ネットワークユーザを支援する計測技術
- ネットワーク見える化- にむけて

小林 克志（産総研）

自己紹介

スケーラブルな
ネットワークアーキテクチャ

IETF 標準化
RFC3189, 3190

産業技術総合研究所

2006

ネットワークの
見える化

マルチメディアシステム

情報通信研究機構
(通信総合研究所)

1998

Gbps クラスの広帯域通信を容易
に実現するアーキテクチャ

電気通信大学
情報処理センター

1993

IP over ATM のホスト実装

インターネット計測の研究開発

- 計測データをサーバで管理
  - 受動、能動計測のデータを収集、提供
    - ネットワーク計測システム e.g., PerfSONAR
- 端末間計測
  - パケットの振る舞いから内部を「推定」
  - 損失、遅延の変化、パケット順序、パケット間隔時間....
    - 計測アプリケーション e.g., ping, traceroute, pchar, iperf
    - トランスポートスタック TCPs
- Knowledge Plane = “sensor” + “actuator”
  - ネットワーク自身が自動構築、再構成、問題発見、解決、報告
    - 本プロジェクトは “sensor” の実現を目指す

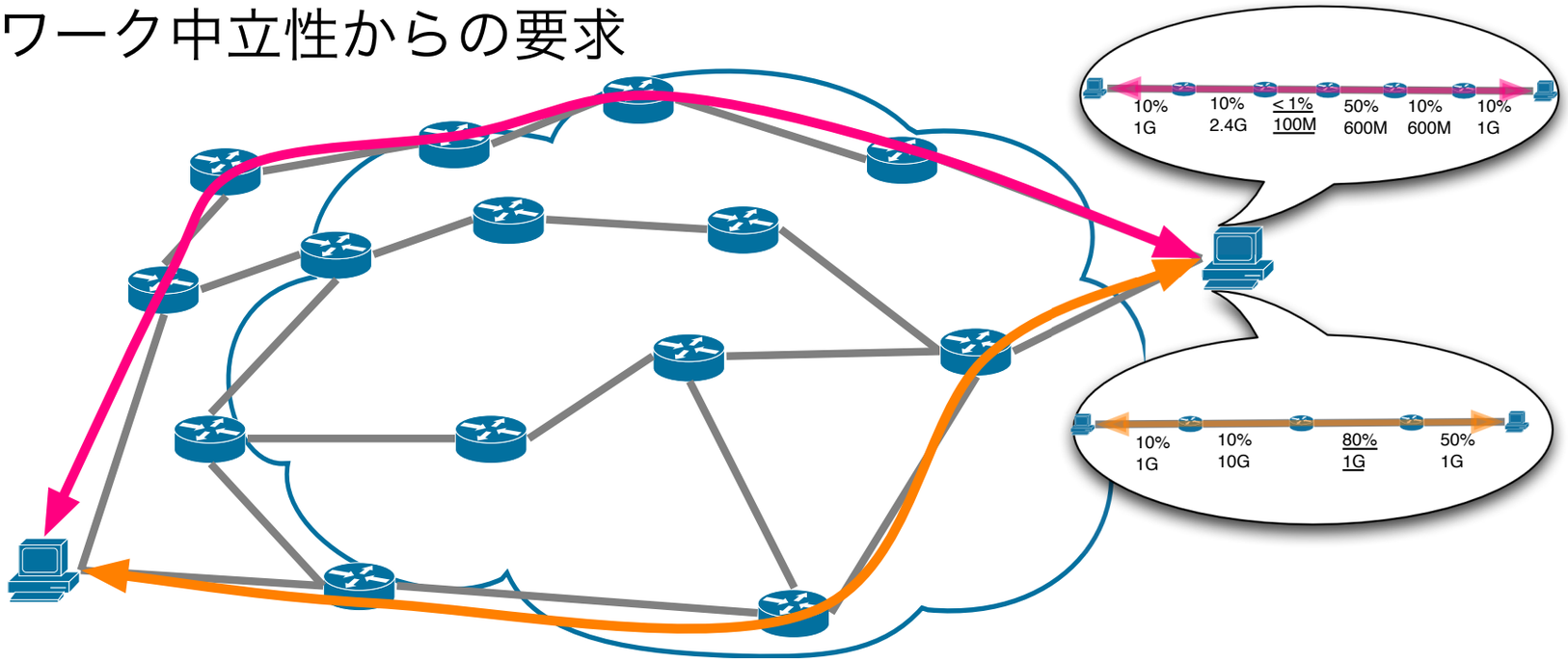
インターネットの『見える化（可視化）』とは

➡基盤の多様化で『推定』では不十分

- 帯域幅 9.6Kbps - 10Gbps、ハンドオーバー、輻輳以外が原因の損失・遅延
- P2P・CDN のピア・サーバ選択、可用帯域に最適化された符号化選択などネットワークの状況に応じた効率的な通信

➡情報開示の方向へ

- 提供サービス品質の検証、通信障害時の状況把握
- ネットワーク中立性からの要求



“The sixth principle is a transparency principle -- stating that providers of broadband Internet access must be transparent about their network management practices.”, http://www.openinternet.gov/read-speech.html, J.Genachowski, 2009

