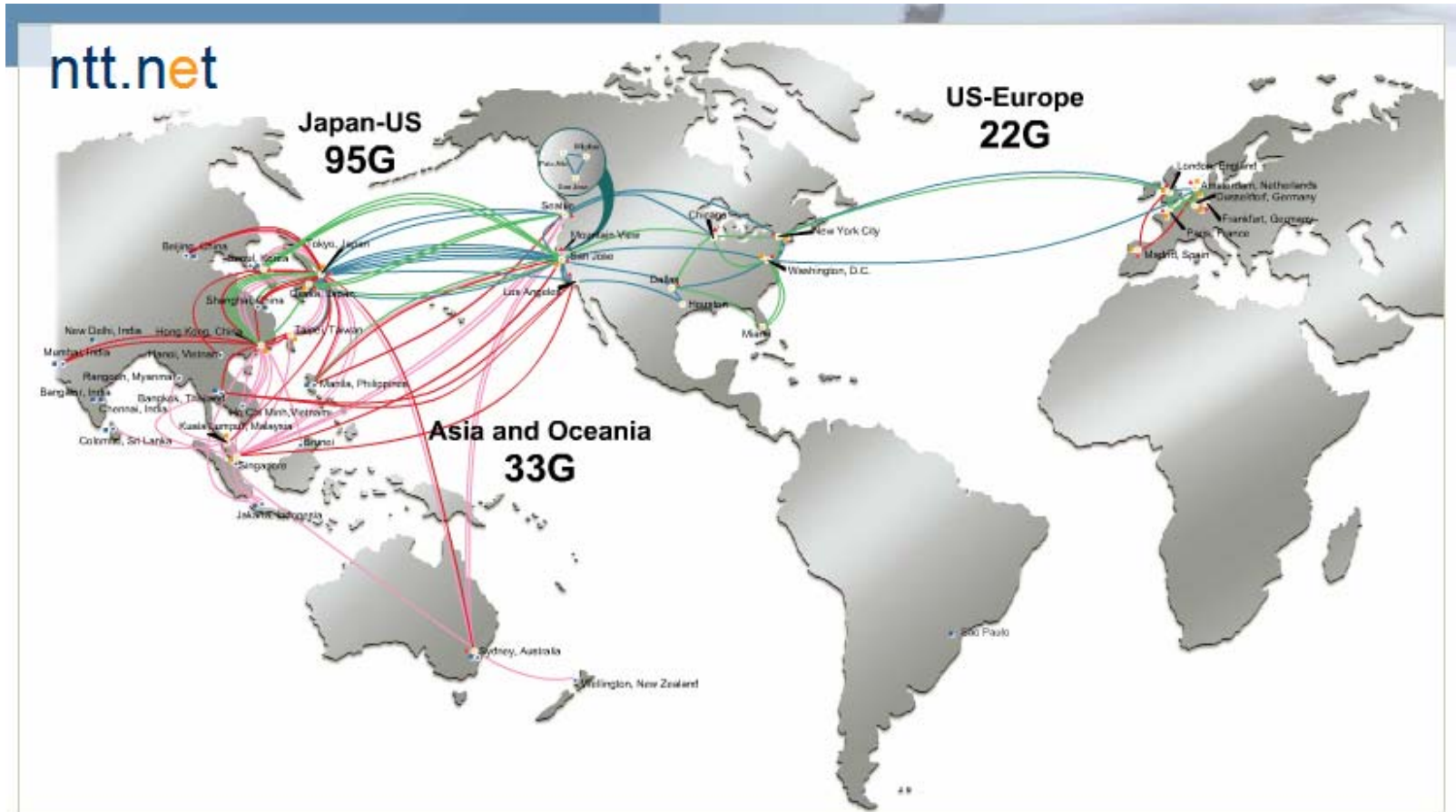
有賀 征爾 <say@ntt.net>
NTTコミュニケーションズ



自己紹介

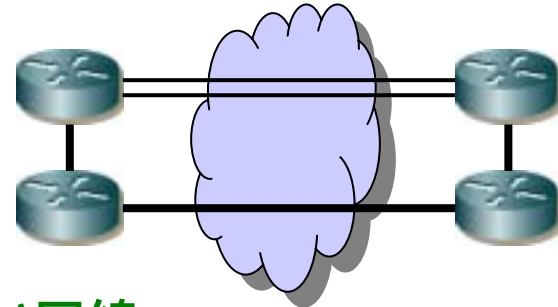
- AS2914
 - NTT Communications Global IP Network
 - “ntt.net”
 - OCN(AS4713)の上流です
 - 昔は Verio (とか“NTT/VERIO”)でした
 - 今は “Verio” はホスティングサービスだけ
- 私
 - 2002年くらいから運用しています

ネットワーク



AS2914 と MPLS

- 2002年初夏くらいから
- 目的
 - トラフィックの統計 (POP-to-POP)
 - たとえば、あるアジアのPOPとUSの通信量は？