『可視化』の要件

- スケーラビリティ
  - 端末数、ルータ数、フロー数、複数ドメイン....
- 遅れのない情報取得
- 単純で正確な処理
- 開示範囲の制御
  - パス利用以外の端末からのアクセス制限
  - 通過 ISP, 利用者の意向の尊重

端末自身の SNMP アクセスによる『可視化』

1.端末からの直接 SNMP アクセス開放

- 管理プレーンの介在
- 高頻度のアクセス
- 通信パスの決定にはルータ経路表からパスを再構成
- 誰でも（パスに関係ない端末から）ルータにアクセス可

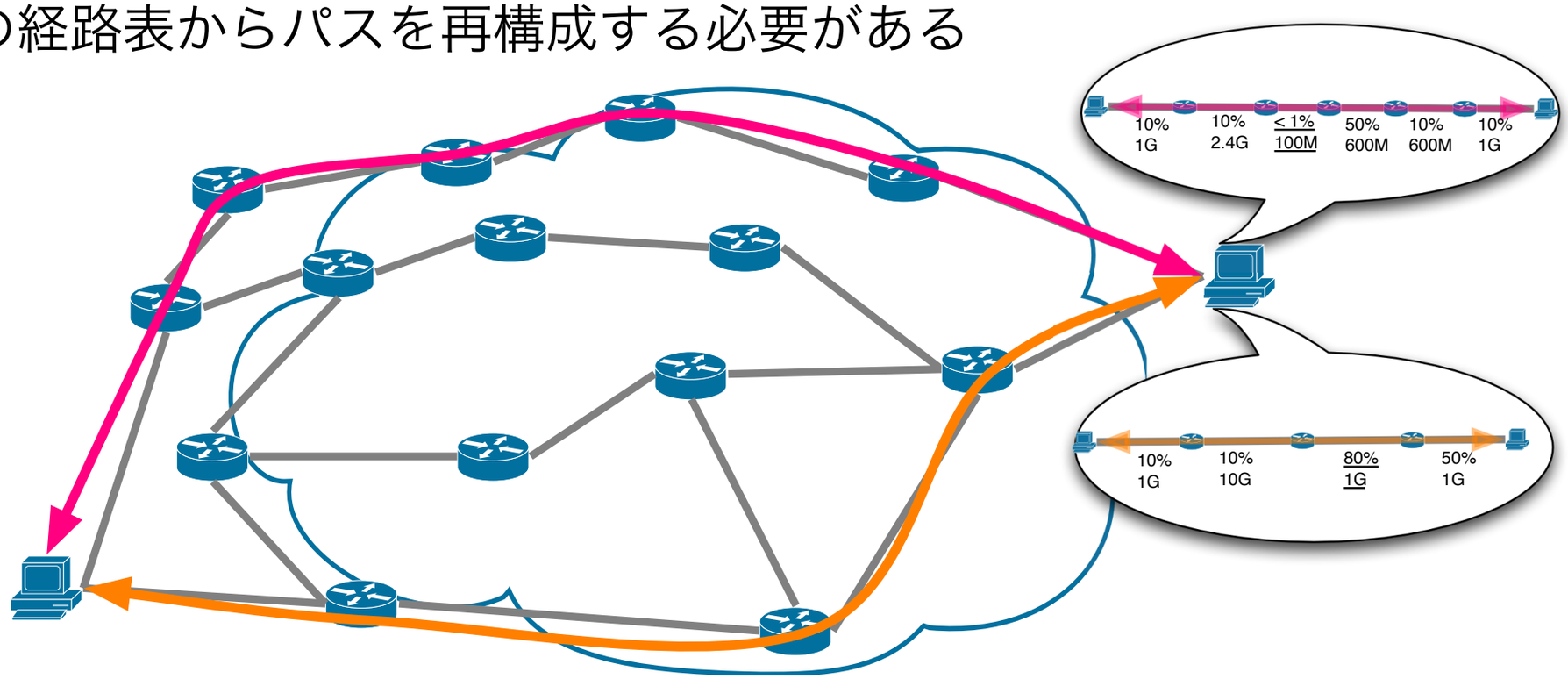
2.Proxy, NMS アクセス

- 端末／サーバ数の適正化、管理ドメインごとに異なるサーバ
- 端末 - サーバ, サーバ - ルータへの高頻度のアクセス
- 通信パスの決定は、サーバ？端末？
- 誰でも（パスに関係ない端末から）情報にアクセス可能
  - ISP vs. P2P サービス提供者の対立 -> draft-kiesel-alto-h12-00.txt

- 運用向け『可視化』データを開放するという手法には限界

「端末」に必要な情報とは？

- 相手までの通信経路に関する情報
  - 輻輳状態,ボトルネック帯域,可用帯域,遅延,corruption loss
  - 複数パスが選択できる場合は、個々のパスの情報を利用
- 一般「端末」にはネットワーク全体の情報は冗長
  - ソースルーティングは無効、端末からの径路指定は不可能
  - 途中ルータの経路表からパスを再構成する必要がある

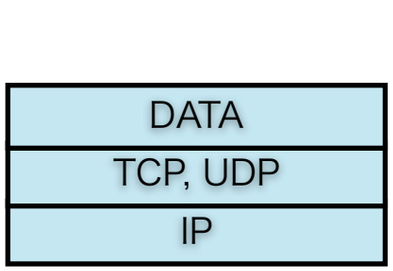


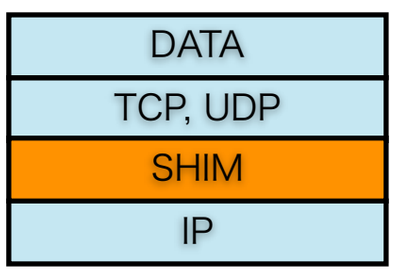
in-band cross-layer approaches - for enhancing transport-

- ルータは多くの情報を管理している
  - Network is dumb, router is also smart, not only terminals.
- Jack up with shim layer between IP and TCP/UDP
  - ETEN, PTP, SIRENS

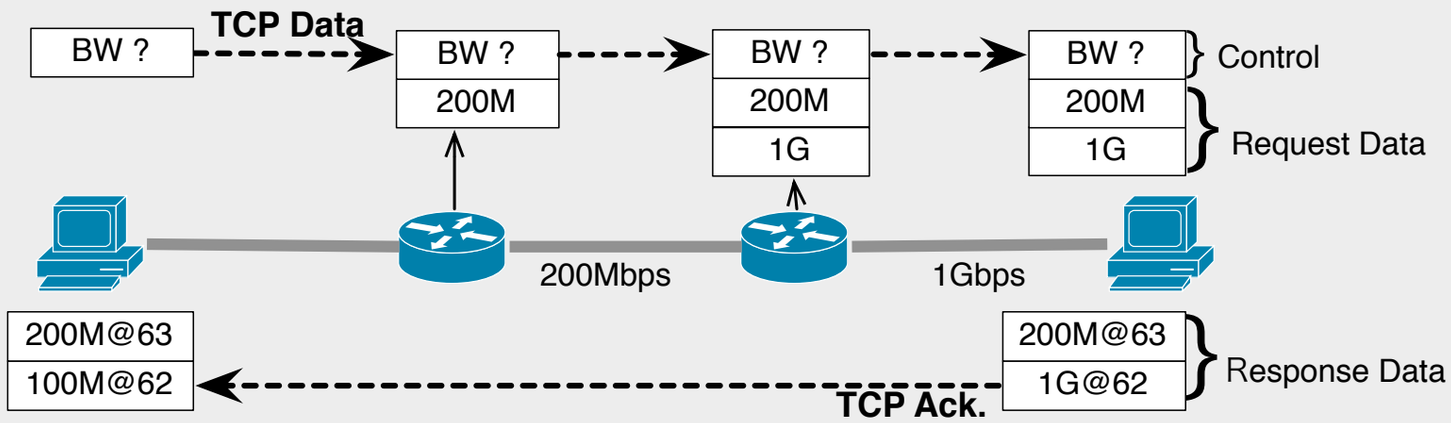
  ➡ ルータが SHIM 層に書き込むことで情報を提供
  - XCP, TCP-QS, ...

  ➡＋シグナリングとリソース（空き帯域）管理機構
- パケット毎の書き込み処理は？
  - 今日のルータはすべてのパケットに大して TTL、L2 アドレスの書き込み処理を要求されている