 - Traffic Engineering
 - 国際線が高い
 - 同時には回線を増やせない
 - » 複数の細かい回線と少数の太い回線
 - » 典型的な非対称の回線



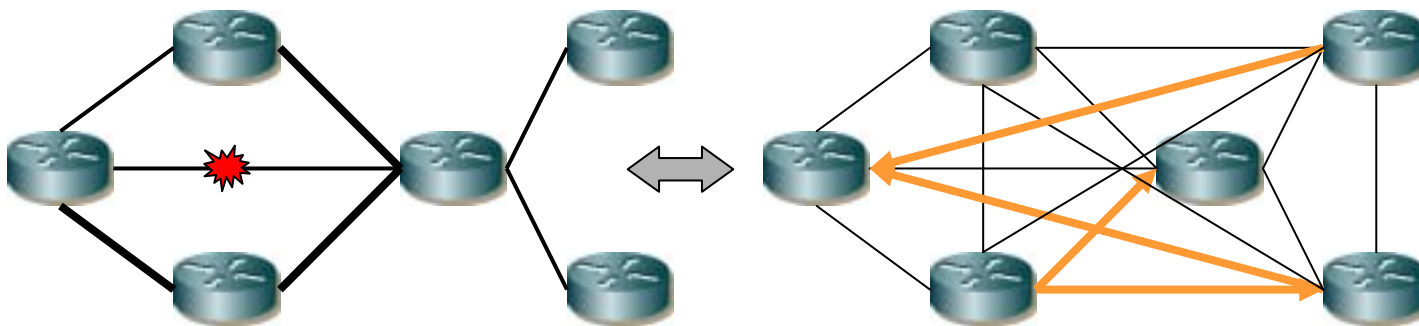


基本設計 (1)

- 「なるべくシンプルに」
- LSP
 - core ルータ間のフルメッシュ
 - 逆に言うと、MPLSするルータとしないルータの判断が面倒
 - 全世界に広がる100台程度のルータ
 - $100 \times 99 = 9900$ 本
 - 内製のツールを使用
 - 一人のエンジニアが全世界のルータの設定を数分で変更できる
 - 「試してみる」ことができる
 - RSVP (というよりIGP)まかせの経路
 - Explicit に経路を決めるのは無理
 - ただし、トラブル時などに局所的に利用することはあり