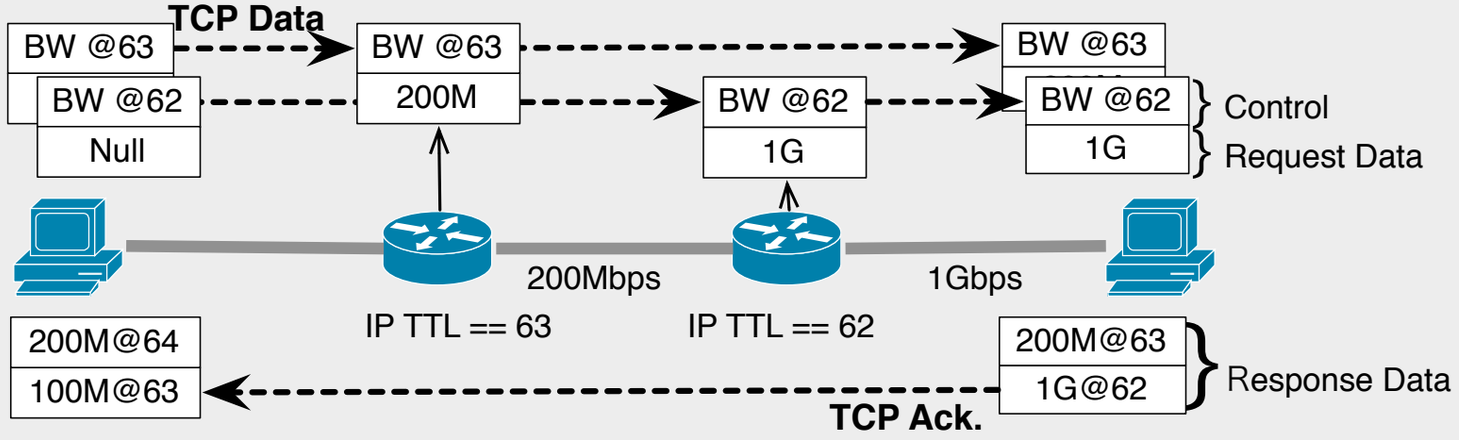




in-band cross-layer approaches



PTP: retrieve every hop's data by single packet with header growth



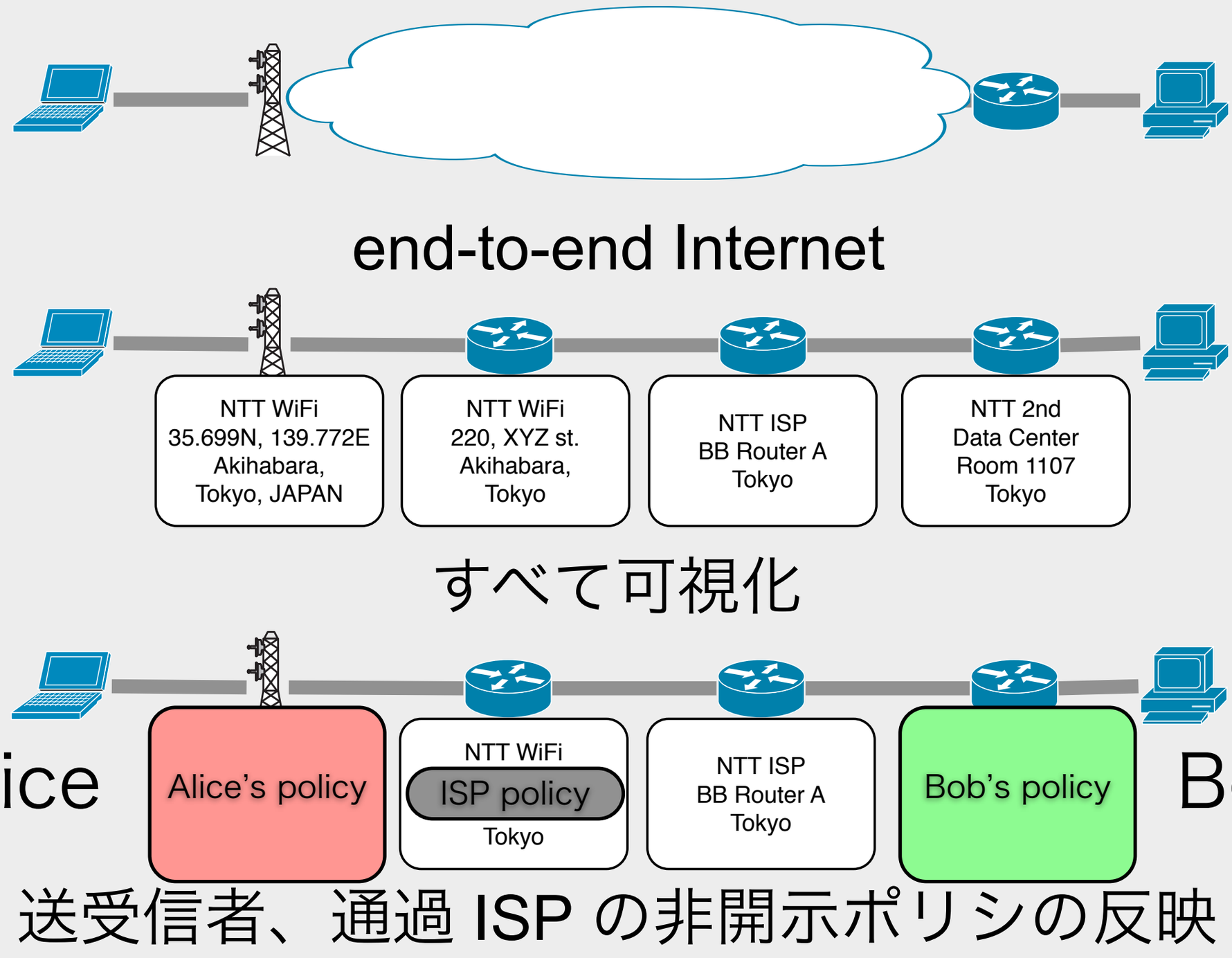
SIRENS: one packet collects one router's data, multiple packets are required to retrieve whole path

送受信側の協業が必須、送受信者間のルータにのみアクセス可

『可視化』の要件

- スケーラビリティ
  - 端末数、ルータ数、フロー数、複数ドメイン....
- 遅れのない情報取得
- 単純で正確な処理
- 開示範囲の制御
  - パス利用以外の端末からのアクセス制限
  - 通過 ISP, 利用者の意向の尊重

通過 ISP, 利用者の意向の尊重

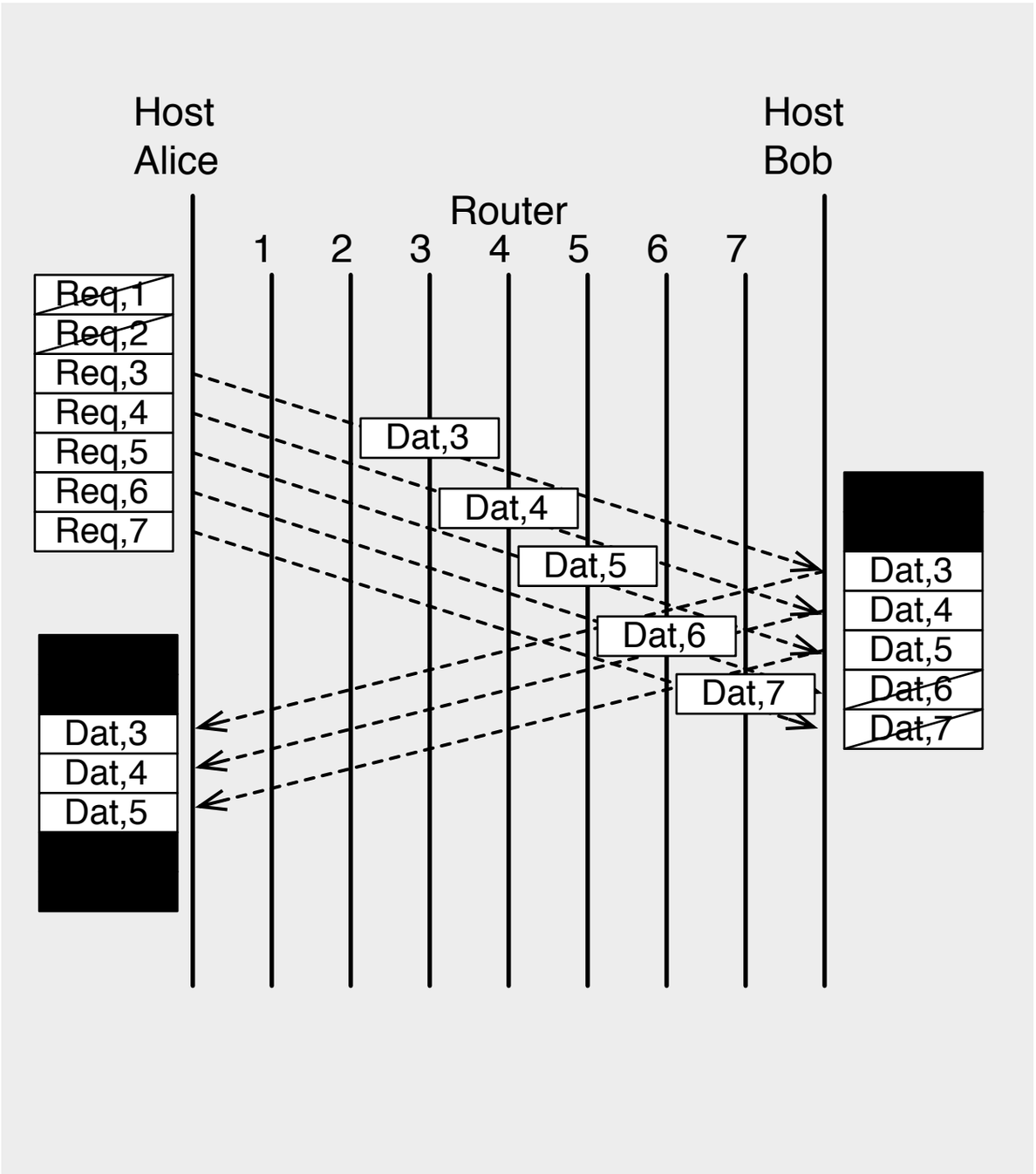


開示ポリシ尊重の実現方法

- 基本：ISP と送受信者が開示に合意した情報だけ開示
- ISP の開示ポリシ：
  - ルータ側で ACL を設定
    - ルータ単独で動作、管理ドメイン毎に独立したポリシ
- 送受信者のポリシ：
  - 送受信者が開示可能な範囲を指定
    - 選択的情報開示