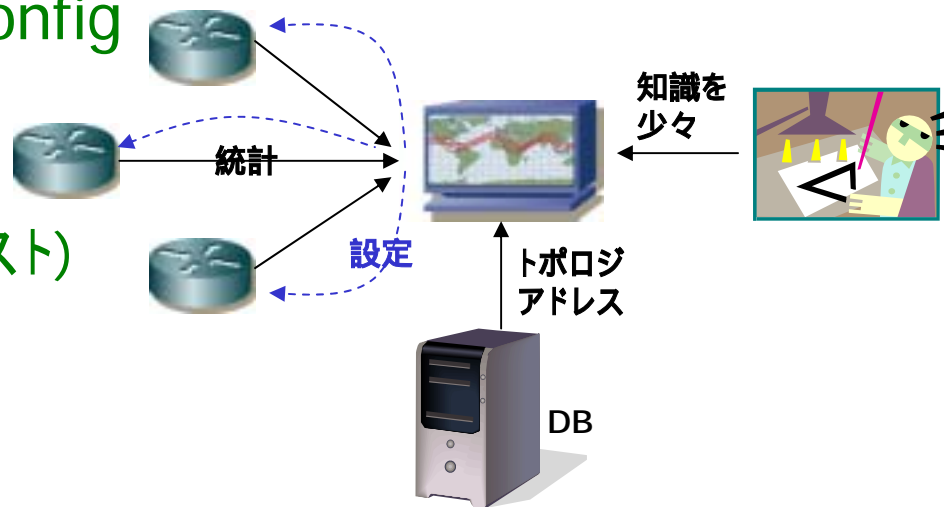
基本設計 (2)

- 統計
 - 統計(実際のトラフィック)にもとづき各LSPの帯域を決める
 - ルータ ルータ間のトラフィック
 - ほぼPOP POP間のトラフィックと同義
 - 適当な時間間隔でLSPのサイズを変更
 - LSPのサイズとトラフィックが合わないとおふれる
 - core ルータ間なので急激にトラフィックトレンドは変わらないだろうという前提(割り切り)
- その他
 - FRR とかは使わない
 - LSP単位での管理はできない



ツール

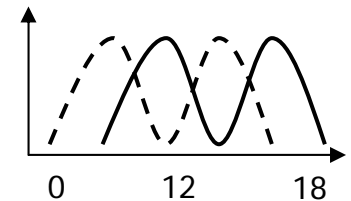
- すべての設定に使ってる
 - LSPの設定だけではなく
- 必要な情報を集めて config に落とす
 - トラフィック統計
 - アドレス (ルータのリスト)
 - IGPのメトリック
- 複雑な条件を入れない
 - 特別なLSPを作らない
 - 特定のルータ、経路を特別視しない



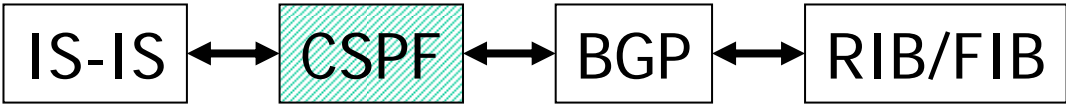
課題 (1)

- LSPのサイズ

- 結局、これが問題 ((3)も同じ問題)
- わりと実トラフィックと設定したLSPのサイズが合わないことが多い
 - 特にまだスタートアップのPOP
- 原因
 - 思ったほどトラフィックのトレンドが安定しない
 - 規模が小さいPOPだと一顧客のトラフィックの変動の影響をもろにうける
 - p95 vs avg
 - 各国のトラフィックのピークにずれがある
 - » 回線の予約はLSPの予約の足し算
 - 統計の平均かp95か？