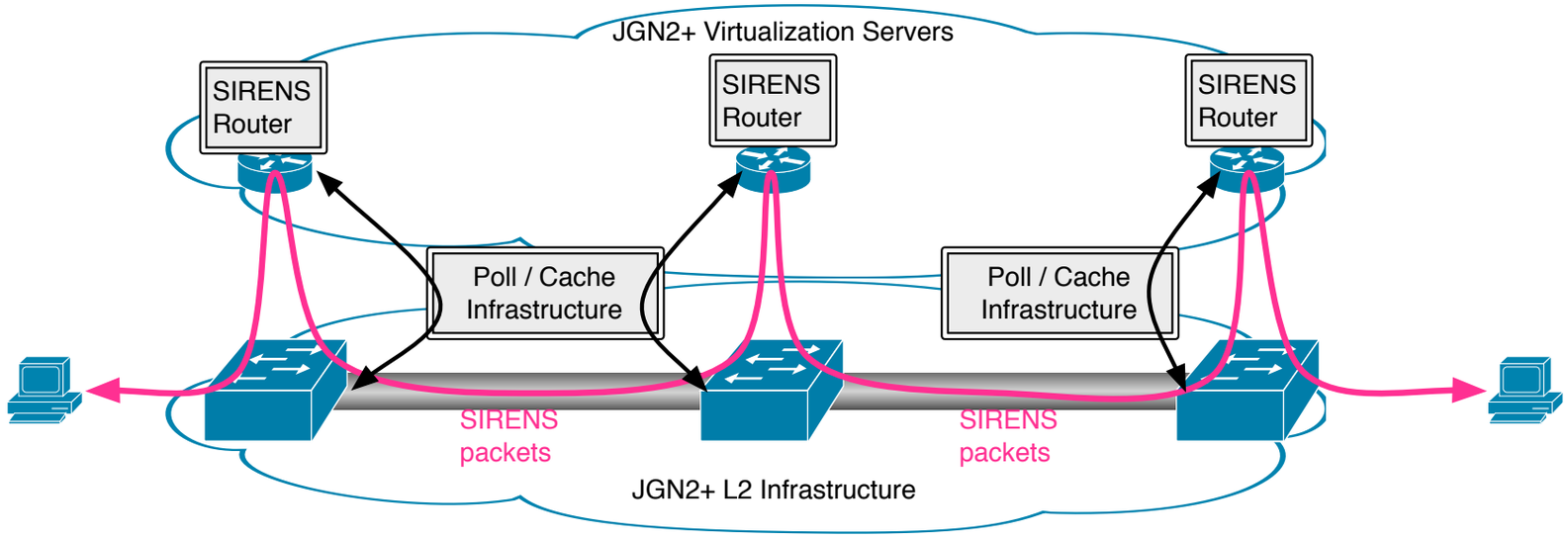
選択的情報開示

- Policy:
  - Alice & Bob allow to disclose beyond 3rd hop routers.
- Implementation:
  - Alice does not send req. for neighbor & next neighbor routers, i.e.,1st & 2nd hop.
  - Bob does not send back res. as Alice, i.e., 6th & 7th hop.
- Result:
  - Alice obtains 3-5 hops' data.
  - Bob obtains 3-7 hops' data



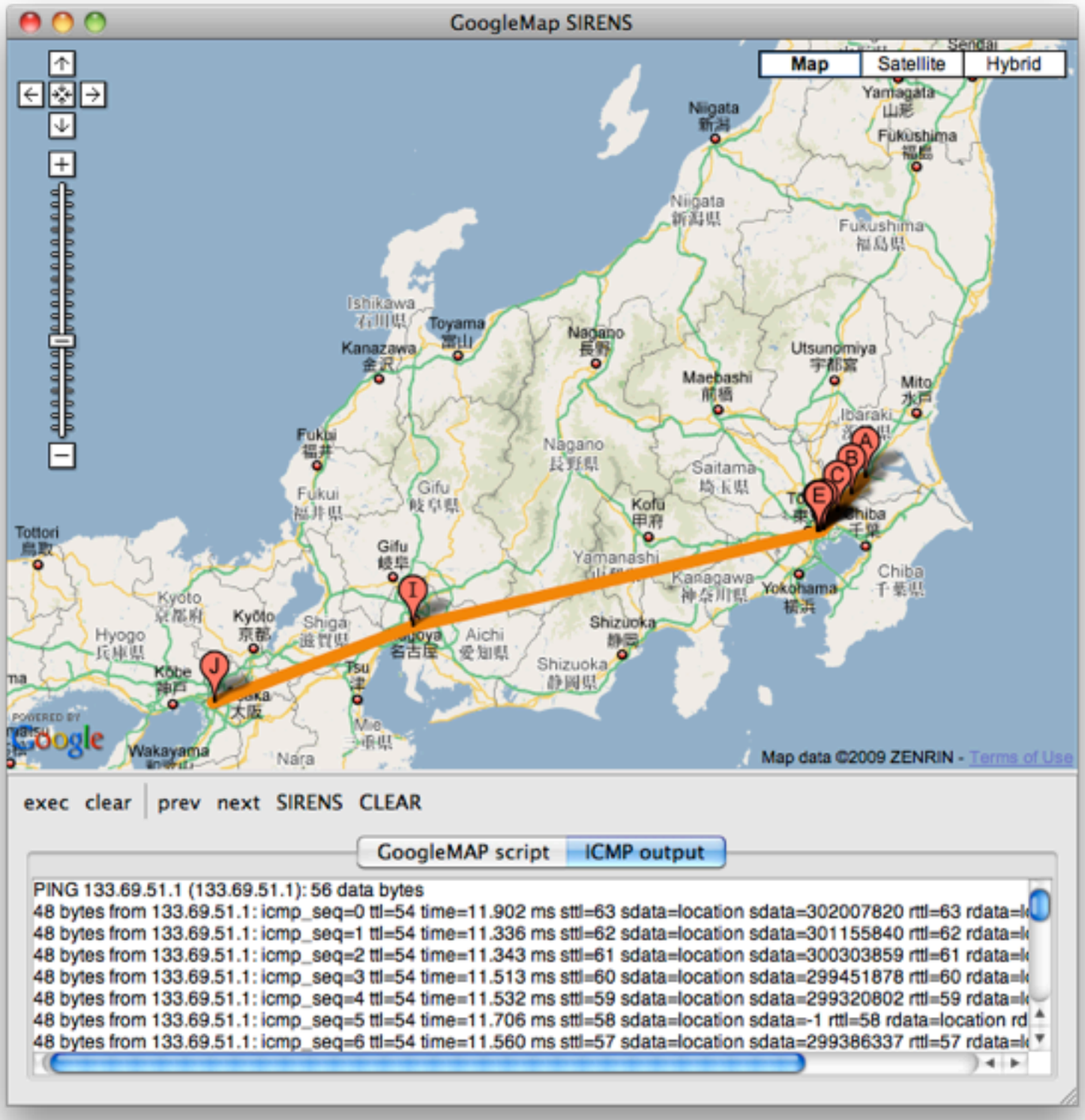
実装、展開状況

- 拡張 SIRENS を FreeBSD 上に実装
  - ルータ機能:基盤情報のキャッシュ・代理応答機構
  - 端末機能:TCP socket API, ICMP
  - C, C++, Java, Pyhon
- 展開
  - 実験室内：13 台のテストベッドを秋葉原、つくばに構築
  - JGN2+：大手町、堂島の物理ホストにルータを展開