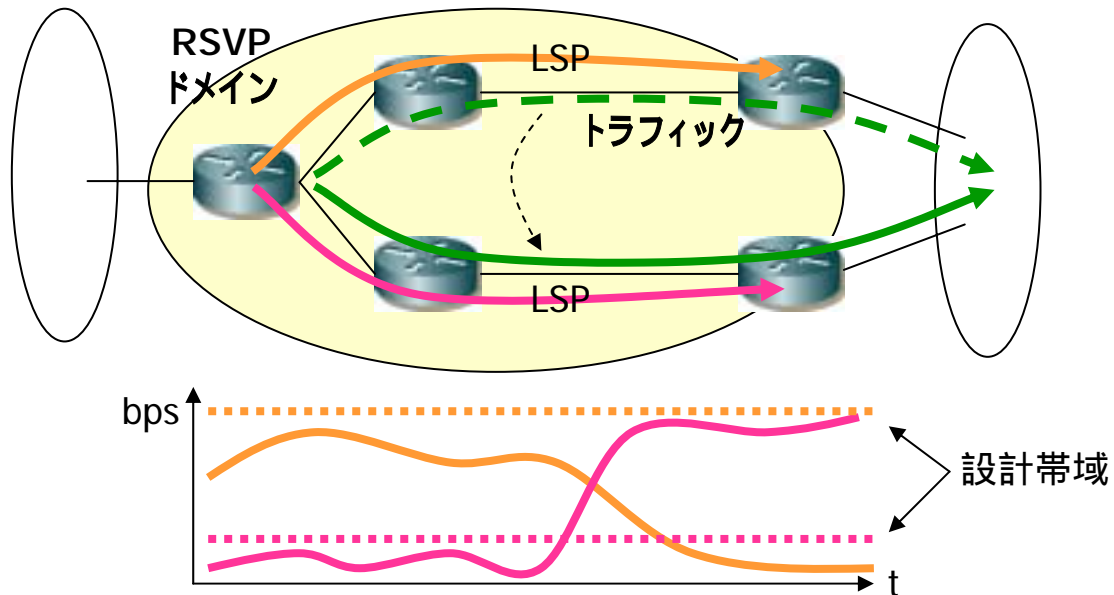


課題 (2)

- パフォーマンス (コンバージェンス)
 - 「state」をもつプロトコルが増える
 - RSVP、LDP、IS-IS、BGP、PIM ...
 - ルータのパフォーマンスに影響
 - コンバージェンスに影響 (forwarding には関係ない)
 - RIB を作るにもステップが増える
- 
- ```
graph LR; A[IS-IS] <--> B[CSPF]; B <--> C[BGP]; C <--> D[RIB/FIB]; style B fill:#add8e6,stroke:#333,stroke-width:1px
```
- 過渡的な問題
    - ルータ/ファームウェアのバージョンアップ

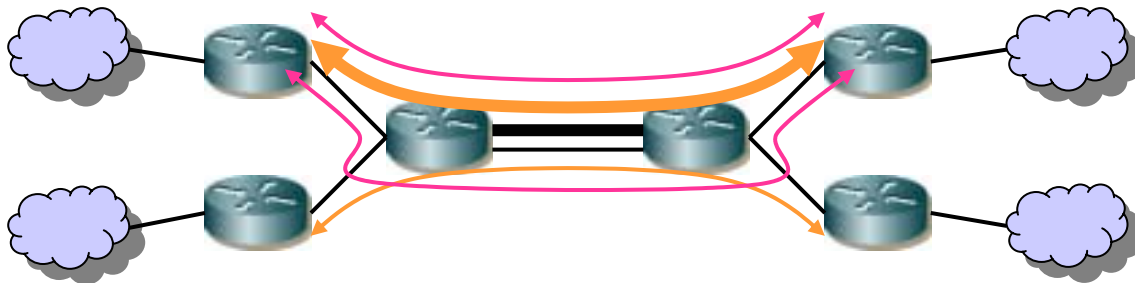
# 課題 (3)

- トポロジが変わると困る
  - IS-IS のメトリック/コスト設計
    - 新規回線や故障はここでは含めません
  - LSPのサイズがまったく実トラフィックに合わなくなる



# 課題 (4)

- LSPが大きすぎる
  - あるPOPとあるPOPだけが大きくなった場合
  - 一つのLSPの動きが回線帯域に影響
  - 対処
    - 大きいLSPを分割する
    - とはいえ、
      - 分割したLSPを別に管理し続ける必要がある
      - LSP の ECMP に問題が起こったこともあった





# 今後

---

- 帯域が大きくなってきた
  - 小さい帯域も bundling で束れば大きくなる
- 一方、帯域がなければ結局限界がある
  - 結局帯域がなければ Engineering できない
- どこまで TE が必要でしょうか？
  - たぶん繰り返しかとは思いますが...