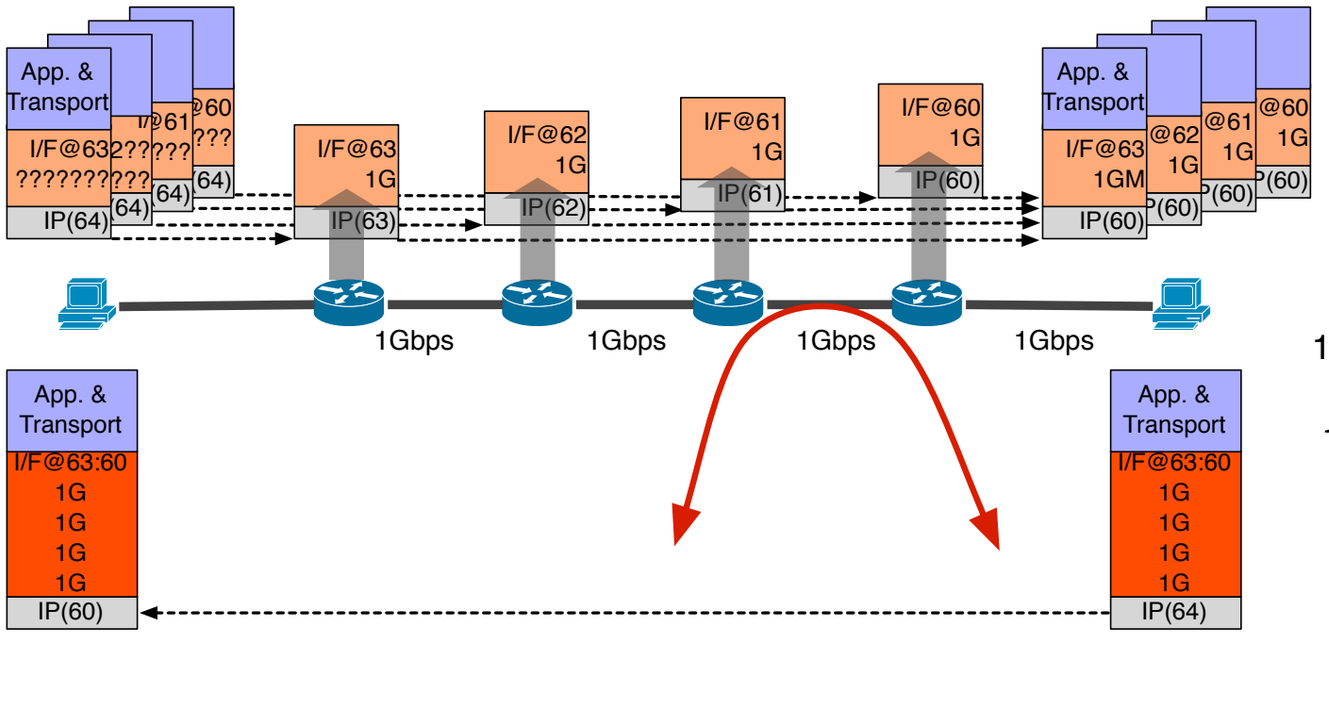


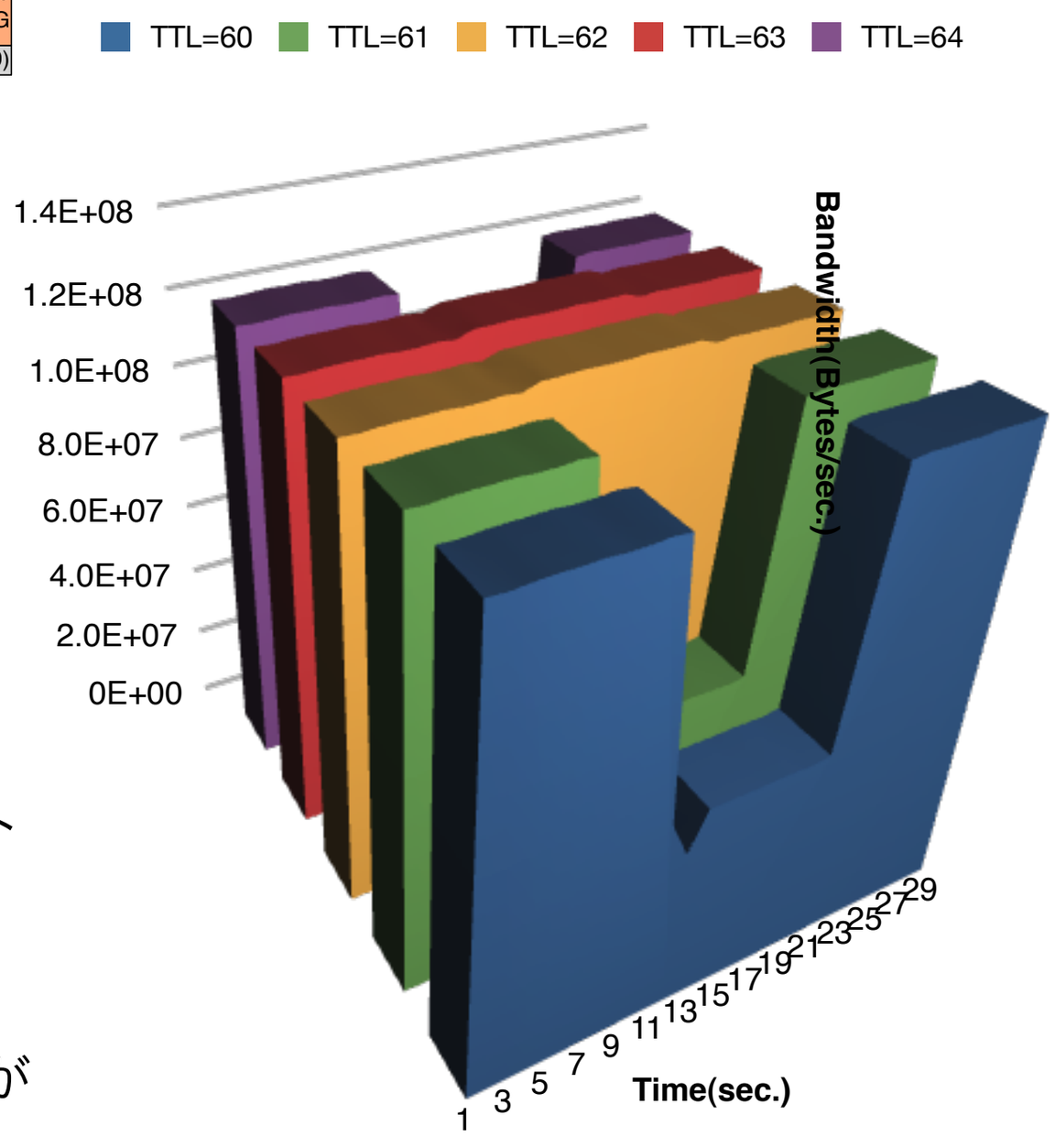
『可視化』した例 (1)

- JGN2+ に展開したサーバ＋産総研のテストベッド
- ICMP echo/reply にクロスレイヤ層を重畳



『可視化』した例





- シナリオ
  1. 端末 + ルータ 4hop の両端で iperf で通信
  2. 端末は、各hop のルータの i/f 通過パケット数を観測、インターフェース（リンクの）利用帯域を計算
  3. 10-20 秒間に 3-4hop 間に別トラフィックが発生

展開にあたっての問題

- ISP は情報開示に否定的
  - 競合他社との関係、技術的制約、セキュリティ
- Future Internet でも状況は同じか？
  - 透明性の提供は世の流れ
    - ISP、利用者の情報は非対称
  - Provisioning が実現すれば「可視化」不要
    - Intserv 普及につながる技術革新は？
    - 提供パスの検証手段として提供

まとめ

- 「端末」にとってのネットワーク「可視化」
  - ➡Knowledge Plane における “sensor” 機能の実現
- in-band クロスレイヤ方式を拡張
- 端末、ISP の開示ポリシのすりあわせ

謝辞

- NICT 委託研究『新世代ネットワークサービス基盤構築技術に関する研究開発』として実施した。

『可視化』とend-to-end 原理

- end-to-end 原理は端末側の性能向上が背景
  - 電話からコンピュータへ
  - 『可視化』によって端末側に情報が提供されれば、振る舞いの最適化余地は大きい
    - end-to-end 原理：端末側で実現不可能な機能をネットワークで実現することを禁止していない
- ネットワークに何かを頼むのではなく、なにができるかを聞く
  ネットワーク内部の『見える化』＝『可視化』が不可欠
- Knowledge Plane[*] に必要な、“sensor”, “actuator” のうち “sensor” を提供

Preliminary

- Policy
  - Alice & Bob allow to disclose beyond 3rd hop routers'.
- Implementation
  - Alice sends req. & One Time Pad (OTP) key pairs for all routers.
  - Original OTP key is overwritten by ciphertext at designated router.
  - Bob does not send back ciphertexts ciphered by 6th & 7th.
  - Alice does not send keys for 1st & 2nd.
- Result:
  - Alice obtains 1-5 hops' data.
  - Bob obtains 3-7 hops' data